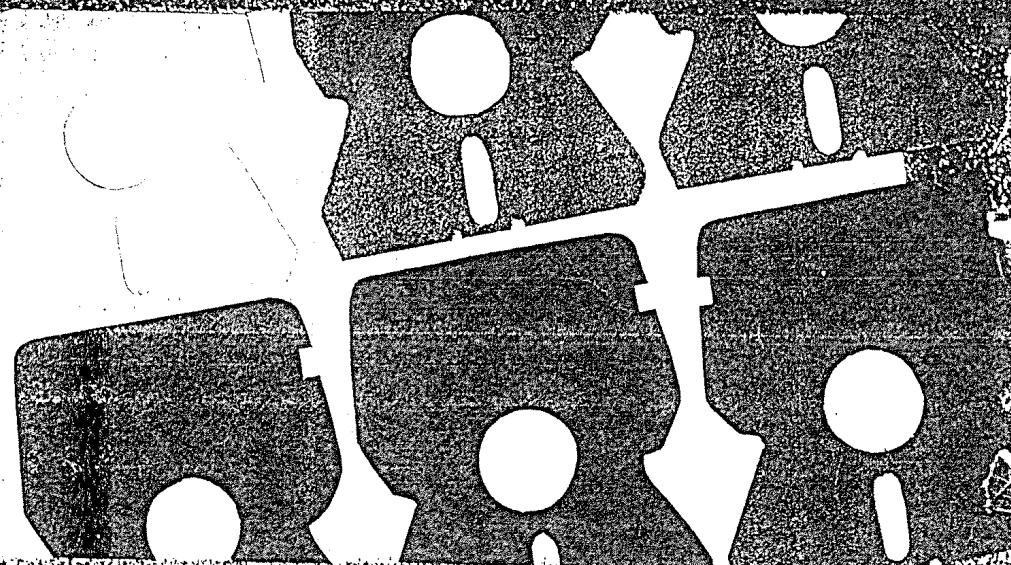


SISTEMAS CONTROLADOS POR COMPUTADOR



Karl J. Astrom
Björn Wittenmark

PARANINFO

Sistemas
controlados
por
Computador

*Karl J. Åström
Björn Wittenmark*

1988



MADRID

A Bia y Karin

Traducido por:

PABLO GONZALEZ DE SANTOS
Instituto de Automática Industrial (C.S.I.C.)

Revisión técnica:

JOSE NO SANCHEZ DE LEON
Instituto de Automática Industrial (C.S.I.C.)

© PRENTICE-HALL, INC., Englewood Cliffs, N.J., U.S.A.
© de la edición española, Editorial Paraninfo, S. A. Madrid (España)
© de la traducción española, Editorial Paraninfo, S. A. Madrid (España)

Título original:

COMPUTER CONTROLLED SYSTEMS. Theory and Design.

Edición original en lengua inglesa
publicada por Prentice-Hall, Inc.

Reservados los derechos para todos los países de lengua española. Ninguna parte de esta publicación, incluido el diseño de la cubierta, puede ser reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea éste electrónico, químico, mecánico, electro-óptico, grabación, fotocopia o cualquier otro, sin la previa autorización escrita por parte de la Editorial.

Printed in Spain
Impreso en España

ISBN: 0-13-164319-3 (edición americana)
ISBN: 84-283-1593-0 (edición española)

Depósito Legal: M.5029-1988



Magallanes, 25 - 28015 MADRID (02209/37/25)

Iberica Grafic, S.A., C/ Atienza, 1 y 3 Fuenlabrada (Madrid)

Indice

PRESENTACION.....	11
1 CONTROL CON COMPUTADOR	15
1.1. Introducción	15
1.2. Tecnología de los computadores.....	16
1.3. Teoría de control con computador.....	23
1.4. Sistemas inherentemente muestrados.....	27
1.5. Cómo se desarrolló la teoría.....	30
1.6. Referencias.....	33
2 MUESTREO DE SEÑALES CONTINUAS	36
2.1. Introducción	36
2.2. Descripción de los mecanismos de muestreo	37
2.3. El teorema de muestreo	37
2.4. Reconstrucción	40
2.5. Enmascaramiento o solapamiento de frecuencias.....	42
2.6. Aspectos prácticos de la elección del periodo de muestreo	47
2.7. Resumen	49
2.8. Problemas	49
2.9. Referencias.....	50
3 MODELOS MATEMATICOS ORIENTADOS AL COMPUTADOR: SISTEMAS DISCRETOS	52
3.1. Introducción	52
3.2. Muestreo de un sistema continuo en el espacio de estados	53
3.3. Transformación de los modelos del espacio de estados	63
3.4. Modelos entrada-salida	66
3.5. La transformada Z	73
3.6. Polos y ceros.....	78
3.7. Selección de la frecuencia de muestreo.....	83
3.8. Problemas	84
3.9. Referencias.....	88

4 MODELOS ORIENTADOS AL PROCESO	90	
4.1. Introducción	90	
4.2. Sistema controlado con computador	91	
4.3. El modelo de modulación	92	
4.4. Respuesta en frecuencia	98	
4.5. Formalismo de la función de transferencia discreta	105	
4.6. Muestreo multiperiódico	114	
4.7. Problemas	118	
4.8. Referencias	118	
5 ANALISIS DE SISTEMAS DISCRETOS	120	
5.1. Introducción	120	
5.2. Estabilidad	121	
5.3. Controlabilidad, alcanzabilidad y observabilidad	132	
5.4. Análisis de lazos de realimentación sencillos	140	
5.5. Problemas	147	
5.6. Referencias	150	
6 MODELOS DE PERTURBACIONES	152	
6.1. Introducción	152	
6.2. Reducción de los efectos de las perturbaciones	153	
6.3. Los modelos clásicos de perturbaciones	156	
6.4. Perturbaciones determinadas a trozos	159	
6.5. Modelos estocásticos de perturbaciones	162	
6.6. Procesos estocásticos continuos	181	
6.7. Muestreo de una ecuación diferencial estocástica	187	
6.8. Conclusiones	187	
6.9. Problemas	188	
6.10. Referencias	190	
7 DISEÑO: VISION DE CONJUNTO	192	
7.1. Introducción	192	
7.2. Aspectos operacionales	194	
7.3. Principios de estructuración	197	
7.4. Aproximación descendente	198	
7.5. Aproximación ascendente	201	
7.6. Diseño de lazos sencillos	204	
7.7. Conclusiones	209	
7.8. Problemas	210	
7.9. Referencias	210	
8 CONVERSION DEL DISEÑO ANALOGICO	212	
8.1. Introducción	212	
8.2. Métodos alternativos para aproximaciones	213	
8.3. Controladores digitales PID	218	
8.4. Rediseño de la realimentación de estado	229	
8.5. Métodos de diseño basados en la respuesta de frecuencia	232	
8.6. Conclusión	234	
8.7. Problemas	234	
8.8. Referencias	236	
9 METODOS DE DISEÑO EN EL ESPACIO DE ESTADOS	238	
9.1. Introducción	238	
9.2. Regulación basada en la asignación de polos por realimentación del estado	238	
9.3. Observadores	249	
9.4. Realimentación de la salida	255	
9.5. El problema del servo	258	
9.6. Conclusiones	262	
9.7. Problemas	263	
9.8. Referencias	265	
10 DISEÑO POR ASIGNACION DE POLOS BASADO EN MODELOS DE ENTRADA-SALIDA	266	
10.1. Introducción	266	
10.2. Formulación del problema	266	
10.3. Solución	270	
10.4. Un problema algebraico	273	
10.5. Un procedimiento de diseño	275	
10.6. Sensibilidad a los errores de modelado	282	
10.7. Relación con otros métodos de diseño	284	
10.8. Aspectos prácticos	288	
10.9. Un ejemplo de diseño	293	
10.10. Conclusiones	300	
10.11. Problemas	301	
10.12. Referencias	303	
11 METODOS DE DISEÑO OPTIMO: FORMULACION EN EL ESPACIO DE ESTADOS	304	
11.1. Introducción	304	
11.2. Control lineal cuadrático	309	
11.3. Teoría de predicción y filtrado	320	
11.4. Control lineal cuadrático Gaussiano	326	
11.5. Aspectos prácticos	328	
11.6. Conclusiones	330	
11.7. Problemas	330	
11.8. Referencias	333	
12 METODOS DE DISEÑO OPTIMO: FORMULACION ENTRADA-SALIDA	335	
12.1. Introducción	335	
12.2. Formulación del problema	336	
12.3. Predicción óptima	340	
12.4. Control de varianza mínima	346	

12.5. Control LQG	355
12.6. Aspectos prácticos	366
12.7. Conclusiones	376
12.8. Problemas	377
12.9. Referencias	381
IDENTIFICACION	383
13.1. Introducción	383
13.2. Construcción matemática del modelo	384
13.3. Identificación de sistemas	385
13.4. El principio de mínimos cuadrados	388
13.5. Cálculos recursivos	392
13.6. Ejemplos	399
13.7. Resumen	401
13.8. Problemas	401
13.9. Referencias	403
CONTROL ADAPTATIVO	404
14.1. Introducción	404
14.2. Control auto-sintonizado	405
14.3. Análisis	408
14.4. Otros métodos de control adaptativo	413
14.5. Usos y abusos de las técnicas adaptativas	417
14.6. Conclusiones	420
14.7. Problemas	421
14.8. Referencias	422
REALIZACION PRACTICA DE CONTROLADORES DIGITALES	424
15.1. Introducción	424
15.2. Visión de conjunto	425
15.3. Prefiltrado y retardo de cálculo	427
15.4. Actuadores no lineales	432
15.5. Aspectos operacionales	438
15.6. Aspectos numéricos	442
15.7. Programación	454
15.8. Conclusiones	460
15.9. Problemas	461
15.10. Referencias	465
EJEMPLOS	468
FUNCIONES MATRICIALES	473
SIMNON: UN PROGRAMA DE SIMULACION INTERACTIVO	476
DEMOSTRACION DEL TEOREMA 10.1	490
INDICE ALFABETICO	493

Presentación

En los años sesenta un ingeniero de control tenía que dominar la tecnología del cálculo analógico debido a que se empleaba en los computadores analógicos para la simulación de sistemas de control y como la herramienta más importante para los cálculos. En los sistemas de control también se utilizaban las tecnologías mecánica, neumática y electrónica de tipo analógico.

La situación está cambiando rápidamente en estos momentos debido a los espectaculares desarrollos habidos en los computadores digitales y la microelectrónica. Los computadores digitales se emplearon originalmente como componentes de complejos sistemas de control de procesos. Sin embargo, debido a su pequeño tamaño y bajo precio, los computadores digitales están siendo utilizados en este momento en los reguladores para el control de lazos individuales. En diversas áreas, los computadores digitales están ya mejorando en rendimiento a sus equivalentes analógicos y también son más baratos.

Los computadores digitales se están empleando también cada vez más como herramientas para el análisis y el diseño de sistemas de control. Así el ingeniero de control dispone hoy día de herramientas mucho más poderosas que antes. Los computadores digitales todavía están en un estadio de rápido desarrollo a causa del progreso en la tecnología VLSI [Very Large Scale Integration (Integración a muy gran escala)]. Por todo ello se pueden esperar en el futuro mejoras tecnológicas sustanciales.

A causa de todos estos desarrollos está cambiando drásticamente el planteamiento del análisis, diseño y realización de sistemas de control. Originalmente se trataba solamente de «traducir» los diseños analógicos anteriores a la nueva tecnología. Sin embargo, poco a poco se ha ido cayendo en la cuenta de que se podía avanzar mucho si se explotaba todo el potencial de la nueva tecnología. Afortunadamente, la teoría de control se ha desarrollado también de una forma sustancial durante los últimos 30 años. Durante un tiempo era poco realista realizar el tipo de reguladores que producía la nueva teoría excepto en unos pocos casos, más bien raros; la mayor parte en la

industria aeroespacial o en control avanzado de procesos. Sin embargo, debido al revolucionario desarrollo de la microelectrónica, pueden realizarse reguladores avanzados incluso para aplicaciones elementales. Y también es posible llevar a cabo el diseño y análisis con un coste razonable mediante las herramientas de diseño interactivo que se encuentran disponibles cada día en mayor cantidad.

El propósito de este libro es presentar la teoría de control que está relacionada con el análisis y diseño de sistemas controlados con computador, haciendo énfasis sobre los conceptos e ideas básicas. Se supone que se dispone de un computador digital con un software adecuado para los cálculos y simulaciones de forma que se le puedan dejar al computador los detalles más tediosos. El diseño de sistemas de control se realiza hasta la etapa de instalación en forma de programas de computador en un lenguaje de alto nivel.

El libro está organizado de la siguiente manera; En el Capítulo 1 se da una visión general del desarrollo del control con computador. También se expone una recopilación de los desarrollos de la teoría para proporcionar una cierta perspectiva. (Aquellos que ignoran la historia se arriesgan a repetirla.)

El muestreo, que es una propiedad fundamental de los sistemas controlados por computador, es el objeto del Capítulo 2. Los modelos matemáticos básicos necesarios se tratan en los Capítulos 3, 4 y 6. El Capítulo 3 presenta los modelos tal como los ve el computador, mientras que el Capítulo 4 los trata desde el punto de vista del proceso. Sin perturbaciones no existen problemas de control; es importante entonces encontrar formas adecuadas para caracterizar las perturbaciones; esto se lleva a cabo en el Capítulo 6.

En el Capítulo 5 se exponen las herramientas más importantes para el análisis y simulación. La simulación juega un papel muy importante porque existen muchas preguntas concretas que son muy difíciles de contestar utilizando solamente el análisis. El lenguaje de simulación interactiva Simon, que se emplea a lo largo de todo el libro, se expone en un apéndice; sus programas no son difíciles de traducir a otros lenguajes de simulación. El hecho de poder disponer de una herramienta de simulación potente produce un cambio drástico en las actitudes y las técnicas. Es muy importante que las simulaciones vayan acompañadas por un análisis que pueda dar una estimación del orden de magnitud para asegurarse de que los resultados de la simulación son razonables. Al mismo tiempo no es necesario disponer de herramientas para realizar cálculos muy exactos, puesto que el computador lo puede hacer muy fácilmente. Los Capítulos 7 al 12 están dedicados al problema del diseño. El Capítulo 7 presenta una visión de conjunto. En el 8 se expone la conversión de los métodos analógicos de diseño. El Capítulo 9 presenta las técnicas de diseño en el espacio de estados para sistemas deterministas basados en la asignación de polos y el 10 aborda el mismo problema pero empleando modelos entrada-salida. Los Capítulos 11 y 12 tratan los métodos de diseño óptimo del control basados en el filtrado de Kalman, el control lineal cuadrático y lineal cuadrático Gaussiano, el primero de ellos por medio de modelos en el espacio de estados y el segundo empleando modelos entrada-salida.

Un rasgo característico de muchos de los nuevos métodos de diseño es que necesitan un modelo del proceso y de sus perturbaciones. En el Capítulo 13 se expone cómo pueden obtenerse estos modelos. El Capítulo 14 trata de forma breve los sistemas de control adaptativos en los parámetros. Esto puede también considerarse como una combinación de los métodos de diseño de los Capítulos 9 al 12 y de los métodos de identificación recursiva del Capítulo 14. El Capítulo 15 aborda diferentes aspectos de la implementación de los algoritmos de control con computador.

La teoría se ha organizado de una forma tal que todos los modelos y las especificaciones se dan en forma continua. Esto hace más fácil las aplicaciones por su proximidad a la física. Los sistemas multivariados se tratan siempre que se emplean las técnicas del espacio de estados; sin embargo, el tratamiento de los modelos entrada-salida empleando la aproximación polinomial se limita al caso de una entrada-salida. En ambas situaciones se abordan los aspectos tanto determinísticos como estocástico del análisis y del problema del diseño.

Cuando se diseña un sistema, suele ser muy provechoso enfocar un problema desde varios puntos de vista. Ya que el objetivo del libro es proporcionar unos buenos fundamentos para el diseño de sistemas controlados por computador, es necesario abordar una amplia gama de temas. Se ha buscado un equilibrio razonable entre los detalles y la visión de conjunto; sin embargo, los Capítulos 6, 11, 12, 13 y 14 requieren libros completos para abordar cada tema en detalle.

Suele ser costumbre en la teoría de sistemas muestreados emplear el mismo símbolo z para indicar la variable compleja y el operador de desplazamiento hacia adelante. Creemos que esta costumbre puede confundir a los estudiantes y, por tanto, hemos introducido el símbolo q para representar al operador de desplazamiento hacia adelante. Esto es similar al empleo de s como variable compleja y $p = d/dt$ como un operador diferencial para los sistemas continuos. La notación q^{-1} se emplea para el operador de desplazamiento hacia atrás.

El libro puede emplearse de muchas formas diferentes. Los Capítulos 2, 3, 5, 7, 8, 9, 10 y 15 y las Secciones 6.1 a 6.3 son adecuadas para un curso de introducción a los sistemas muestreados. Una representación detallada de los Capítulos 4, 6, 7 y del 9 al 15 puede servir de núcleo para un curso de especialización en diseño de sistemas controlados por computador. Hemos dado cursos a personal de las industrias basándonos en los Capítulos 3, 4, 5, 8, 9, 10, 13, 14 y 15. En todos los casos nos ha sido muy útil poder usar la simulación por computador y complementar las clases y ejercicios con experiencias de laboratorio. En el manual de soluciones se dan algunas sugerencias sobre este tema.

Agradecimientos

Durante la redacción de este libro hemos tenido el placer y el privilegio de estar en contacto con muchas personas. Es un especial placer poder expresar nuestro agradecimiento a uno de los padres de la teoría de los sistemas muestreados, profesor

E. I. Jury, quien ha leído pacientemente varias versiones del manuscrito y nos ha dado muchos consejos y sugerencias. También damos las gracias a Per Hagander del Lund Institute of Technology, George Axelby, editor de *Automatica*, y a Brian Anderson de la Australian National University, Canberra; cada uno de los cuales nos ha hecho críticas muy útiles. Queremos también dejar constancia de la ayuda recibida de los profesores Paul Houp del MIT, Rick Johnson de Cornell y Howard Elliot de la Universidad de Massachusetts, quienes han utilizado el manuscrito en sus clases. Estamos en deuda con muchos estudiantes que nos han hecho comentarios muy útiles de diferentes versiones del manuscrito. También queremos hacer extensiva nuestra gratitud a Eva Dagnegård y Agneta Tusynski por el mecanografiado de las diferentes versiones del manuscrito y a Doris Nilsson por la realización de las ilustraciones.

KARL J. ÅSTRÖM
BJÖRN WITTENMARK

Control con computador

Objetivo:

Introducir el tema y presentar algunos fundamentos históricos sobre el desarrollo de la tecnología y de la teoría del control con computador.

1.1. Introducción

Los computadores digitales están siendo usados cada vez más para realizar sistemas de control. Por tanto, es importante comprender bien los sistemas controlados con computador. Se pueden ver los sistemas controlados por computador como aproximaciones de los sistemas controlados analógicamente, pero éste es un tratamiento pobre, ya que no se emplea todo el potencial del control con computador. En el mejor de los casos los resultados serán solamente tan buenos como los obtenidos con el control analógico. Otra solución es estudiar los sistemas controlados con computador, de forma que se explote todo el potencial del control con computador. El principal objetivo de este libro es proporcionar los conocimientos básicos que esto requiere.

En la Figura 1.1 se describe de forma esquemática un sistema controlado por computador. La salida del proceso $y(t)$ es una señal continua. Dicha salida se pasa a forma digital mediante un convertidor analógico-digital (A-D). El convertidor A-D puede incluirse en el computador o considerarse como una unidad independiente, según la preferencia de cada uno. La conversión se realiza en los instantes de muestreo, t_k . El computador interpreta las señales ya convertidas, $\{y(t_k)\}$, como una secuencia de números, procesa las medidas por medio de un algoritmo y genera una nueva secuencia de números, $\{u(t_k)\}$. Un convertidor digital-analógico (D-A) se encarga de convertir esta secuencia en una señal analógica. Nótese que el sistema funciona en lazo abierto en el intervalo entre las conversiones A-D y D-A. Los sucesos se sincronizan por medio del reloj de tiempo real del computador. El computador

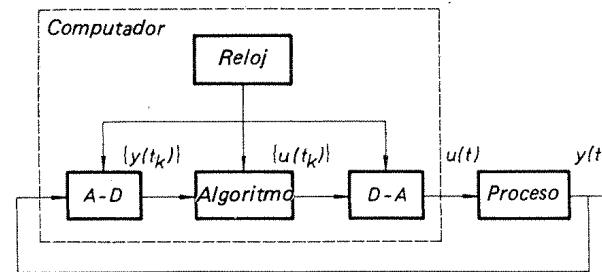


Figura 1.1 Diagrama esquemático de un sistema controlado con computador.

digital opera de forma secuencial en el tiempo y cada operación dura un cierto tiempo. Sin embargo, el convertidor D-A debe producir una señal continua. Esto se realiza normalmente manteniendo constante la señal de control entre las conversiones. El sistema de control por computador contiene señales continuas y *muestreadas*, o *discretas*. Estos sistemas se han llamado tradicionalmente *sistemas de datos muestreados*, o simplemente *sistemas muestreados*; utilizaremos aquí este término como sinónimo de *sistemas controlados por computador*.

La mezcla de diferentes tipos de señal a veces causa dificultades. Sin embargo, en la mayor parte de los casos es suficiente con describir el comportamiento del sistema en los instantes de muestreo. Sólo interesan entonces las señales en instantes discretos. A los sistemas de este tipo les denominaremos *sistemas discretos*. Los sistemas discretos tratan secuencias de números, por lo que una forma natural de representarlos es mediante ecuaciones en diferencias.

El propósito del libro es presentar la teoría de control necesaria para el análisis y el diseño de los sistemas controlados por computador. Este capítulo proporciona algunos conocimientos básicos. En la Sección 1.2 se presenta una breve visión de conjunto de la tecnología de los computadores. En la Sección 1.3 se expone la necesidad de una teoría apropiada. La comprobación de que los sistemas de control por computador no pueden comprenderse *totalmente* con la teoría de los sistemas lineales invariantes y continuos en el tiempo se hace mediante ejemplos. Un ejemplo muestra no sólo que los sistemas de control por computador pueden diseñarse empleando la teoría de los sistemas continuos y aproximaciones, sino que pueden conseguirse mejoras sustanciales con otras técnicas que emplean todo el potencial del control por computador. La Sección 1.4 presenta algunos ejemplos de sistemas que son en sí mismos muestreados y en la Sección 1.5 se hace un bosquejo del desarrollo de la teoría de los sistemas muestreados.

1.2. Tecnología de los computadores

La idea de utilizar computadores digitales como componentes de sistemas de control surgió alrededor de 1950. Las primeras aplicaciones ensayadas fueron el

control de misiles y de aviones. Los estudios mostraron que no era posible utilizar los computadores digitales disponibles por aquel entonces. Los computadores eran demasiado grandes, consumían demasiada potencia y no eran suficientemente fiables. Por esta razón se desarrollaron computadores de propósito especial —analizadores diferenciales digitales (DDA, de Digital Differential Analyzers)— para las primeras aplicaciones aeroespaciales.

Los desarrollos más importantes en el control por computador se produjeron en las industrias de proceso. El progreso de estos desarrollos aparece en la Figura 1.2, que muestra el crecimiento del número de computadores usados en control de procesos durante los últimos 25 años.

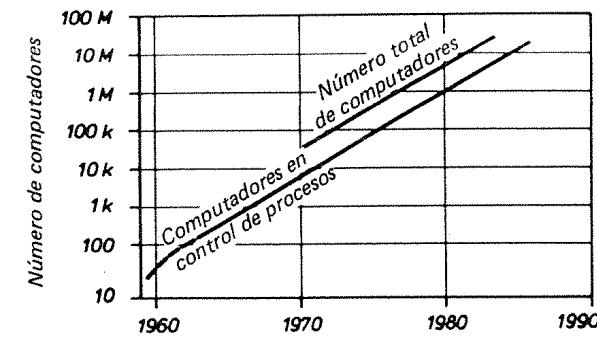


Figura 1.2 Crecimiento del número de computadores empleados para el control de procesos industriales. Para poder comparar se da también el número total de computadores. El gráfico se ha obtenido de varias fuentes diferentes: Control Engineering, A. D. Little, Frost y Sullivan y Diebold. (Dibujo rehecho con autorización a partir de los datos publicados en Control Engineering, 1980, Technical Publishing Co.)

Por otra parte, la idea de emplear computadores digitales para el control de procesos surgió a mitad de los años cincuenta. El trabajo en serio comenzó en marzo de 1956 cuando la compañía aeroespacial Thomson Ramo Woolridge (TRW) contactó con Texaco para realizar un estudio de factibilidad. Después de las conversaciones preliminares se decidió probar en una unidad de polimerización de la refinería de Port Arthur, Texas. Un grupo de ingenieros de TRW y Texaco hicieron un estudio de factibilidad completo, que requirió alrededor de 30 hombres/año. Se diseñó un sistema de control por computadora para la unidad de polimerización basado en un computador RW-300. El sistema de control se puso en marcha el 12 de marzo de 1959. El sistema controlaba 26 flujos, 72 temperaturas, 3 presiones y 3 composiciones. Las funciones esenciales eran minimizar la presión del reactor, determinar la distribución óptima entre la alimentación de los cinco reactores, controlar la entrada de agua caliente a partir de la actividad catalítica y determinar la recirculación óptima.

Muchos fabricantes de computadores, que vieron un gran mercado potencial para sus productos, tomaron nota del trabajo de pioneros realizado por TRW. Se iniciaron muchos estudios de factibilidad diferentes y comenzó un desarrollo muy activo. El resultado de estos esfuerzos se refleja en el crecimiento que presenta la Figura 1.2.

Para tratar este tremendo desarrollo es útil considerar cuatro períodos.

Período de los pioneros	≈ 1955
Período de control digital directo	≈ 1962
Período del minicomputador	≈ 1967
Período del microcomputador	≈ 1972

Es difícil dar fechas precisas porque el desarrollo estuvo muy diversificado. Hubo una gran diferencia entre diferentes áreas de aplicación y diferentes industrias; también hubo un considerable solapamiento. Las fechas dadas se refieren a la primera aparición de nuevas ideas.

El periodo de los pioneros

El trabajo hecho por TRW y Texaco hizo surgir un interés muy grande en las industrias de procesos, entre los fabricantes de computadores y en las organizaciones de investigación. Las industrias vieron una posible herramienta para una mayor automatización, las industrias de computadores vieron nuevos mercados y las universidades un nuevo campo de investigación. Los fabricantes de computadores iniciaron muchos estudios de factibilidad debido a que estaban ansiosos de aprender la nueva tecnología y muy interesados en conocer cómo debería ser un control de procesos por computador adecuado. Los estudios de factibilidad continuaron durante los años 60.

Los sistemas de computador que se utilizaban entonces eran lentos, caros y poco fiables. Los sistemas primitivos empleaban válvulas de vacío. Datos típicos de un computador de alrededor de 1958 son: un tiempo de suma de 1 ms, un tiempo de multiplicación de 20 ms y un tiempo medio entre fallos (MTBF, de Mean Time Between Failures) de entre 50 y 100 horas para la unidad central. Para aprovechar al máximo unos computadores tan caros, era necesario encomendarles la ejecución de muchas tareas. Como tenían tan poca fiabilidad, controlaban el proceso imprimiendo instrucciones al operador del proceso o cambiando las consignas de los regulares analógicos. A estas formas de operación de supervisión se les conocía como *guía del operador* y *control de consigna*.

Las tareas más importantes del computador eran encontrar las condiciones de operación óptimas, realizar la organización y la planificación de la producción y elaborar informes sobre la producción y el consumo de materias primas. El problema de encontrar las mejores condiciones de operación se planteaba como un problema de optimización estática. Para realizar la optimización hacían falta los modelos matemáticos de los procesos. Los modelos que se empleaban —que eran bastante complicados— se obtenían de los modelos físicos de los procesos y del análisis de regresión de los datos del proceso. También se llevaron a cabo intentos de realizar una optimización en línea.

A menudo los avances se veían entorpecidos por una falta de conocimientos de proceso. También fue quedando claro que no era suficiente con ver simplemente los

problemas como cuestiones de optimización estática; hacían falta modelos dinámicos. Una parte significativa de los estudios de factibilidad estaban dedicados al modelado, lo que consumía bastante tiempo debido a la falta de una buena metodología de modelado. Esto estimuló las investigaciones sobre los métodos de identificación de sistemas.

Durante los estudios de factibilidad se adquirió una gran experiencia. Se vio claramente que el control de procesos impone a los computadores unas exigencias especiales. La necesidad de una rápida respuesta a las peticiones del proceso llevó a desarrollar la *característica de las interrupciones*, que consiste en un dispositivo físico especial que permite a un acontecimiento exterior interrumpir al computador en el trabajo que está haciendo para que pueda responder a tareas del proceso que son más urgentes. Muchos de los sensores que se necesitaban no estaban disponibles. También había dificultades de diferente tipo para intentar introducir una tecnología nueva en industrias anticuadas.

Se seguían muy de cerca los progresos en las conferencias y reuniones y en las revistas. La revista *Control Engineering* publicó una serie de artículos describiendo el empleo de los computadores en control de procesos. En marzo de 1961 se habían instalado 37 sistemas. Un año más tarde el número de sistemas había crecido hasta 159. Las aplicaciones comprendían el control de acerías, industrias químicas y plantas de generación de energía eléctrica. El desarrollo prosiguió a diferente ritmo en las distintas industrias. Los estudios de factibilidad continuaron a lo largo de los años 60 y los 70.

El control digital directo

Las primitivas instalaciones de computadores de control operaban en modo supervisor, bien como guía de operador, bien como control de consigna. En ambos casos hacía falta el equipo de control analógico ordinario. Imperial Chemical Industries (ICI) en Inglaterra se apartó drásticamente de este enfoque en 1962. Se sustituyó la totalidad de la instrumentación analógica de control de procesos por un computador, un Argus de Ferranti. El computador media 224 variables y controlaba 129 válvulas directamente. Era el comienzo de una nueva era en el control de procesos: simplemente la tecnología analógica se sustituía directamente por la tecnología digital; la función del sistema era la misma. Se acuñó el nombre *Control Digital Directo* (DDC, del inglés Direct Digital Control) para destacar que el computador controlaba el proceso directamente. En 1962 un típico computador de control de procesos podía sumar dos números en $100 \mu s$ y multiplicarlos en 1 ms. El MTBF era de alrededor de 1.000 horas.

El coste era el argumento más fuerte para cambiar la tecnología. El coste de la tecnología analógica aumentaba linealmente con el número de lazos de control; el coste inicial de un computador digital era grande, pero el coste de añadir un lazo adicional era pequeño. El sistema digital era entonces más barato para los sistemas grandes. Otra ventaja era que la comunicación con el operador podía cambiar de forma drástica; un panel de comunicación con el operador podía sustituir a una larga



pared de instrumentos analógicos. El panel empleado en el sistema de ICI era muy sencillo: una pantalla digital y unos pocos pulsadores.

La flexibilidad era otra de las ventajas de los sistemas DDC. Los sistemas analógicos se cambiaban recableándolos; los sistemas controlados por computador lo eran reprogramándolos. La tecnología digital ofrecía también otras ventajas. Era fácil realizar interacciones entre varios lazos de control. Los parámetros de un lazo de control podían hacerse función de las condiciones de operación. La programación se simplificaba introduciendo lenguajes especiales para DDC. Un usuario de un lenguaje de este tipo no necesitaba saber nada sobre programación, sino simplemente introducir las entradas, salidas, tipos de reguladores, factores de escala y parámetros de los reguladores en unas tablas. Entonces, para el usuario, los sistemas se asemejaban a una conexión de reguladores ordinarios. Uno de los inconvenientes del sistema es que era difícil de realizar con él estrategias de control no convencionales. Esto entorpeció de forma evidente durante muchos años el desarrollo del control.

El DDC fue un cambio de dirección fundamental en el desarrollo de los sistemas de control por computador. El interés se centró en las funciones básicas de control en lugar de hacerlo en las funciones de supervisión de los sistemas anteriores. En los años 1963-65 se realizó un progreso considerable. Las especificaciones de los sistemas DDC se elaboraban conjuntamente entre los usuarios y los vendedores. Se discutían en profundidad problemas tales como la elección del período de muestreo y de los algoritmos de control, así como el problema clave de la fiabilidad. El concepto DDC se aceptó rápidamente a pesar del hecho que los sistemas DDC a menudo resultaban más caros que sus correspondientes analógicos.

El periodo de los minicomputadores

Durante los años 60 hubo un desarrollo sustancial en la tecnología de los computadores digitales. Los requisitos de los computadores de control de procesos se cumplían perfectamente gracias al progreso en la tecnología de los circuitos integrados. Los computadores resultaban más pequeños, más rápidos, más fiables y más baratos. Se elaboró el término *minicomputador* para los nuevos computadores que surgieron.

El desarrollo de la tecnología de los minicomputadores combinado con el creciente conocimiento alcanzado sobre el control de procesos con computador durante los períodos inicial y de DDC produjo un rápido incremento en las aplicaciones del control por computador. Varios fabricantes anunciaron computadores especiales de control de procesos. Un computador de procesos típico de ese período tenía una longitud de palabra de 16 bits. La memoria principal era de 8 K a 124 K palabras. Se solía emplear un disco como memoria secundaria. Un computador típico de ese período fue el CDC 1700, con un tiempo de suma de $2 \mu\text{s}$ y un tiempo de multiplicación de $7 \mu\text{s}$. El MTBF de la unidad central era de alrededor de 20.000 horas.

Un factor importante en el rápido incremento del control por computador en este período fue que el control digital con computador ahora venía en «unidades»

más pequeñas. Esto permitía utilizar el control con computador para proyectos menores y para problemas más pequeños. Debido a los minicomputadores, el número de computadores de procesos creció desde unos 5.000 en 1970 hasta alrededor de 50.000 en 1975.

Los microcomputadores

El minicomputador era todavía un sistema demasiado grande. Incluso aunque las prestaciones siguieron aumentando y los precios disminuyendo, el precio de un minicomputador en 1975 era todavía de unos 10.000 \$. Esto significa que un pequeño sistema rara vez costaba menos de 100.000 \$. El control por computador estaba todavía fuera del alcance de un gran número de problemas de control. Pero con la llegada del microcomputador en 1972, el precio de un computador en una tarjeta con las prestaciones de un minicomputador de 1975 descendió a 500 \$ en 1980. Otra consecuencia fue que la potencia de cálculo digital en 1980, comprada en cantidad, llegó a valer solamente 50 \$. Esto significa en la práctica que el control por computador podía considerarse en ese momento como una alternativa que había que tener en cuenta, sin importar lo reducido del tamaño de la aplicación.

Ya que durante los años 80 van a producirse unos desarrollos todavía más drásticos en microelectrónica con la tecnología VLSI, es una predicción segura que se producirá un gran aumento en las aplicaciones del control con computador. Los microcomputadores ya han producido un impacto apreciable sobre los equipos de control: los microcomputadores están sustituyendo a los circuitos analógicos incluso en los controladores de un solo lazo; se han construido pequeños sistemas DDC empleando microcomputadores; la comunicación con el operador se ha mejorado extraordinariamente en estos sistemas con la introducción de pantallas gráficas de color; se han construido sistemas jerarquizados de control con un gran número de microcomputadores y se han diseñado reguladores de propósito especial basados en microcomputadores.

El futuro

Basándose en los extraordinarios desarrollos habidos en el pasado se está tentado de especular sobre el futuro. Hay cuatro áreas que son importantes para el desarrollo del control de procesos por computador.

- Conocimiento del proceso.
- Tecnología de la medida.
- Tecnología de los computadores.
- Teoría de control.

El conocimiento acerca del control de procesos y de la dinámica del proceso está aumentando lenta pero continuamente. Las posibilidades de aprender sobre las características de los procesos está creciendo de forma sustancial con la instalación de sistemas de control de procesos, ya que entonces es fácil recoger datos, realizar experimentos y analizar los resultados. También han proporcionado una información muy valiosa los progresos en identificación de sistemas y análisis de datos.

Sin embargo, es muy difícil predecir los progresos en tecnología de la medida. Empleando las técnicas existentes se pueden hacer muchas cosas. Es muy interesante la posibilidad de combinar las salidas de diferentes sensores mediante modelos matemáticos. También es posible realizar la calibración automática mediante un computador. La aparición de nuevos sensores ofrecerá, sin embargo, nuevas posibilidades.

En tecnología de los computadores se esperan desarrollos espectaculares con la introducción de los circuitos VLSI. La relación de precio a prestaciones descenderá de forma sustancial. Se espera que los microcomputadores de finales de los 80 tengan una potencia de cálculo mayor que la de los grandes computadores de finales de los 70. También se esperan mejoras sustanciales en las técnicas de presentación y en las comunicaciones.

La programación ha sido hasta ahora uno de los cuellos de botella. Desde 1950 a 1970 sólo hubo mejoras marginales en la productividad en programación. A finales de los 70 muchos sistemas de control por computador se programaban todavía en ensamblador. En el campo de control por computador, ha sido habitual solucionar alguno de los problemas de programación mediante tablas. A los usuarios de sistemas DDC se les suministraba un paquete de programas conocido como DDC que les permitía generar un sistema DDC simplemente rellenando una tabla, con lo que se necesitaba muy poco esfuerzo para generar un sistema. La amplia utilización de estos paquetes retrasan, sin embargo, el desarrollo, ya que, aunque es muy fácil utilizar el DDC, la realización de cualquier cosa diferente exige mucho mayor esfuerzo. De esta forma sólo se empleaban los métodos muy ensayados.

Por su parte, la teoría de control tuvo un progreso sustancial durante el período 1955-70. Sin embargo, muy poco de esta teoría se ensayó en los sistemas de control por computador, incluso a pesar de que los estudios de factibilidad habían indicando que podían conseguirse mejoras muy significativas. Una razón para esto es el coste de la programación. Como ya mencionamos antes, se requiere poco esfuerzo para emplear un paquete suministrado por el vendedor. Sin embargo, supone un esfuerzo mucho mayor intentar cualquier cosa que sea diferente. Hay indicios de que se puede esperar que esta situación cambie. Los computadores personales con lenguajes interactivos como el BASIC están comenzando a utilizarse para el control de procesos. Con un lenguaje interactivo es muy fácil intentar cosas nuevas. En gran proporción, también ocurre lo mismo con aquellos que utilizan BASIC de tiempo real. Sin embargo, desgraciadamente es muy difícil escribir sistemas seguros en BASIC. Esto cambiará cuando se disponga de mejores sistemas interactivos.

Hay, por tanto, muchos signos que apuntan a desarrollos interesantes en el campo de los sistemas controlados por computador. ¡Una buena forma de prepararse es aprender la teoría que se expone en este libro!

1.3. Teoría de control con computador

En la Figura 1.1 se presenta un diagrama esquemático de un sistema controlado por computador. El sistema contiene esencialmente cinco partes: el proceso, los convertidores A-D y D-A, el algoritmo de control y el reloj. Su operación se controla por el reloj. Los instantes en que las señales medidas se convierten a forma digital se llaman *instantes de muestreo*; el tiempo entre muestreos sucesivos se llama *periodo de muestreo* y se representa por h . Normalmente se utiliza un muestreo periódico, pero por supuesto hay otras muchas posibilidades. Por ejemplo, es posible muestrear cuando las señales de salida han variado en una determinada cantidad. También es posible emplear diferentes períodos de muestreo en diferentes lazos de un sistema. Esto se conoce como *muestreo multifrecuencia* o *multiperiódico*.

La única diferencia entre un sistema controlado por computador y un sistema de realimentación analógica ordinario es que la ley de control se realiza empleando un computador digital, con lo que aumenta considerablemente la clase de leyes de control que se pueden utilizar de forma conveniente. Por ejemplo, es fácil emplear cálculos no lineales, incorporar decisiones lógicas y realizar un volumen de cálculo importante en el controlador. También pueden emplearse tablas para almacenar datos a fin de ir acumulando conocimiento acerca de las propiedades del sistema.

¿Es necesaria una teoría de los sistemas controlados con computador?

Una buena teoría debería hacer posible comprender cómo trabaja un sistema como el de la Figura 1.1 y cómo debería diseñarse. Parece claro que un sistema muestreado se comportaría como un sistema continuo si el período de muestreo fuera suficientemente pequeño. Esto es realmente cierto bajo unas hipótesis muy razonables. ¿Hace falta entonces una teoría especial para los sistemas controlados con computador?

Se usarán algunos ejemplos para mostrar que el sistema de la Figura 1.1 no puede entenderse completamente mediante la teoría de los sistemas lineales invariantes, aunque el proceso a controlar sea un sistema continuo lineal invariante.

Ejemplo 1.1 – Dependencia temporal

Supongamos que queremos realizar un compensador que es simplemente un retardo de primer orden. Este compensador puede realizarse usando una conversión A-D, un computador digital y una conversión D-A. La ecuación diferencial de primer orden se approxima por una ecuación en diferencias de primer orden. La respuesta a un salto de un sistema de este tipo es la que aparece en la Figura 1.3. La figura muestra claramente que el sistema muestreado no es invariante en el tiempo, ya que la respuesta depende del instante en que ocurre el salto. Si se retrasa la entrada, la salida también se retrasa en la misma cantidad sólo si el retraso es múltiplo del período de muestreo.

El fenómeno ilustrado en la Figura 1.3 depende del hecho de que el sistema está controlado por un reloj (compárese con la Figura 1.1). La respuesta del sistema a un

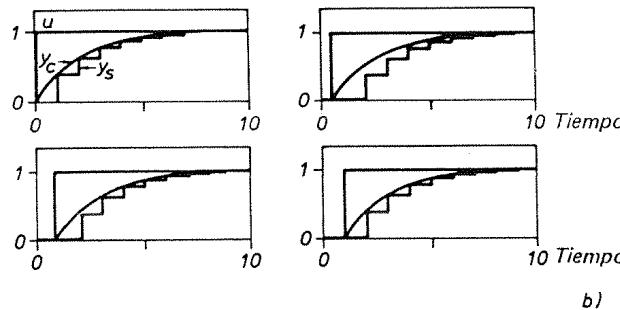
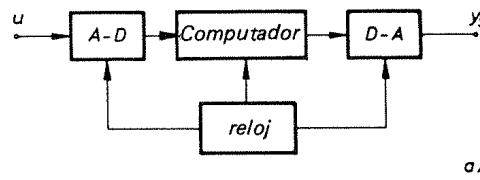


Figura 1.3 a. Diagrama de bloques de un filtro digital. b. Respuestas, y_s , a un salto de la realización en un computador digital de un retardo de primer orden para diferentes retardos en el salto de entrada comparado con el período de muestreo. Para poder comparar se da también la respuesta del sistema continuo correspondiente, y_c .

estímulo externo dependerá de cómo está sincronizado el suceso externo con el reloj interno del sistema computador.

Un sistema controlado por computador con un muestreo periódico es un *sistema periódico*. Los efectos de la periodicidad pueden hacerse arbitrariamente pequeños con sólo escoger una velocidad de muestreo suficientemente elevada; sin embargo, es necesario considerar la naturaleza periódica del sistema muestreado para comprender completamente cómo trabaja. Esto se ilustra mediante otro ejemplo.

Ejemplo 1.2 – Armónicos de orden superior

Es sabido que una señal de entrada sinusoidal aplicada a un sistema estable, lineal, continuo e invariante en el tiempo dará —después de un transitorio— unas señales de salida en el sistema que serán sinusoidales y de la misma frecuencia que la señal de entrada. La Figura 1.4 muestra lo que puede ocurrir cuando un sistema controlado por computador se somete a una excitación periódica. Se aplica una señal sinusoidal de frecuencia 4,9 Hz al sistema de la Figura 1.1. Se ve claramente que el fenómeno que se muestra en la Figura 1.4 no puede explicarse en términos de sistemas lineales invariantes. Las pulsaciones en la salida del sistema controlado por computador se deben a la interferencia entre la frecuencia de entrada y una frecuencia mayor generada durante el proceso de muestreo.

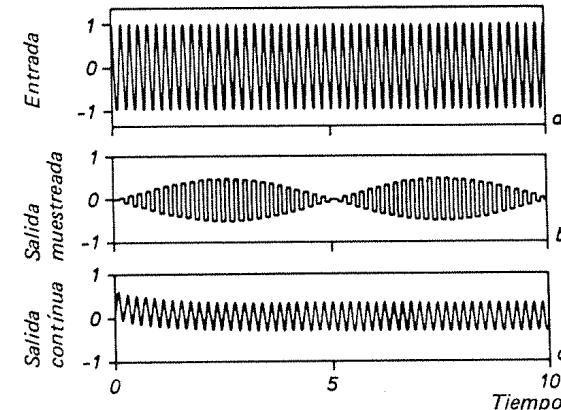


Figura 1.4 Excitación sinusoidal del sistema muestreado del Ejemplo 1.1.
a. Entrada sinusoidal con frecuencia 4,9 Hz.
b. Salida del sistema muestreado. El período de muestreo es de 0,1 s.
c. Salida del correspondiente sistema continuo.

Hay muchos aspectos de los sistemas muestreados que realmente se pueden comprender bien mediante la teoría de los sistemas lineales invariantes. Los ejemplos presentados indican, sin embargo, que los sistemas muestreados no pueden comprenderse totalmente dentro de esta estructura. Por tanto, es útil disponer de otras herramientas de análisis.

Al resolver el problema del diseño de controladores, ciertamente se puede emplear la teoría ordinaria del control continuo para diseñar la ley de control y entonces usar una aproximación discreta adecuada. Veamos esto mediante un ejemplo.

Ejemplo 1.3 – Aproximación discreta

Una planta con un integrador doble puede controlarse fácilmente por una realimentación del estado. Una forma directa de hacerlo es sencillamente calculando las ganancias de realimentación empleando la teoría continua y entonces realizar la realimentación del estado usando un control con computador. Si el período de muestreo es suficientemente pequeño, se puede esperar que el control digital tenga las mismas propiedades, desde un punto de vista práctico, que el control continuo. La Figura 1.5 muestra que esto es lo que ocurre en realidad. La concordancia con el regulador continuo puede incluso mejorarse reduciendo el período de muestreo.

Basándose en los resultados del Ejemplo 1.3 se puede estar tentado de concluir que *no* es necesaria una teoría para los sistemas muestreados. Sin embargo, esto no es correcto, ya que los sistemas controlados por computador pueden sin duda comportarse mejor que sus equivalentes continuos. ¡Esta es la principal razón por la que es útil una teoría de los sistemas muestreados! Un ejemplo aclarará esto aún más.

Ejemplo 1.4 – Control en tiempo finito

Consideremos la planta con un integrador doble del Ejemplo 1.3. En la Figura 1.6 se muestra el resultado de una realimentación del estado controlada por computador con

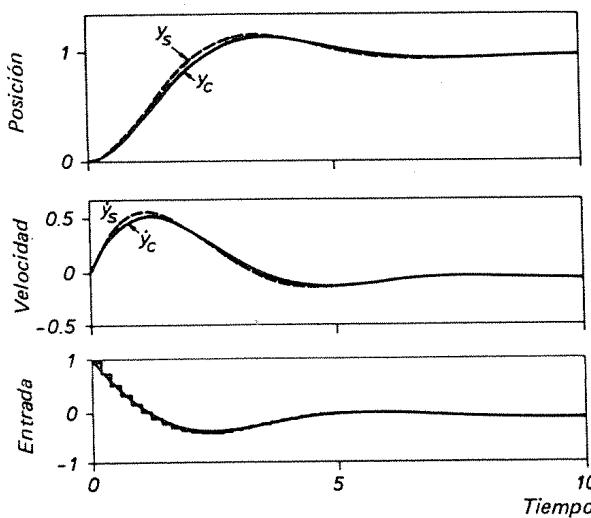


Figura 1.5 Respuesta salto de un integrador doble con realimentación del estado, solución continua y_c y aproximación muestreada y_s , cuando se utiliza un período de muestreo de 0.2 s.

una estrategia especial de realimentación lineal. La estrategia es del mismo tipo que la estrategia de control empleada en el Ejemplo 1.3, es decir, una realimentación lineal a partir de los valores muestreados de los estados. Los coeficientes de realimentación y el período de muestreo son, sin embargo, diferentes. Esta particular estrategia de control se llama control *en tiempo finito*.

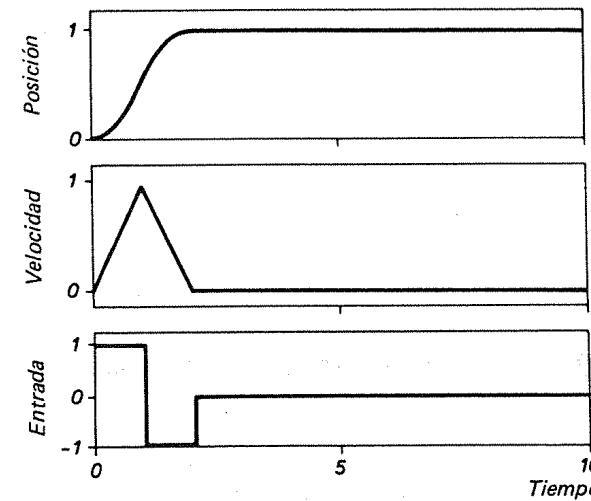


Figura 1.6 Control con computador del integrador doble con una estrategia de tiempo finito. El período de muestreo es 1 s.

Comparando las Figuras 1.5 y 1.6 se ve que el sistema de la Figura 1.6 se estabiliza más rápidamente que el sistema continuo, incluso aunque el valor máximo de las señales de control sea el mismo en los dos casos. El valor de la velocidad es, sin embargo, mayor en la Figura 1.6.

El sistema muestreado no tiene rebasamiento. Se puede observar también que las señales se estabilizan en valores constantes después de un tiempo finito. Esto no puede ocurrir en los sistemas continuos porque la solución de éstos son sumas de funciones que son productos de funciones polinómicas y exponenciales. Nótese también que el período de muestreo empleado en la Figura 1.6 es cinco veces mayor que el de la aproximación del sistema continuo de control de la Figura 1.5.

El ejemplo demuestra claramente que *incluso en el caso lineal* es posible conseguir algo mejor que la mera aproximación de un regulador continuo.

1.4. Sistemas inherentemente muestreados

Los modelos muestreados son la forma natural de describir muchos fenómenos. La teoría de los sistemas muestreados tiene, por tanto, muchas aplicaciones fuera del campo del control por computador. A continuación se dan algunos ejemplos.

Sistemas discretos como modelos de algoritmos de computador

Los algoritmos de los computadores pueden describirse mediante sistemas discretos. Lo ilustraremos con un algoritmo iterativo y una aplicación de tiempo real.

Ejemplo 1.5 – Solución iterativa

Los algoritmos iterativos son ejemplos de sistemas inherentemente muestreados. Supongamos que se desea la solución de la ecuación de la forma

$$x - f(x) = 0$$

Una manera de hallar dicha solución es suponer un valor inicial y entonces utilizar el algoritmo de Picard; es decir, emplear el esquema iterativo

$$x(k+1) = f[x(k)]$$

donde $x(k)$ es la k -ésima iteración. El algoritmo numérico puede interpretarse como un sistema discreto en el que el tiempo representa el número de iteraciones.

Concretamente, supongamos que $f(x) = 3 - \sqrt{x}$. En este caso es fácil ver que la solución es $x = (7 - \sqrt{13})/2 \approx 1.697$. La secuencia de números de la Figura 1.7 es la que se obtiene si se toma como estimación inicial $x(0) = 0$.

Ejemplo 1.6 – Algoritmo de control

Un algoritmo de control sencillo para un controlador proporcional e integral (PI) es el siguiente:

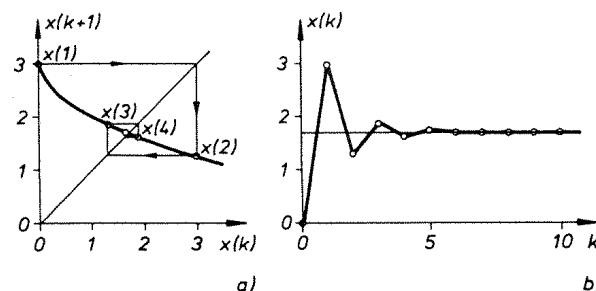


Figura 1.7 Dos ilustraciones gráficas del esquema iterativo del Ejemplo 1.5.

```

uc:=adin(in1)      {leer valor referencia}
y:=adin(in2)       {leer valor proceso}
e:=uc-y
u:=k*(e+i)
dout(u)           {sacar señal de control}
i:=i+h*e/ti

```

El programa se ejecuta cada período de muestreo por un programa de planificación temporal, como lo ilustra la Figura 1.8. El código de computador es equivalente al siguiente conjunto de ecuaciones en diferencias:

$$\begin{aligned} e(k) &= uc(k) - y(k) \\ u(k) &= k[e(k) + i(k-1)] \\ i(k) &= i(k-1) + \frac{he(k)}{ti} \end{aligned}$$

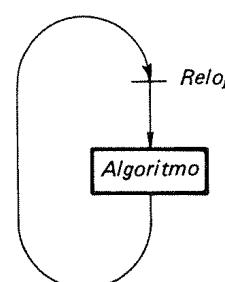


Figura 1.8 Planificación de un programa de computador.

Muestreo debido al sistema de medida

En muchos casos, el muestreo tendrá lugar de forma natural debido al procedimiento de medida. Veamos algunos ejemplos.

Ejemplo 1.7 – Radar

Cuando una antena de radar gira, la información sobre la distancia y la dirección se obtiene una vez por cada revolución de la antena. Por tanto, la forma natural de describir un sistema de radar es un modelo muestreado. De hecho, uno de los puntos de arranque de la teoría de sistemas muestreados fueron los intentos de describir los sistemas de radar.

Ejemplo 1.8 – Instrumentos de análisis

En los sistemas de control de procesos hay muchas variables que no pueden medirse en línea, por lo que se analiza fuera del proceso una muestra del producto mediante un instrumento de análisis, como puede ser un espectrógrafo de masas o un cromatógrafo.

Ejemplo 1.9 – Sistemas económicos

Los procedimientos contables en los sistemas económicos suelen ir unidos al calendario. Aunque las transacciones pueden tener lugar en cualquier momento, la información sobre las variables importantes se acumula sólo en ciertos momentos; por ejemplo: diaria, semanal, mensual, trimestral o anualmente.

Muestreo debido a operación discreta

Muchos sistemas son inherentemente muestreados porque la información se transmite empleando información pulsada. Los circuitos electrónicos son un ejemplo. También ellos fueron fuente de inspiración para el desarrollo de la teoría de los sistemas muestreados. Se dan también algunos ejemplos.

Ejemplo 1.10 – Control de un tiristor

La electrónica de potencia usando tiristores da lugar a sistemas muestreados. Consideremos el circuito de la Figura 1.9. La corriente puede circular sólo cuando la tensión es positiva. Cuando circula corriente, sigue haciéndolo hasta que pasa por el valor cero. La corriente se sincroniza con la periodicidad de la fuente principal.

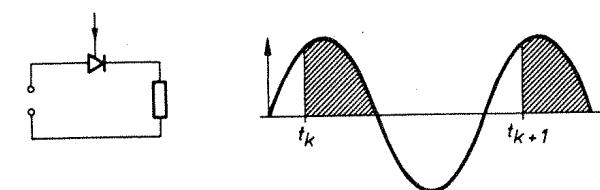


Figura 1.9 Circuito de control de un tiristor.

Ejemplo 1.11 – Sistemas biológicos

Los sistemas biológicos son fundamentalmente de tipo muestreado porque la transmisión de señal en el sistema nervioso se realiza en forma de pulsos.

Ejemplo 1.12 – Motores de combustión interna

Un motor de combustión interna es un sistema muestreado. La ignición puede verse como un reloj que sincroniza la operación del motor. En cada ignición se genera un pulso de par.



Los sistemas expuestos en estos ejemplos son periódicos debido a su funcionamiento pulsante. Los sistemas periódicos son bastante difíciles de manejar, pero pueden simplificarse de forma considerable estudiándolos en instantes sincronizados con los pulsos; esto es, utilizando modelos muestrados. Los procesos pueden describirse como sistemas discretos invariantes en el tiempo en los instantes de muestreo. Los Ejemplos 1.10 y 1.12 son de este tipo.

1.5. Cómo se desarrolló la teoría

Aunque las aplicaciones más importantes de la teoría de los sistemas muestrados están normalmente dentro del campo del control por computador, muchos de los problemas ya se habían encontrado antes. En esta sección expondremos algunas de las principales ideas que intervinieron en el desarrollo de la teoría.

El teorema de muestreo

Debido a que todos los sistemas controlados por computador operan sobre los valores de las variables de proceso solamente en los instantes de muestreo, es muy importante saber las condiciones bajo las cuales se puede reconstruir una señal a partir únicamente de sus valores en instantes discretos. El tema clave lo investigó Nyquist, quien mostró que para reconstruir una señal sinusoidal a partir de sus muestras es necesario muestrear al menos dos veces en cada período. La solución completa la dio Shannon en un importante trabajo en 1949.

Las ecuaciones en diferencias

Los primeros gérmenes de una teoría de los sistemas muestrados aparecieron en conexión con el análisis de sistemas de control específicos. Una de las primeras contribuciones a la teoría fue el comportamiento de un galvanómetro estabilizado por un troceador investigado por Oldenburg y Sartorius (1948). Mostró cómo muchas propiedades podían comprenderse analizando una ecuación en diferencias lineal y de coeficientes constantes. La ecuación en diferencias sustituía a la ecuación diferencial en la teoría de los sistemas continuos. Por ejemplo, se podía estudiar la estabilidad con el método de Schur-Cohn, que es equivalente al criterio de Routh-Hurwitz.

Los métodos de las transformadas

Durante y después de la Segunda Guerra Mundial se dedicó una gran cantidad de actividad al análisis de los sistemas de radar. Estos sistemas son de los intrínsecamente muestrados debido a que se obtiene una medida de posición por cada revolución de la antena. Como la teoría de la transformada había resultado tan

útil para los sistemas continuos, era natural intentar desarrollar una teoría similar para los sistemas muestrados. El primer paso en esta dirección lo dio Hurewicz (1947). Introdujo la transformación de una secuencia $\{f(kh)\}$, definida por

$$Z\{f(kh)\} = \sum_{k=0}^{\infty} z^{-k} f(kh)$$

Esta transformación es similar a la *función generatriz*, que se había usado con tanto éxito en muchas ramas de las matemáticas aplicadas. Más tarde Ragazzini y Zadeh (1952) definieron la transformación como la *transformada z*. La teoría de la transformada se desarrolló de forma independiente en la Unión Soviética, los Estados Unidos y Gran Bretaña. Tsypkin (1950) la llamó la *transformada discreta de Laplace* y desarrolló una teoría sistemática para los sistemas controlados por pulsos basada en la transformada. El método de la transformada fue desarrollado también de forma independiente por Barker (1952) en Inglaterra.

En los Estados Unidos Jury desarrolló el tema de forma más amplia en su tesis doctoral en la Universidad de Columbia. Jury desarrolló herramientas tanto para el análisis como para el diseño. Demostró también que los sistemas muestrados podían ser mejores que sus equivalentes continuos (véase el Ejemplo 1.4 en la Sección 1.3). Jury también destacó que era posible obtener un sistema en lazo cerrado que alcanzara exactamente el estado estacionario en un tiempo finito. En trabajos posteriores demostró también que el muestreo puede producir la cancelación de polos y ceros. Una investigación posterior más profunda de esta propiedad dio lugar a las nociones de observabilidad y alcancabilidad.

La teoría de la transformada z conduce a resultados comparativamente muy sencillos. Sin embargo, una limitación de la teoría es que dice lo que ocurre a los sistemas sólo en los instantes de muestreo. El comportamiento entre los instantes de muestreo no es sólo una cuestión académica, porque se vio que estos sistemas podían presentar *oscilaciones ocultas*. Estas oscilaciones son nulas en los instantes de muestreo, pero muy notables entre éstos.

Otra aproximación a la teoría de los sistemas muestrados fue la proporcionada por Linvill (1951). Siguiendo las ideas de MacColl (1945), consideró el muestreo como una modulación en amplitud. Empleando la aproximación de la función descriptiva, Linvill consiguió describir de forma eficaz el comportamiento entre instantes de muestreo. Otra aproximación diferente al análisis del problema fue la *transformada z retardada*, que fue desarrollada por Tsypkin en 1950, Barker en 1951 y Jury en 1956. Se la conoce también como la *transformada z modificada*.

Una gran parte del desarrollo de la teoría la llevó a cabo un grupo de la Universidad de Columbia en Estados Unidos, dirigido por Ragazzini. Jury, Kalman, Bertram, Zadeh, Franklin, Friedland, Kranc, Freeman, Sarachik y Sklansky hicieron todos su tesis doctoral con Ragazzini.

Al final de los años 50, la aproximación de la transformada z para los sistemas muestrados había madurado, y aparecieron casi simultáneamente varios libros de texto: Ragazzini y Franklin (1958), Jury (1958), Tsypkin (1958) y Tou (1959). Esta teoría, que se había creado siguiendo el modelo de la teoría de los sistemas lineales

invariantes y continuos, proporcionó buenas herramientas para el análisis y la síntesis de los sistemas muestreados. Sólo fue preciso hacerle unas pocas modificaciones debido a la naturaleza variable en el tiempo de los sistemas muestreados. Por ejemplo, ¡todas las operaciones en la representación de diagramas de bloques no comutan!

Teoría del espacio de estados

Un acontecimiento muy importante a finales de los años cincuenta fue el desarrollo de la teoría del espacio de estados. La inspiración principal vino por parte de las matemáticas y la teoría de las ecuaciones diferenciales ordinarias y de matemáticos tales como Lefschetz, Pontryagin y Bellman. A Kalman hay que atribuir el mérito del enfoque del espacio de estados en la teoría de control. Fue él quien formuló muchos de los conceptos básicos y resolvió muchos de los problemas importantes.

Varios de los conceptos fundamentales se elaboraron durante el análisis del problema de si sería posible obtener sistemas en los que las variables alcanzaran el estado estacionario en un tiempo finito. El análisis de este problema llevó a las nociones de alcanzabilidad y observabilidad. El trabajo de Kalman condujo también a una formulación mucho más simple del análisis de los sistemas muestreados: las ecuaciones básicas podían deducirse de forma sencilla comenzando con las ecuaciones diferenciales e integrándolas con la hipótesis de que la señal de control es constante durante el período de muestreo.

Control óptimo y control estocástico

Hubo también otros desarrollos importantes a finales de los cincuenta. Bellman (1957) y Pontryagin con otros (1962) demostraron que muchos problemas de diseño podían formularse como problemas de optimización. Para los sistemas no lineales esto llevó al cálculo de variaciones no clásico. Bellman y otros (1958) dieron una solución explícita para los sistemas lineales con función de coste cuadrática. Kalman (1960) demostró en un célebre trabajo que el problema lineal cuadrático podía reducirse a la solución de una ecuación de Riccati. También demostró que el clásico problema del filtrado de Wiener podía reformularse en el espacio de estados. Esto permitió dar una «solución» en forma de ecuaciones recursivas, lo que se adaptaba muy bien al cálculo con computador.

Al comienzo de los años sesenta se formuló un problema variacional estocástico suponiendo que las perturbaciones eran procesos aleatorios. Entonces se pudo formular y resolver el problema del control óptimo para sistemas lineales para el caso de funciones de coste cuadráticas. Esto condujo al desarrollo de la *teoría de control estocástico*. El trabajo dio como resultado la llamada teoría Lineal Cuadrática Gaussiana (LQG). Esta es en la actualidad una herramienta fundamental de diseño para los sistemas lineales multivariados.

Teoría algebraica de sistemas

A finales de los años sesenta y comienzo de los setenta se reconsideraron los problemas fundamentales de la teoría de los sistemas lineales. Se reestableció el carácter algebraico de los problemas, lo que condujo a una mejor comprensión de los fundamentos de la teoría de sistemas lineales. Otro resultado fueron las técnicas para resolver problemas específicos empleando métodos polinomiales [véase Kalman y otros (1969), Rosenbrock (1970), Wonham (1974), Blomberg y Ylinen (1983) y Kucera (1979)].

Identificación de sistemas

Cuando se intentó aplicar la teoría a problemas prácticos, quedó patente la necesidad de herramientas eficaces de modelado. Esta necesidad estimuló muchísimo la investigación sobre técnicas para obtener los modelos directamente de los datos de la planta. Estos esfuerzos dieron como fruto una buena comprensión del problema y una gran cantidad de algoritmos [véase Åström y Eykhoff (1971) y Goodwin y Payne (1977)].

Control adaptativo

La llegada del control con computador ha hecho posible la realización práctica de algoritmos de control más complicados. Esto ha renovado el interés en el control adaptativo. Durante los años setenta se ha progresado en el desarrollo tanto de algoritmos como de la teoría. Además, se ha demostrado mediante estudios de viabilidad la utilidad del control adaptativo [véase, por ejemplo, Åström (1983)].

1.6. Referencias

Para lograr un conocimiento sólido sobre un tema es útil conocer su historia y leer alguno de los trabajos originales. Los trabajos

JURY, E. I. (1980): «Sampled-data Systems, Revisited: Reflections, Recollections, and Reassessments,» *Trans. of the ASME, Journal of Dynamic Systems, Measurement, and Control*, 102 (December 1980) 208-16.

JURY, E. I. y Y. Z. TSYPKIN (1971): «On the Theory of Discrete Systems», *Automatica* 7, 89-107.

escritos por dos de los originadores de la teoría de los sistemas muestreados dan una perspectiva muy útil. El trabajo inicial sobre los sistemas muestreados se encuentran en

MACCOLL, L. A. (1945): *Fundamental Theory of Servomechanisms*. New York: D. Van Nostrand.

OLDENBURG, R. C. y H. SARTORIUS (1984): *The Dynamics of Automatic Control*. New York: ASME.



HUREWICZ, W. (1947): «Filters and Servo Systems with Pulsed Data», en *Theory of Servomechanisms*, ed. H. M. James, N. B. Nichols, y R. S. Philips. New York: McGraw-Hill.

El teorema de muestreo se da en

SHANNON, C. E. (1949): «Communication in Presence of Noise», *Proc. IRE*, 37, 10-21.

y

KOTELNIKOV, V. A. (1933): «On the Transmission Capacity of "Ether" and Wire in Electrocommunications», *Proc. First All-union Conference on Questions of Communication*. Moscow.

Las contribuciones más importantes a la teoría inicial de los sistemas de datos muestreados las obtuvieron en Inglaterra

BARKER, R. H. (1952): «The Pulse Transfer Function and Its Applications to Sampling Servosystems», *Proc. IEE*, 99, pt. IV, 302-17.

LAWDEN, D. F. (1951): «A General Theory of Sampling Servomechanisms», *Proc. IEE*, 98, pt. IV (October) 31-36.

en los Estados Unidos

LINVILL, W. K. (1951): «Sampled-Data Control Systems Studied Through Comparison of Sampling with Amplitude Modulation», *AIEE Trans.*, 70, pt. II, 1778-88.

RAGAZZINI, J. R. y L. A. ZADEH (1952): «The Analysis of Sampled-Data Systems», *AIEE Trans.*, 71, pt. II (November) 225-34.

JURY, E. I. (1956): «Synthesis and Critical Study of Sampled-Data Control Systems», *AIEE Trans.*, 75, pt. II, 141-51.

y en la Unión Soviética

TSYPKIN, Y. Z. (1949) y (1950): «Theory of Discontinuous Control», *Avtomat i Telemekh.*, 3 (1949), 5 (1949), 5 (1950).

Los primeros libros de texto sobre la teoría de datos muestreados aparecieron hacia finales de los años cincuenta

RAGAZZINI, J. R. y G. F. FRANKLIN (1958): *Sampled-Data Control Systems*. New York: McGraw-Hill.

TSYPKIN, Y. Z. (1958): *Theory of Impulse Systems*. Moscow: Editora Estatal de Literatura Física y Matemática.

JURY, E. I. (1958): *Sampled-Data Control Systems*. New York: John Wiley.

TOU, J. T. (1959): *Digital and Sampled-Data Control Systems*. New York: McGraw-Hill.

Desde entonces han aparecido un gran número de libros de texto. Entre los más conocidos podemos citar

ACKERMANN, J. (1972): *Abtastregelung*. Berlin: Springer-Verlag.

ISERMANN, R. (1977): *Digitales Regelsysteme*. Trad. Inglesa (1981): *Digital Control Systems*. Berlin: Springer-Verlag.

KUO (1980): *Digital Control Systems*. Tokyo: Holt-Saunders.

FRANKLIN, G. E. y J. D. POWELL (1980): *Digital Control of Dynamic Systems*. Reading, Mass.: Addison-Wesley.

La idea de formular los problemas de control en el espacio de estados también llevó consigo una reformulación de la teoría de los sistemas muestreados. El trabajo

KALMAN, R. E. (1960): «On the General Theory of Control Systems», *Proc. First IFAC Congress*, Moscow, *Butterworths*, 1, 481-92.

es fundamental.

Algunas referencias básicas sobre control óptimo y estocástico son

BELLMAN, R. (1957): *Dynamic Programming*. Princeton, N. J.: Princeton University Press.

BELLMAN, R., I. GLICKSBERG, y O. A. GROSS (1958): «Some Aspects of the Mathematical Theory of Control Processes», Report R-313. Santa Monica, Calif.: The RAND Corporation.

PONTRYAGIN, L. S., V. G. BOLTYANSKII, R. V. GAMKRELIDZE, y E. F. MISCHENKO (1962): *The Mathematical Theory of Optimal Processes*. New York: John Wiley.

KALMAN, R. E. (1960): «Contributions to the Theory of Optimal Control», *Boletín de la Sociedad Matemática Mexicana*, 5, 102-19.

La aproximación algebraica a los sistemas se expone en

KALMAN, R. E., P. L. FALB, y ARBIB, M. A. (1969): *Topics in Mathematical System Theory*. New York: McGraw-Hill.

ROSENBROCK, H. H. (1970): *State-Space and Multivariable Theory*. London: Nelson.

WONHAM, W. M. (1974): *Linear Multivariable Control: A Geometric Approach*. New York: Springer-Verlag.

BLOMBERG, H. y R. YLINEN (1983): *Algebraic Theory for Multivariable Linear Systems*. New York: Academic Press.

KUČERA, V. (1979): *Discrete Linear Control*. Prague: Academia.

La identificación de sistemas se trata en

ÅSTRÖM, K. J. y P. EYKHOFF (1971): «System Identification: A Survey», *Automatica*, 7, 123-62.

GOODWIN, G. C. y R. L. PAYNE (1977): *Dynamic System Identification: Experiment Design and Data Analysis*. New York: Academic Press.

El control adaptativo se expone en

BELLMAN, R. (1961): *Adaptive Control: A Guided Tour*. Princeton, N. J.: Princeton University Press.

ÅSTRÖM, K. J. y B. WITTENMARK (1973): «On Self-tuning Regulators», *Automatica*, 9, 185-99.

ÅSTRÖM, K. J. y B. WITTENMARK (1980): «Self-tuning Controllers Based on Pole-Zero Placement», *Proc. IEE pt. D*, 127, 120-30.

ÅSTRÖM, K. J. (1983): «Theory and Applications of Adaptive Control», *Automatica*, 19.

En las siguientes Secciones se dan muchas más referencias. También recomendamos las Actas de los Simposium de la IFAC sobre «Digital Computer Applications to Process Control» y sobre «Identification and System Parameter Estimation», que publica Pergamon Press,



Muestreo de señales continuas

Objetivo:

Comprender los mecanismos del muestreo e introducir algunos de los conceptos fundamentales y notaciones para los sistemas muestreados. Ilustrar el problema del enmascaramiento de frecuencias.

2.1. Introducción

Se entiende por *muestreo* «la acción o proceso de tomar una pequeña parte o cantidad de algo para su examen o análisis». En el contexto del control y de las comunicaciones, muestreo significa que *una señal continua en el tiempo se reemplaza por una secuencia de números*, los cuales representan los valores de la señal en ciertos instantes.

El muestreo es una propiedad fundamental de los sistemas controlados con computador debido a la naturaleza discreta de los computadores digitales. Considerese, por ejemplo, el sistema de la Figura 1.1. Las variables analógicas del proceso se muestrean en combinación con la conversión analógica y posteriormente se digitalizan para su procesamiento. De este modo, la señal continua que representa la variación en el tiempo de las variables del proceso se convierte en una secuencia de números, los cuales se procesan en el computador digital. El procesamiento da una nueva secuencia de números, que se convierten en una señal continua y se aplican al proceso. En el sistema de la Figura 1.1 esto lo realiza el convertidor D-A. El proceso de convertir una secuencia de números en una señal continua se denomina *reconstrucción de la señal*.

Debido a que el muestreo es una propiedad básica de los sistemas controlados con computador, es necesario tener un buen conocimiento del proceso de muestreo. El muestreo de señales continuas se expone en este capítulo; la descripción de sistemas con señales muestreadas se discute en el Capítulo 3. Los mecanismos de muestreo se tratan en la Sección 2.2. Aunque en la teoría se incluyen diferentes tipos de muestreo,

la mayoría de ellos se dedican al caso especial de muestreo periódico. La Sección 2.3 trata el problema que aparece de representar exactamente una señal continua por sus valores muestreados. El teorema de muestreo de Shannon proporciona el resultado para muestreo periódico. En la Sección 2.4 se discuten diferentes modos de reconstrucción de una señal continua a partir de sus valores muestreados. También se estudia la reconstrucción de Shannon y los mantenedores de primer orden y de orden superior.

Un efecto peculiar denominado *enmascaramiento de frecuencias* se presenta cuando la señal muestreada contiene frecuencias superiores a la mitad de la frecuencia de muestreo. La Sección 2.5 trata el problema del enmascaramiento de frecuencias o solapamiento de frecuencias y el uso de filtros previos para evitarlo. Los aspectos prácticos de la elección de los períodos de muestreo se discuten en la Sección 2.6.

2.2. Descripción de los mecanismos de muestreo

Para la tarea del análisis resulta útil tener una descripción del muestreo. Muestrear una señal continua simplemente significa que se reemplaza la señal por sus valores en un conjunto de puntos discretos. Consideremos el conjunto de los números enteros $Z = \{\dots, -1, 0, 1, \dots\}$ y sea $\{t_k : k \in Z\}$ un subconjunto de los números reales denominado instantes de muestreo. Entonces, la versión muestreada de la señal f es la secuencia $\{f(t_k) : k \in Z\}$. El muestreo es una operación lineal. A menudo, los instantes de muestreo se distribuyen en intervalos regulares en el tiempo; es decir,

$$t_k = k \cdot h$$

Este caso se denomina *muestreo periódico* y h es el período de muestreo o intervalo de muestreo. La frecuencia correspondiente $f_s = 1/h$ (Hz) se denomina *frecuencia de muestreo*.

También pueden utilizarse otros esquemas más complejos. Por ejemplo, pueden utilizarse diferentes períodos de muestreo para diferentes lazos de control. Esta técnica se denomina *muestreo multiperíodo* y puede considerarse como la superposición de varios muestreos periódicos.

El caso del muestreo periódico es bien conocido. La mayor parte de la teoría se centra en este caso; sin embargo, el muestreo multiperíodo está incrementando su importancia debido al uso creciente de los sistemas multiprocesador. Con los modernos programas para procesos concurrentes es posible diseñar sistemas como si estuviesen formados por diferentes procesadores funcionando asíncronamente. También se encuentran ventajas técnicas en el uso de diferentes velocidades de muestreo para diferentes variables.

2.3. El teorema de muestreo

Se pierde muy poca información muestreando una señal continua si los instantes de muestreo están muy próximos; sin embargo, gran parte de la información referente

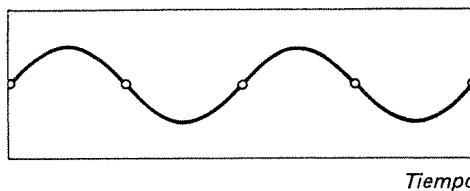


Figura 2.1 Pérdida de información debida a un muestreo lento. La curva senoidal se muestrea a una velocidad de dos muestras por periodo.

a la señal puede perderse si los puntos de muestreo están alejados. Esta propiedad se ilustra en la Figura 2.1, en la cual se muestrea una señal senoidal. La figura muestra que una señal senoidal no puede distinguirse de una señal nula si la frecuencia de la señal senoidal es la mitad de la frecuencia de muestreo.

Esto es, por supuesto, esencial para conocer con precisión cuándo una señal continua está definida de forma única por su versión muestreada. El siguiente teorema da las condiciones para el caso de muestreo periódico.

Teorema 2.1 – Teorema de muestreo de Shannon. Una señal continua con una transformada de Fourier nula fuera del intervalo $(-\omega_0, \omega_0)$ está determinada de forma única por sus valores en puntos equidistantes si la frecuencia de muestreo es superior a $2\omega_0$. La señal continua puede calcularse a partir de la señal muestreada mediante la fórmula de interpolación

$$f(t) = \sum_{k=-\infty}^{\infty} f(kh) \frac{\sin \omega_s(t - kh)/2}{\omega_s(t - kh)/2} \quad (2.1)$$

donde ω_s es la frecuencia angular de muestreo en radianes por segundo (rad/s).

*Demuestra*ón. Consideremos la señal f y sea F su transformada de Fourier.

$$F(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} f(t) dt \quad (2.2)$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} F(\omega) d\omega \quad (2.3)$$

Y definamos

$$F_s(\omega) = \frac{1}{h} \sum_{k=-\infty}^{\infty} F(\omega + k\omega_s) \quad (2.4)$$

La demostración se basa en la observación de que las muestras $f(kh)$ pueden considerarse como los coeficientes de la serie de Fourier de la función periódica $F_s(\omega)$. La demostración es directa. El desarrollo de Fourier de F_s es

$$F_s(\omega) = \sum_{k=-\infty}^{\infty} C_k e^{-ik\omega h} \quad (2.5)$$

donde los coeficientes están definidos por

$$C_k = \frac{1}{\omega_s} \int_0^{\omega_s} e^{ik\omega h} F_s(\omega) d\omega$$

Utilizando la definición de los coeficientes de Fourier, definidos en (2.3) y (2.4), es inmediato demostrar que

$$C_k = f(kh) \quad (2.6)$$

De este modo se obtiene que la señal muestreada $\{f(kh), k = \dots, -1, 0, 1, \dots\}$, define de modo único la función $F_s(\omega)$. La función F , por consideraciones del teorema, es nula fuera del intervalo $(-\omega_0, \omega_0)$. Si $\omega_s > 2\omega_0$ de (2.4) se obtiene

$$F(\omega) = \begin{cases} hF_s(\omega) & |\omega| \leq \frac{\omega_s}{2} \\ 0 & |\omega| > \frac{\omega_s}{2} \end{cases} \quad (2.7)$$

De este modo, la transformada de Fourier de la señal continua está determinada únicamente por F_s , la cual está definida por la función muestreada $\{f(kh), k = \dots, -1, 0, 1, \dots\}$. La primera parte del teorema queda demostrada. Para demostrar la expresión (2.1), debe observarse que de (2.3) y (2.7) se obtiene

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} F(\omega) d\omega \\ &= \frac{h}{2\pi} \int_{-\omega_s/2}^{\omega_s/2} e^{i\omega t} F_s(\omega) d\omega \\ &= \frac{h}{2\pi} \int_{-\omega_s/2}^{\omega_s/2} e^{i\omega t} \sum_{k=-\infty}^{\infty} e^{-ik\omega h} f(kh) d\omega \end{aligned}$$

donde la última igualdad se obtiene a partir de (2.5) y (2.6). Intercambiando el orden de la integración y la suma,

$$\begin{aligned} f(t) &= \sum_{k=-\infty}^{\infty} f(kh) \frac{h}{2\pi} \int_{-\omega_s/2}^{\omega_s/2} e^{i\omega t - ik\omega h} d\omega \\ &= \sum_{k=-\infty}^{\infty} f(kh) \frac{h}{2\pi i(t - kh)} e^{i\omega t - ik\omega h} \Big|_{-\omega_s/2}^{\omega_s/2} \\ &= \sum_{k=-\infty}^{\infty} f(kh) \frac{\sin \omega_s(t - kh)/2}{\pi(t - kh)/h} \end{aligned}$$

Y como $\omega_s h = 2\pi$, la ecuación (2.1) queda demostrada. \square



Nota 1. La frecuencia $\omega_N = \omega_s/2$ evidentemente juega un papel muy importante. Esta frecuencia se denomina *frecuencia de Nyquist*.

Nota 2. Debe observarse que la ecuación (2.1) define la reconstrucción de la señal cuya transformada de Fourier es nula para frecuencias superiores a la frecuencia de Nyquist $\omega_N = \omega_s/2$.

Nota 3. En algunas ocasiones se dice que la operación de muestreo tiene una ganancia $1/h$ debido a la presencia del factor $1/h$ en la ecuación (2.4).

2.4. Reconstrucción

La inversa de la operación de muestreo, es decir, la conversión de una secuencia de números $\{f(t_k): k \in \mathbb{Z}\}$ en una función continua $f(t)$, se denomina *reconstrucción*. En sistemas controlados por computador es necesario convertir las acciones de control, calculadas por el computador como una secuencia de números, en una señal continua que puede aplicarse al proceso. En filtrado digital es del mismo modo necesario convertir la representación de la señal filtrada, dada como una secuencia de números, en una función continua. En esta sección se estudian diferentes tipos de reconstrucciones.

Reconstrucción de Shannon

Para el caso de señales periódicas de banda limitada se deduce del teorema de muestreo que la reconstrucción viene dada por (2.1). Esta reconstrucción se denomina *reconstrucción de Shannon*. La ecuación (2.1) define la operación inversa del muestreo, que puede considerarse como un operador lineal. Sin embargo, no es un operador causal, puesto que los valores de f en los instantes t están expresados en función de los valores pasados $\{f(kh): k \leq t/h\}$ y de los valores futuros $\{f(kh): k > t/h\}$. Esto implica que la reconstrucción de Shannon no es útil en control con computador, pero puede usarse en algunos problemas de comunicaciones donde se aceptan determinados retardos. Otros inconvenientes de la reconstrucción de Shannon residen en su complejidad y en que sólo puede aplicarse en muestreo periódico. Resulta, por tanto, útil disponer de otros tipos de reconstrucciones.

Mantenedor de orden cero (ZOH)

Una reconstrucción causal simple está definida por

$$f(t) = f(t_k), \quad t_k \leq t < t_{k+1} \quad (2.8)$$

Esto significa que la señal reconstruida es constante a intervalos, continua por la derecha e igual a la señal muestreada en los instantes de muestreo. De este modo, el valor reconstruido se mantiene constantes hasta el siguiente período de muestreo.

Debido a su simplicidad, el mantenedor de orden cero (ZOH, del inglés Zero-Order Hold) es muy común en sistemas controlados con computador. Los conversores D-A comerciales suelen diseñarse de tal modo que cada valor de una conversión se mantiene constante hasta que se realiza una nueva conversión.

El mantenedor de orden cero tiene la ventaja de que puede utilizarse para muestreo aperiódico. Debe observarse, sin embargo, que la reconstrucción definida por (2.8) proporciona una representación exacta de la operación inversa al muestreado solamente para señales que son continuas por la derecha y constantes sobre cada intervalo de muestreo. Para otro tipo de señal, la reconstrucción de (2.8) produce errores (véase la Fig. 2.2). Para el muestreo periódico de una señal con una derivada primera suave, puede obtenerse la siguiente estimación del error. El máximo valor del error viene dado por

$$\epsilon_{ZOH} = \max_k |f(t_{k+1}) - f(t_k)| \leq h \max |f'(t)| \quad (2.9)$$

donde f' es la derivada de f .

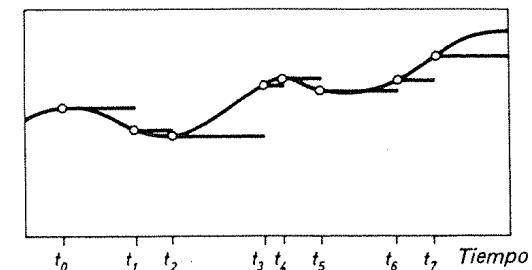


Figura 2.2 Muestreo y reconstrucción con un mantenedor de orden cero de una señal continua.

Mantenedores de orden superior

El mantenedor de orden cero puede considerarse como una extrapolación que utiliza un polinomio de grado cero. Para funciones suaves, pueden obtenerse errores de reconstrucción más pequeños extrapolando con polinomios de orden superior. Una extrapolación polinómica causal de primer orden da

$$f(t) = f(t_k) + \frac{t - t_k}{t_k - t_{k-1}} [f(t_k) - f(t_{k-1})], \quad t_k \leq t < t_{k+1}$$

La reconstrucción se obtiene dibujando una línea recta entre las muestras más cercanas. El mantenedor de primer orden se ilustra en la Figura 2.3.

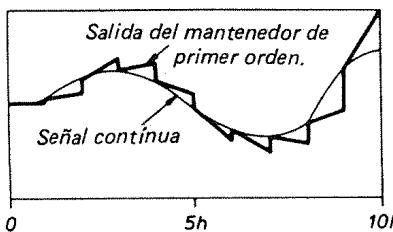


Figura 2.3 Muestreo y reconstrucción con un mantenedor de primer orden de una señal continua.

El mayor error que se obtiene cuando se utiliza un mantenedor de primer orden viene dado por

$$e_{FOH} = \max_k \max_t \left| f(t) - f(t_k) - \frac{t - t_k}{t_k - t_{k-1}} [f(t_k) - f(t_{k-1})] \right| \quad (2.10)$$

Para muestreo periódico de señales con derivada segunda suave, el error se puede estimar por

$$e_{FOH} \leq h^2 \max_t |f''(t)| \quad (2.11)$$

Para señales con derivadas suaves de mayor orden puede realizarse una extrapolación polinómica de orden superior. Estos tipos de reconstrucciones no son muy comunes en la práctica debido a su dificultad de realización.

2.5. Enmascaramiento o solapamiento de frecuencias

Si una señal continua que tiene una transformada de Fourier F se muestrea periódicamente, se obtiene de (2.5) y (2.6) que las señales muestreadas $f(kh)$, $k = \dots, -1, 0, 1, \dots$, pueden interpretarse como los coeficientes de Fourier de la función F_s definida por (2.4).

De este modo, la función F_s puede interpretarse como la transformada de Fourier de la señal muestreada. La función (2.4) es periódica de periodo igual a la frecuencia de muestreo ω_s . Si la señal continua no tiene componentes de frecuencia superiores a la frecuencia de Nyquist, la transformada de Fourier es simplemente una repetición periódica de la transformada de Fourier de la señal continua (véase la Fig. 2.4).

Se obtiene de (2.4) que el valor de la transformada de Fourier de la señal muestreada a la frecuencia ω es la suma de los valores de la transformada de Fourier de la señal continua a las frecuencias $\omega + n\omega_s$. Después de muestrear, ya no es posible separar la contribución de cada una de las frecuencias. La frecuencia ω puede, de este modo, considerarse como el enmascaramiento de $\omega + n\omega_s$. Se acostumbra a considerar únicamente las frecuencias positivas. La frecuencia ω es entonces la máscara de $\omega_s - \omega, \omega_s + \omega, 2\omega_s - \omega, 2\omega_s + \omega, \dots$, donde $0 \leq \omega < \omega_N$. Después de muestrear, una

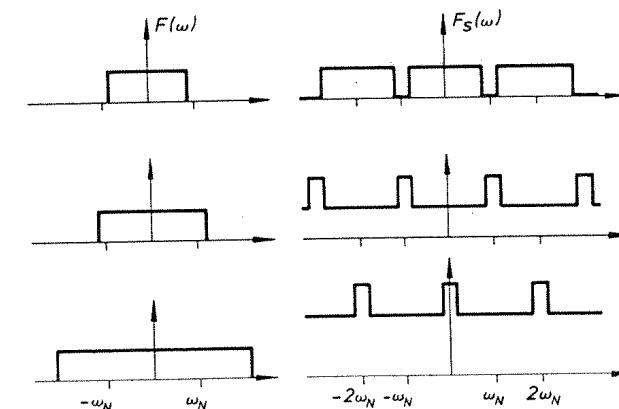


Figura 2.4 Relación entre la transformada de Fourier para señales continuas y muestreadas para diferentes frecuencias de muestreo. Por simplicidad se ha supuesto que la transformada de Fourier es real.

frecuencia no puede distinguirse de las que enmascaran. El enmascaramiento fundamental para una frecuencia $\omega_1 > \omega_N$ viene expresada por

$$\omega = |(\omega_1 + \omega_N) \text{mod}(\omega_1) - \omega_N| \quad (2.12)$$

Debe observarse que aunque el muestreo es una operación lineal, no es invariantes en el tiempo. Esto explica por qué se crean nuevas frecuencias en la operación de muestreo. Esto se discutirá, posteriormente, en el Capítulo 4.

La Figura 2.5 ilustra el efecto del enmascaramiento de frecuencias. Dos señales de 0,1 Hz y 0,9 Hz de frecuencia se muestrean con una frecuencia de 1 Hz ($h = 1$ s). En la figura se observa que ambas señales tienen los mismos valores en los instantes de muestreo.

Este problema se ilustra con un ejemplo.

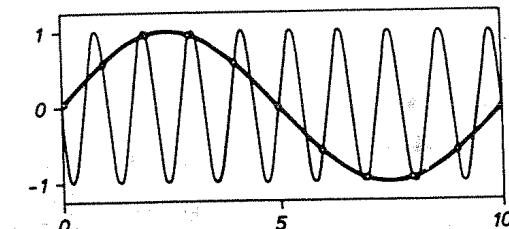


Figura 2.5 Dos señales con frecuencias diferentes, 0,1 Hz y 0,9 Hz, pueden tener el mismo valor en todos los instantes de muestreo.



Ejemplo 2.1 – Enmascaramiento de frecuencias

La Figura 2.6 muestra el diagrama del proceso de calentamiento del agua de alimentación de la caldera de un barco. El flujo de agua está controlado por una válvula. Existe un juego en el posicionador de la válvula debido al desgaste. Esto causa oscilaciones en la temperatura y la presión. La Figura 2.7 representa un registro muestreado de la temperatura y un registro continuo de la presión. Del registro de la temperatura puede pensarse que hay una oscilación con un período de 38 min. El registro de la presión revela, sin embargo, que la oscilación en la presión tiene un período de 2,11 min. Físicamente las dos variables están acopladas y deberían oscilar con la misma frecuencia. La temperatura se muestrea cada minuto. La frecuencia de muestreo es $\omega_s = 2\pi/2 = 3,142$ rad/min y la frecuencia a la que oscila la presión es $\omega_0 2\pi/2,11 = 2,978$ rad/min. La frecuencia enmascarada más pequeña es $\omega_s = \omega_0 = 0,1638$ rad/min. Esto corresponde a un período de 38 min., el cual es el período de oscilación del registro de temperatura.

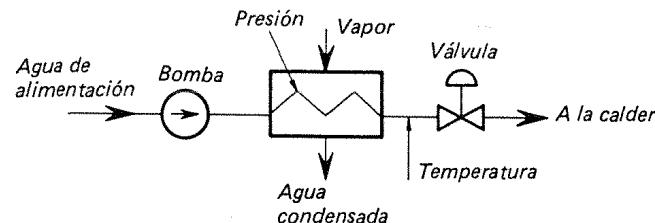


Figura 2.6 Diagrama de proceso del sistema de calentamiento del agua de alimentación de una caldera.

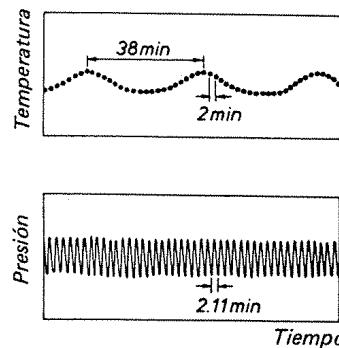
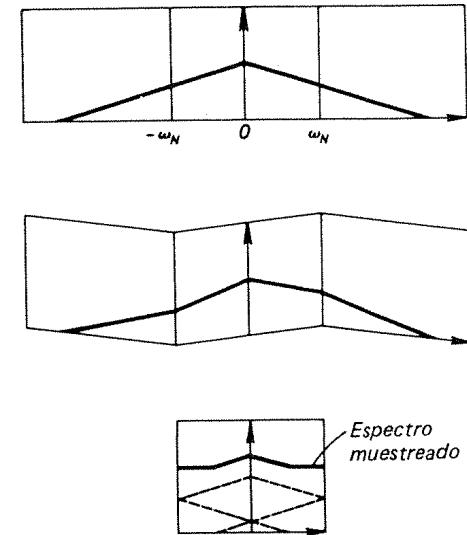


Figura 2.7 Registro de la temperatura y la presión.

Solapamiento de frecuencias

A la ecuación (2.4) puede dársele otra interpretación. Para ello, en primer lugar, se dibuja en un papel la gráfica del espectro de la señal continua. A continuación se dobla el papel en las abscisas que son múltiplos impares de la frecuencia de Nyquist,

Figura 2.8 Solapamiento de frecuencias.



como se indica en la Figura 2.8. El espectro muestreado puede obtenerse sumando, con la fase adecuada, la contribución de todas las hojas así formadas.

Prefiltrado

Una dificultad práctica que se presenta, radica en que las señales reales no tienen una transformada de Fourier que se anule fuera de un ancho de banda determinado. Las componentes de alta frecuencia pueden aparecer como componentes de baja frecuencia debido al enmascaramiento. El problema puede ser particularmente serio si hay componentes periódicas de alta frecuencia. Para evitar el problema del enmascaramiento, es necesario filtrar la señal analógica antes de muestrear. Esto puede hacerse de diferentes formas.

En la práctica todos los sensores tienen algún tipo de filtro, pero es muy poco frecuente que los filtros se elijan para un problema particular de control. Por tanto, suele ser posible modificar el filtro para que las señales obtenidas no contengan frecuencias superiores a la frecuencia de Nyquist.

En algunas ocasiones, la solución más simple consiste en introducir un filtro analógico delante del muestreador. La Figura 2.9 muestra un circuito común para un filtro de segundo orden de función de transferencia

$$G_f(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2} \quad (2.13)$$

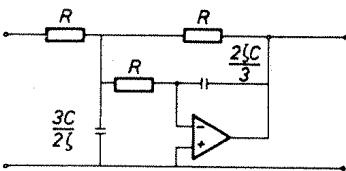


Figura 2.9 Realización mediante un amplificador operacional de un filtro de segundo orden. La función de transferencia del filtro es la dada por (2.13) con $\omega = 1/RC$.

Pueden obtenerse filtros de orden superior disponiendo en cascada sistemas de primer y segundo orden. En la tabla 2.1 se dan diferentes filtros.

El filtro de Bessel tiene una curvatura de fase lineal, lo cual significa que la forma de la señal no se distorsiona demasiado. Los filtros de Bessel son, por tanto, muy corrientes en sistemas de altas prestaciones.

TABLA 2.1 Amortiguamiento ζ y frecuencia natural ω para filtros Butterworth, ITAE (Integral en el tiempo del error absoluto), y de Bessel. Los filtros de orden superior se obtienen poniendo en cascada filtros de la forma de (2.13). Aquí, ω_0 es la frecuencia natural deseada del filtro.

Order	Butterworth		ITAE		Bessel	
	ω/ω_0	ζ	ω/ω_0	ζ	ω/ω_0	ζ
2	1	0.71	1.00	0.71	1.00	0.87
4	1	0.38	1.33	0.32	3.39	0.62
	1	0.92	0.75	0.83	3.02	0.96
6	1	0.26	1.30	0.32	5.14	0.49
	1	0.71	0.98	0.60	5.57	0.82
	1	0.97	0.79	0.93	4.34	0.98

Ejemplo 2.2 – Prefiltrado

La utilidad del prefiltro se ilustra en la Figura 2.10. En a) se muestra una señal analógica formada por una onda cuadrada con una perturbación senoidal (0,9 Hz) superpuesta. En c) se representa el resultado de muestrear la señal analógica con una frecuencia de 1 Hz. La frecuencia de Nyquist es 0,5 Hz. La perturbación de 0,9 Hz tiene un enmascaramiento de 0,1 Hz [véase (2.12)]. Esta frecuencia es claramente perceptible en la señal muestreada. La salida del prefiltro, un filtro ITAE de cuarto orden con un ancho de banda de 0,25 Hz, se muestra en b), y el resultado de muestrear con el prefiltro se representa en d). De este modo la amplitud de la perturbación se reduce mucho por la acción del prefiltro.

En el control de procesos hay una situación en la cual el prefiltroado no puede utilizarse: esto ocurre cuando se muestrea una cadena de procesos y el resultado se envía a un instrumento para su análisis. En este caso es aconsejable tomar muchas muestras y mezclarlas adecuadamente antes de enviarlas al analizador.

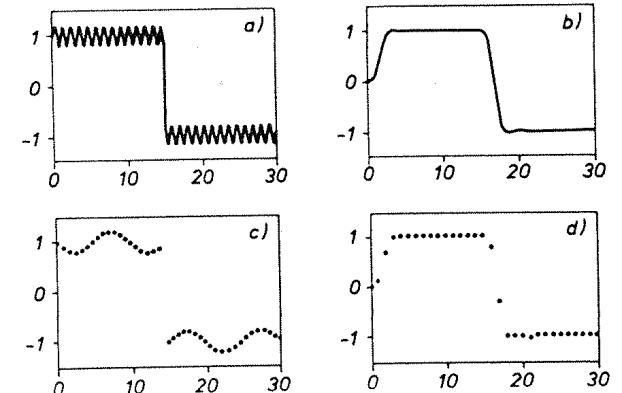


Figura 2.10 Utilidad de un prefiltrado.

- a. Señal más perturbación sinusoidal.
- b. Se filtra la señal mediante un filtro ITAE de cuarto orden.
- c. Muestreo de la señal de (a).
- d. Muestreo de la señal de (b).

2.6. Aspectos prácticos de la elección del período de muestreo

La selección del período de muestreo en los sistemas muestreados es un problema fundamental que discutiremos varias veces en este libro. La elección adecuada depende de las propiedades de la señal, del método de reconstrucción y de la finalidad del sistema. En un problema de reconstrucción de señal puro, el propósito consiste simplemente en registrar una señal digitalmente y reconstruirla a partir de sus muestras. Un criterio razonable para la selección puede ser el valor del error entre la señal original y la señal reconstruida.

El teorema de muestreo de Shannon da una regla muy simple en el caso ideal de que en la reconstrucción se admite un gran retardo y cuando el contenido de frecuencia de la señal se limita a un determinado ancho de banda. En la práctica es, con frecuencia, necesario restringir el retardo en la reconstrucción de la señal. También es importante admitir que las señales pueden estar contaminadas con perturbaciones de alta frecuencia.

Elección del período de muestreo para procesamiento de señal

Considérese un problema de procesado de señales cuyo propósito puede expresarse en términos del error en la reconstrucción de la señal. Supongamos que la transformada de Fourier de la señal es nula para $|\omega| \geq \omega_0$. Si se acepta un retardo en la reconstrucción, el teorema de Shannon asegura que una frecuencia de muestreo de $2\omega_0$ es suficiente. Sin embargo, se necesitan velocidades de muestreo considerablemente mayores, si se limita el retardo en la reconstrucción.

Entonces es necesario emplear una reconstrucción causal como la proporcionada por un mantenedor de primer orden o de orden cero. El error de reconstrucción puede estimarse en este caso por (2.9) ó (2.11). El Ejemplo 2.3 ilustra los órdenes de magnitud típicos.

Ejemplo 2.3

Supongamos que la señal es una onda senoidal pura de frecuencia ω . Los errores máximos de la amplitud pico a pico para reconstrucciones con mantenedores de orden cero y primer orden están definidos por

$$e_0 = \frac{\omega h}{2} = \frac{\pi}{N}$$

$$e_1 = \frac{(\omega h)^2}{2} = \frac{2\pi^2}{N^2}$$

donde N es el número de muestras por período. Algunos valores numéricos típicos se muestran en la Tabla 2.2.

TABLA 2.2 Errores relativos obtenidos cuando se muestrea y se reconstruye una curva sinusoidal utilizando diferentes frecuencias de muestreo.

Muestras por período N	Error relativo máximo Mantenedor de orden cero	Error relativo máximo Mantenedor de primer orden
2	1.5	2.5
5	0.6	0.8
10	0.3	0.19
20	0.15	0.05
50	0.06	0.008
100	0.03	0.002
200	0.015	$5 \cdot 10^{-4}$
500	0.006	$8 \cdot 10^{-5}$

Para obtener un error relativo del 1 % para una reconstrucción con mantenedor de orden cero se necesita tomar alrededor de 300 muestras por período. La tabla muestra también que el efecto de utilizar un mantenedor de primer orden mejora apreciablemente, solamente si N es mayor de 20.

El resultado es similar para el muestreo y reconstrucción de otras señales.

Del ejemplo se deduce que la toma de varios cientos de muestras por período pueden justificarse en aplicaciones de procesamiento de señales.

Control en lazo cerrado

Una elección racional de la velocidad de muestreo en sistemas de control en lazo cerrado debe basarse en el conocimiento de su influencia en las características del

sistema de control. Parece razonable que las frecuencias de interés más altas, estén intimamente relacionadas con el ancho de banda del sistema en lazo cerrado. De este modo, la selección de la velocidad de muestreo pueden basarse en el ancho de banda, o lo que es equivalente, en el tiempo de subida del sistema en lazo cerrado. Las velocidades razonables de muestreo están entre seis y diez veces el ancho de banda, o dos o tres veces el tiempo de subida, lo cual puede parecer una velocidad baja en relación con el problema típico de procesamiento de señales. Comparativamente, las velocidades bajas de muestreo pueden utilizarse en problemas de control debido a que la dinámica de algunos sistemas tiene características de filtro pasa bajo y sus constantes de tiempo son normalmente mayores que los tiempos de respuesta en lazo cerrado. Entonces, la contribución a la salida de un período de muestreo depende del área del pulso; es comparativamente insensible a la forma del pulso.

La elección del período de muestreo se discutirá varias veces en este libro.

2.7. Resumen

En este capítulo se ha discutido el muestreado de señales continuas. Muestrear una señal, significa sustituirla por una secuencia de sus valores en una serie discreta de instantes de tiempo. Para muestreo periódico, la señal puede reconstruirse a partir de sus muestras si se verifican las siguientes condiciones:

La transformada de Fourier de la señal continua es nula para $|\omega| > \omega_1$.
La frecuencia de muestreo ω_s es mayor de $2\omega_1$.

Por tanto, cuando se trabaja con señales muestreadas, resulta muy importante elegir frecuencias de muestreo suficientemente altas en comparación con las frecuencias de interés más altas.

También es importante filtrar la señal continua antes de muestreala, de modo que las componentes de frecuencia superiores a la frecuencia de Nyquist $\omega_N = \omega_s/2$ no se transformen en componentes de baja frecuencia por el efecto de solapamiento.

2.8. Problemas

- 2.1. Esbozar la respuesta impulsional del reconstructor de Shannon definida por (2.1).
- 2.2. La señal

$$f(t) = a_1 \operatorname{sen} 2\pi t + a_2 \operatorname{sen} 20t$$

es la entrada a un circuito muestreador y mantenedor de orden cero. ¿Qué frecuencias hay en la salida si el período de muestreo es $h = 0.2$?

- 2.3. Una señal que se va a muestrear tiene un espectro como el indicado en la Figura 2.11. Interesan las frecuencias comprendidas entre 0 y f_1 Hz. Se tiene una perturbación de frecuencia fija $f_2 \approx 5f_1$. Discutir la elección del período de muestreo y el filtro premuestreo.

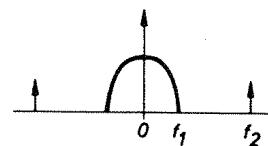


Figura 2.11

- 2.4. Demostrar que el sistema de la Figura 2.12 es una realización de un mantenedor de orden cero y determinar su respuesta a un pulso unidad y de duración un intervalo de muestreo.

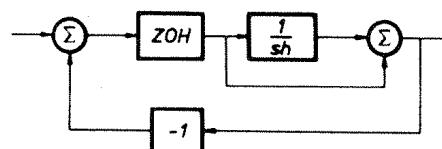


Figura 2.12

- 2.5. La Figura 2.13 muestra el valor del espectro de una señal. Esbozar el valor del espectro cuando la señal se muestrea con (a) $h=2\pi/10$ s, (b) $h=2\pi/20$ s. y (c) $h=2\pi/50$ s.

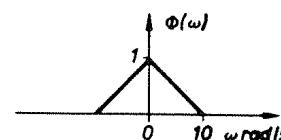


Figura 2.13

- 2.6. Considérese la señal del Problema 2.5, pero con el espectro centrado en $\omega = 100$ rad/seg y con a) $\omega_s = 120$ rad/s y b) $\omega_s = 240$ rad/s.
 2.7. Se utiliza una cámara fotográfica para obtener imágenes de una rueda que tiene dibujada una marca. La rueda gira a r revoluciones por segundo. La cámara toma una fotografía cada h segundos. Discutir cómo serán las imágenes cuando se muestren en una pantalla. (Comárese con el efecto de las ruedas móviles en las películas cinematográficas.)
 2.8. La señal $y(t) = \sin 3\pi t$ se muestrea con un período h . Calcular h para que la señal muestreada sea periódica.

2.9. Referencias

El hecho de que una sinusoides puede recuperarse a partir de sus muestras si se muestrea al menos dos veces por período fue enunciada en

NYQUIST, H. (1928): «Certain topics in Telegraph Transmission Theory», *AIEE Trans.*, 47, 617-44.

El teorema de muestreo en la forma presentada en este capítulo se introdujo en SHANNON, C. E. (1949): «Communications in the Presence of Noise», *Proc. IRE*, 37, 10-21.

en el cual se puso énfasis en el tema de comunicaciones. Sin embargo, el resultado ya se conocía anteriormente, como un teorema matemático. En la literatura soviética sobre comunicaciones, el teorema se introdujo en

KOTELNIKOV, V. A. (1933): «On the Transmission Capacity of "Ether" and Wire in Electrocommunications», *Proc. First All-Union Conference on Questions of Communications*. Moscow.

Una revisión del teorema de muestreo con gran número de referencias se da en

JERRI, A. J. (1977): «The Shannon Sampling Theorem—Its Various Extensions and Applications: A Tutorial Review», *Proc. IEEE*, 65, 1565-95.

Existen muchas formas de muestreo. Una revisión de los diferentes esquemas se da en JURY, E. I. (1961): «Sampling Schemes in Sampled-Data Control Systems», *IRE Trans., AC-6*, 88-90.

Diferentes tipos de circuitos mantenedores se discuten con detalle en

RAGAZZINI, J. R. y G. F. FRANKLIN (1958): *Sampled Data Control Systems*. New York: McGraw-Hill.

La selección del período de muestreo para procesamiento de señales se discute en

GARDENHIRE, L. W. (1964): «Selection of Sample Rates», *ISA Journal* (April), 59-64.

La elección de períodos de muestreo para problemas de control se discutirá más adelante y, por tanto, las referencias se darán en esa ocasión. Los diferentes compromisos que pueden darse entre las áreas de control y de procesamiento de señal pueden llevar a reglas muy diferentes para la elección de la velocidad de muestreo.

DEPARTAMENTO ELECTROTECNIA

Modelos matemáticos orientados al computador: sistemas discretos

Objetivo:

*Proporcionar descripciones matemáticas de los sistemas controlados por computador desde el punto de vista del computador.
Desarrollar los conceptos y métodos necesarios para trabajar con los sistemas discretos.*

3.1. Introducción

En este capítulo se introducen los modelos matemáticos para los sistemas controlados por computador. Se estudia el sistema tal como lo ve el computador. El computador recibe medidas desde el proceso en instantes discretos y transmite las nuevas señales de control en instantes discretos. El objetivo, entonces, es describir el cambio en las señales de un muestreo a otro y no tener en cuenta el comportamiento entre muestreos. Debe destacarse que los modelos matemáticos orientados al computador dan el comportamiento solamente en los puntos de muestreo; el proceso físico sigue siendo un sistema continuo. Sin embargo, enfocar el problema de esta forma simplifica mucho el tratamiento. En el Capítulo 4 se presenta una descripción de los modelos orientados al proceso, que tiene en cuenta el comportamiento continuo de éste.

Un punto que hay que tratar con algún cuidado es la dependencia del tiempo de un sistema muestreado (véase el Ejemplo 1.1). Este problema se trata también en el Capítulo 4. En el presente capítulo se elude el problema de la dependencia del tiempo estudiando las señales en unos instantes de tiempo que están sincronizados con el reloj del computador. Esto es lo que a veces se llama *modelo estroboscópico*, ya que sólo se obtienen observaciones intermitentes de las señales. Esto produce modelos descritos por ecuaciones en diferencias en las formas del espacio de estados y de entrada-salida. La Sección 3.2 tratará el problema de encontrar la representación discreta de un modelo continuo dado en el espacio de estados utilizando un mantenedor de orden cero. También se tratan la solución general de las ecuaciones forzadas en diferencias y el

problema inverso de encontrar el sistema continuo que corresponde a un sistema discreto dado. Las Secciones 3.3 y 3.4 estudian la transformación de los modelos en el espacio de estados y la conexión entre los modelos en el espacio de estados y de entrada-salida. Para describir los modelos de entrada-salida se emplean operadores de desplazamiento. El cálculo con el operador de desplazamiento es equivalente al empleo de operadores diferenciales en los sistemas continuos. El equivalente discreto de la transformada de Laplace es la transformada z , que se expone en la Sección 3.5.

El tratamiento de los modelos en el espacio de estados en la Sección 3.2 comprende el caso multivariable. El estudio de los modelos de entrada-salida, sin embargo, se restringe a los sistemas de una entrada y una salida. Se pueden hacer extensiones al caso multivariable, pero no se emplean en este libro porque requieren las matemáticas de las matrices polinomiales.

Para diseñar sistemas controlados por computador, es importante comprender cómo están relacionados los polos y ceros de los modelos continuos y discretos. Esto es lo que trata la Sección 3.6. En la Sección 3.7 se representa la selección del período de muestreo. Para ello se dan reglas empíricas basadas en el aspecto de las respuestas transitorias en términos de número de muestras por tiempo de subida.

3.2. Muestreo de un sistema continuo en el espacio de estados

Un problema fundamental es cómo describir un sistema continuo que está conectado a un computador por convertidores A-D y D-A. Consideremos el sistema de la Figura 3.1. Las señales en el computador son las secuencias $\{u(t_k)\}$ y $\{y(t_k)\}$. El

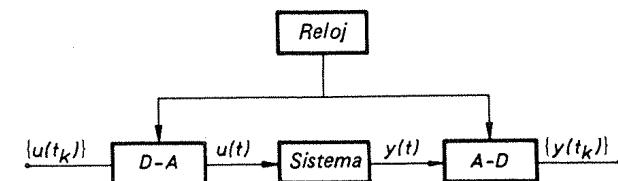


Figura 3.1 Diagrama de bloques de un sistema continuo conectado a unos convertidores A-D y D-A.

problema clave consiste en encontrar la relación entre estas dos secuencias. Encontrar el equivalente discreto de un sistema continuo es lo que se llama *muestreo de un sistema continuo*. El modelo obtenido se llama también modelo estroboscópico porque da la relación entre las variables del sistema sólo en los instantes de muestreo. Para obtener las descripciones deseadas es necesario describir los convertidores y el sistema.

Supongamos que el sistema continuo viene dado en el espacio de estados en la siguiente forma:

$$\begin{aligned}\frac{dx}{dt} &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t)\end{aligned}\quad (3.1)$$

El sistema tiene r entradas, p salidas y es de orden n .

Muestreo de un sistema con mantenedor de orden cero

Una situación corriente en control con computador es que el convertidor D-A está construido de tal forma que mantiene constante la señal analógica hasta que se ordena una nueva conversión. Por tanto, es natural tomar como instantes de muestreo, t_k , los momentos en que cambia el control. Como la señal de control es discontinua, es preciso especificar su comportamiento durante las discontinuidades. El convenio adoptado es que la señal es continua por la derecha. La señal de control se representa entonces por la señal muestreada $\{u(t_k) : k = \dots, -1, 0, 1, \dots\}$.

A continuación se determina la relación entre las variables del sistema en los instantes de muestreo. Dado el estado en el instante de muestreo t_k , el estado en algún instante futuro t , se obtiene resolviendo la ecuación (3.1).

$$x(t) = e^{A(t-t_k)}x(t_k) + \int_{t_k}^t e^{A(t-s)}Bu(s)ds' \quad (3.2)$$

El estado en el siguiente instante de muestreo, t_{k+1} , viene dado por

$$\begin{aligned}x(t_{k+1}) &= e^{A(t_{k+1}-t_k)}x(t_k) + \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-s)}Bu(s)ds' \\ &= e^{A(t_{k+1}-t_k)}x(t_k) + \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-s)}ds' Bu(t_k) \\ &= \Phi(t_{k+1}, t_k)x(t_k) + \Gamma(t_{k+1}, t_k)u(t_k)\end{aligned}$$

La segunda igualdad se deduce del hecho que u es constante entre los instantes de muestreo.

El vector de estado en el instante t_{k+1} es entonces una función lineal de $x(t_k)$ y $u(t_k)$. Si los convertidores A-D y D-A de la Figura 3.1 están perfectamente sincronizados y si los tiempos de conversión son despreciables, puede considerarse que la entrada u y la salida y se muestrean en el mismo instante. La ecuación de un sistema muestreado es entonces

$$\begin{cases} x(t_{k+1}) = \Phi(t_{k+1}, t_k)x(t_k) + \Gamma(t_{k+1}, t_k)u(t_k) \\ y(t_k) = Cx(t_k) + Du(t_k) \end{cases} \quad (3.3)$$

donde

$$\Phi(t_{k+1}, t_k) = e^{A(t_{k+1}-t_k)} \quad (3.4)$$

$$\Gamma(t_{k+1}, t_k) = \int_0^{t_{k+1}-t_k} e^{As} ds B \quad (3.5)$$

Entonces la relación entre las señales muestreadas puede expresarse por la ecuación lineal en diferencias, (3.3). Debe notarse que la ecuación (3.3) no encierra aproximación alguna. Da los valores exactos de las variables de estado y de la salida en los instantes de muestreo, ya que la señal de control es constante entre los instantes de muestreo. Al modelo representado por (3.3) se le llama, por tanto, *muestreo con mantenedor de orden cero del sistema* dado en (3.1). El sistema de (3.3) también puede llamarse el equivalente de (3.1) con mantenedor de orden cero.

En la mayor parte de los casos se tiene $D=0$. Una razón de esto es que, en los sistemas controlados por computador, se mide primero la salida y entonces se genera la señal de control $u(t_k)$ como una función de $y(t_k)$. En la práctica sucede a menudo que entre las conversiones A-D y D-A hay un retardo significativo. Sin embargo, es fácil realizar las modificaciones necesarias. Puede calcularse el vector de estado en instantes de tiempo entre los instantes de muestreo empleando (3.2). Esto permite investigar el comportamiento del sistema entre muestreos. Nótese que las respuestas entre los puntos de muestreos son parte de respuestas del sistema a saltos escalón, con condiciones iniciales no nulas. Esto implica, por ejemplo, que entre instantes de muestreo el sistema funciona en lazo abierto.

Para un muestreo periódico con período h ,

$$t_k = k \cdot h$$

el modelo dado en (3.3) se simplifica al sistema invariante

$$\begin{aligned}x(kh+h) &= \Phi x(kh) + \Gamma u(kh) \\ y(kh) &= Cx(kh) + Du(kh)\end{aligned} \quad (3.6)$$

donde

$$\begin{cases} \Phi = e^{Ah} \\ \Gamma = \int_0^h e^{As} ds B \end{cases} \quad (3.7)$$

Cómo calcular Φ y Γ

Los cálculos requeridos para muestrear un sistema continuo son la evaluación de una matriz exponencial y la integración de una matriz exponencial. Esto se puede hacer de muchas formas, por ejemplo, utilizando las siguientes técnicas:



Desarrollo en serie de la matriz exponencial.

La transformada de Laplace —la transformada de Laplace de $\exp(At)$ es $(sI - A)^{-1}$. El teorema de Cayley-Hamilton (véase el Apéndice B). Transformación a la forma de Jordan.

Los cálculos a mano son factibles para sistemas de bajo orden, $n \leq 2$, y para sistemas de orden superior con estructuras especiales. Una forma de simplificar los cálculos es calcular

$$\Psi = \int_0^h e^{At} ds = I h + \frac{A h^2}{2!} + \frac{A^2 h^3}{3!} + \cdots + \frac{A^i h^{i+1}}{(i+1)!} + \cdots \quad (3.8)$$

Las matrices Φ y Γ vienen dadas por

$$\begin{cases} \Phi = I + A\Psi \\ \Gamma = \Psi B \end{cases} \quad (3.9)$$

La evaluación con computador puede hacerse mediante diferentes algoritmos numéricos.

Ejemplo 3.1 – Integrador doble

El integrador doble (véase el Apéndice A) se describe por el sistema de ecuaciones

$$\begin{aligned} \dot{x} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}x + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u \\ y &= [1 \ 0]x \end{aligned} \quad (3.10)$$

Por tanto,

$$\begin{aligned} \Phi &= e^{Ah} = I + Ah + A^2 h^2/2 + \cdots \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & h \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \\ \Gamma &= \int_0^h \begin{bmatrix} s \\ 1 \end{bmatrix} ds = \begin{bmatrix} \frac{h^2}{2} \\ h \end{bmatrix} \end{aligned}$$

El modelo discreto de (3.10) es

$$\begin{aligned} x(kh + h) &= \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}x(kh) + \begin{bmatrix} \frac{h^2}{2} \\ h \end{bmatrix}u(kh) \\ y(kh) &= [1 \ 0]x(kh) \end{aligned} \quad (3.11)$$

Ejemplo 3.2 – Motor

Un modelo sencillo y normalizado de un motor eléctrico de corriente continua (véase el Ejemplo A.2 en el Apéndice A) es el dado por

$$\begin{aligned} \dot{x} &= \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}x + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u \\ y &= [0 \ 1]x. \end{aligned}$$

El método de la transformada de Laplace da

$$\begin{aligned} (sI - A)^{-1} &= \begin{bmatrix} s+1 & 0 \\ -1 & s \end{bmatrix}^{-1} = \frac{1}{s(s+1)} \begin{bmatrix} s & 0 \\ 1 & s+1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{s+1} & 0 \\ \frac{1}{s(s+1)} & \frac{1}{s} \end{bmatrix} \end{aligned}$$

Por tanto,

$$\Phi = e^{Ah} = \begin{bmatrix} e^{-h} & 0 \\ 1 - e^{-h} & 1 \end{bmatrix}$$

y

$$\Gamma = \int_0^h \begin{bmatrix} e^{-v} \\ 1 - e^{-v} \end{bmatrix} dv = \begin{bmatrix} 1 - e^{-h} \\ h - 1 + e^{-h} \end{bmatrix}$$

Solución de la ecuación del sistema

Los sistemas discretos invariantes se pueden describir mediante la ecuación en diferencias

$$\begin{aligned} x(k+1) &= \Phi x(k) + \Gamma u(k) \\ y(k) &= Cx(k) \end{aligned} \quad (3.12)$$

Por simplicidad se considera el periodo de muestreo como la unidad de tiempo, $h = 1$. Supongamos que tanto $x(k_0)$ como las señales de entrada $u(k_0), u(k_0+1), \dots$, están dadas. ¿Cómo evolucionará entonces el estado? Es posible resolver (3.12) por simple iteración.

$$\begin{aligned} x(k_0 + 1) &= \Phi x(k_0) + \Gamma u(k_0) \\ x(k_0 + 2) &= \Phi x(k_0 + 1) + \Gamma u(k_0 + 1) \\ &= \Phi^2 x(k_0) + \Phi \Gamma u(k_0) + \Gamma u(k_0 + 1) \\ &\vdots \end{aligned}$$

$$\begin{aligned} x(k) &= \Phi^{k-k_0} x(k_0) + \Phi^{k-k_0-1} \Gamma u(k_0) + \cdots + \Gamma u(k-1) \\ &= \Phi^{k-k_0} x(k_0) + \sum_{j=k_0}^{k-1} \Phi^{k-j-1} \Gamma u(j) \end{aligned} \quad (3.13)$$

La solución consta de dos partes: una depende de las condiciones iniciales y la otra es una suma ponderada de las señales de entrada.

El inverso del muestreo

El muestreo define una aplicación de los sistemas continuos, como (3.1), sobre los sistemas discretos, como (3.6). Un ejemplo sencillo nos muestra que la aplicación no siempre puede invertirse.

Ejemplo 3.3

No existe ninguna ecuación diferencial de primer orden que, al muestrearla, nos dé la ecuación en diferencias

$$x(kh + h) = -0.5x(kh) + u(kh)$$

porque la ecuación

$$e^{sh} = -0.5$$

no tiene solución real, ya que la exponencial siempre es positiva.

El modelo de (3.6) es entonces más general que (3.1). Sin embargo, si la matriz Φ no tiene valores propios en el eje real negativo, existe el correspondiente sistema continuo, que puede obtenerse resolviendo (3.7) y (3.9) para obtener A y B . De (3.7) se deduce que

$$A = \frac{\ln(\Phi)}{h} = \ln(\Phi^{1/h})$$

donde $\ln(\Phi)$ es la matriz función logarítmica. Nótese que algunas veces $\ln(\Phi)$ no es única. El cálculo de $\ln(\Phi)$ se detalla en el Apéndice B. La ecuación (3.9) da

$$B = \Psi^{-1}\Gamma$$

Si el sistema no contiene ningún integrador, entonces $\Phi - I$ es no singular, y de (3.9) se deduce que

$$B = \Psi^{-1}\Gamma = (A\Psi)^{-1}A\Gamma = (\Phi - I)^{-1}A\Gamma \quad (3.14)$$

donde Ψ viene dada por (3.8).

El muestreo inverso de un oscilador armónico da otra visión del problema del enmascaramiento presentado en la Sección 2.5.

Ejemplo 3.4 – El oscilador armónico

El sistema discreto

$$x(kh + h) = \begin{bmatrix} \cos \alpha h & \sin \alpha h \\ -\sin \alpha h & \cos \alpha h \end{bmatrix} x(kh) + \begin{bmatrix} 1 - \cos \alpha h \\ \sin \alpha h \end{bmatrix} u(kh)$$

puede obtenerse muestreando un sistema continuo con

$$A = \begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 \\ \beta \end{bmatrix}$$

donde

$$\beta = \alpha + \frac{2\pi}{h} \cdot n \quad n = 0, 1, \dots$$

En este caso el problema inverso tiene muchas soluciones (compárense los Ejemplos A.3 y B.1). Esto es lo que ocurre generalmente si la matriz Φ tiene valores propios complejos. Obsérvese que siempre existe un único β en el intervalo $-\omega_N \leq \beta \leq \omega_N$, donde $\omega_N = \pi/h$ es la frecuencia de Nyquist asociada con el período de muestreo h .

Muestreo de un sistema con retardo temporal

Los retardos temporales son corrientes en los modelos matemáticos de los procesos industriales. La teoría de los sistemas continuos con retardos es complicada debido a que los sistemas tienen dimensión infinita.

Vamos a muestrear un sistema compuesto de un retardo y un sistema continuo en el espacio de estados. Sea el sistema descrito por

$$\dot{x} = Ax(t) + Bu(t - \tau) \quad (3.15)$$

Se supone en principio que el retardo τ es menor que el período de muestreo. Vamos a calcular el muestreo con mantenedor de orden cero del sistema (3.15).

La integración de (3.15) durante un período de muestreo da

$$x(kh + h) = e^{th}x(kh) + \int_{kh}^{kh+h} e^{A(kh+h-s')}Bu(s' - \tau) ds' \quad (3.16)$$

Debido a que la señal $u(t)$ es constante a tramos sobre el intervalo de muestreo, la señal retardada $u(t - \tau)$ también lo es. La señal retardada, sin embargo, cambiará entre los instantes de muestreo (véase la Fig. 3.2). Para evaluar la integral de (3.16), conviene descomponer el intervalo de integración en dos partes de forma que $u(t - \tau)$ sea constante en cada parte. Por tanto,

$$\begin{aligned} \int_{kh}^{kh+h} \dots &= \int_{kh}^{kh+\tau} e^{A(kh+h-s')} B ds' u(kh - h) \\ &+ \int_{kh+\tau}^{kh+h} e^{A(kh+h-s')} B ds' u(kh) \\ &= \Gamma_1 u(kh - h) + \Gamma_0 u(kh) \end{aligned}$$

El muestreo del sistema continuo de (3.15) da entonces

$$x(kh + h) = \Phi x(kh) + \Gamma_0 u(kh) + \Gamma_1 u(kh - h) \quad (3.17)$$

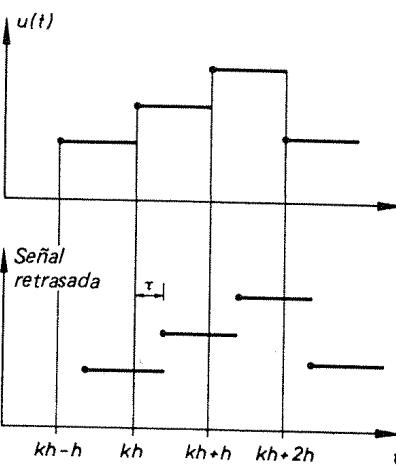


Figura 3.2 Relación entre $u(t)$, la señal retrasada $u(t - \tau)$ y los instantes de muestreo.

$$\Phi = e^{At} = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}$$

$$\Gamma_1 = e^{A(h-\tau)} \int_0^{\tau} e^{As} ds B = \begin{bmatrix} 1 & h-\tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\tau^2}{2} \\ \tau \end{bmatrix} = \begin{bmatrix} \tau(h - \frac{\tau}{2}) \\ \tau \end{bmatrix}$$

$$\Gamma_0 = \int_0^{h-\tau} e^{As} ds B = \begin{bmatrix} \frac{(h-\tau)^2}{2} \\ h-\tau \end{bmatrix}$$

Retardos de tiempo más largos

Si el retardo es mayor que h , hay que modificar un poco el análisis anterior. Si

$$\tau = (d - 1)h + \tau' \quad 0 < \tau' \leq h$$

donde d es un número entero, se obtiene la siguiente ecuación:

$$x(kh + h) = \Phi x(kh) + \Gamma_0 u(kh - dh + h) + \Gamma_1 u(kh - dh) \quad (3.20)$$

donde Γ_0 y Γ_1 están dados por (3.18) y (3.19) con τ sustituido por τ' . La descripción correspondiente en el espacio de estados es

$$\begin{bmatrix} x(kh + h) \\ u(kh - dh + h) \\ \vdots \\ u(kh - h) \\ u(kh) \end{bmatrix} = \begin{bmatrix} \Phi & \Gamma_1 & \Gamma_0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x(kh) \\ u(kh - dh) \\ \vdots \\ u(kh - 2h) \\ u(kh - h) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I \end{bmatrix} u(kh) \quad (3.21)$$

Hay que hacer notar que si $\tau > 0$ hay que utilizar $d \cdot r$ variables de estado extra para describir el retardo, donde r es el número de entradas. El polinomio característico de la descripción en el espacio de estados es $\lambda^{dr} A(\lambda)$, donde $A(\lambda)$ es el polinomio característico de Φ .

Un ejemplo ilustra el uso de la fórmula general.

Ejemplo 3.6 – Modelo sencillo de una máquina de papel

Determinar el equivalente muestrado con mantenedor de orden cero del sistema

$$\dot{x}(t) = -x(t) + u(t - 2.5)$$

con un intervalo de muestreo $h = 1$. En este caso $d = 3$ y $\tau' = 0.5$ y (3.17) cambiará de la forma siguiente:

$$x(k + 1) = \Phi x(k) + \Gamma_0 u(k - 2) + \Gamma_1 u(k - 3)$$

donde

$$\Phi = e^{At}$$

$$\Gamma_1 = e^{A(h-\tau)} \int_0^{\tau} e^{As} ds B \quad (3.18)$$

$$\Gamma_0 = \int_0^{h-\tau} e^{As} ds B \quad (3.19)$$

Un modelo en el espacio de estados de (3.17) es el dado por

$$\begin{bmatrix} x(kh + h) \\ u(kh) \end{bmatrix} = \begin{bmatrix} \Phi & \Gamma_1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x(kh) \\ u(kh - h) \end{bmatrix} + \begin{bmatrix} \Gamma_0 \\ I \end{bmatrix} u(kh)$$

Hay que notar que se han introducido r variables de estado extra, $u(kh - h)$, que representan los valores anteriores de la señal de control. El sistema continuo de (3.15) es de dimensión infinita; el sistema muestrado correspondiente, sin embargo, es un sistema de dimensión finita. Por tanto, los retardos son mucho más sencillos de manejar si el sistema está muestrado, por la razón siguiente: para especificar el estado del sistema es necesario almacenar la entrada durante un intervalo de tiempo igual al retardo. Con la reconstrucción del mantenedor de orden cero, la señal de entrada siempre puede representarse por un número finito de valores.

Ejemplo 3.5 – Integrador doble con retardo

Consideremos el integrador doble del Ejemplo 3.1 e introduzcamos un retardo $0 \leq \tau \leq h$. Entonces

donde

$$\Phi = e^{-1} \approx 0.37$$

$$\Gamma_0 = \int_0^{0.5} e^{-s} ds = 1 - e^{-0.5} \approx 0.39$$

$$\Gamma_1 = e^{-0.5} \int_0^{0.5} e^{-s} ds = e^{-0.5} - e^{-1} \approx 0.24$$

Otros tipos de mantenedor para la señal de control

Para calcular el sistema muestrado a partir del sistema continuo debe conocerse la forma de la señal de control durante el intervalo de muestreo. Estos cálculos son particularmente fáciles de hacer cuando se emplean mantenedores de orden cero, ya que la señal de control es continua a trozos. El muestreo con mantenedor de orden cero es también el más común en la práctica, ya que se instala muy fácilmente empleando convertidores D-A ordinarios. En los ejercicios se dan ejemplos en los que se emplean otros tipos de mantenedores. Un caso típico es un sistema con actuadores hidráulicos, donde se necesitan señales de control suaves para evitar transitorios de presión.

A menudo los sistemas a controlar tienen características de paso bajo. Entonces la forma de la señal de control durante el período de muestreo no es crucial. Si el período de muestreo es pequeño comparado con las constantes de tiempo del sistema, la respuesta está determinada fundamentalmente por la integral de la señal de control durante el período de muestreo.

De (3.5) se deduce que un cambio en la forma de la señal de control cambiará Γ pero no la matriz de transición Φ . Esto se ve claramente con un ejemplo.

Ejemplo 3.7

Consideremos un integrador doble que se ha muestreado con un período h . Supongamos que el circuito del mantenedor es tal que la señal de control es $\alpha u(k)$ durante la primera mitad del período de muestreo y $\beta u(k)$ durante la segunda mitad. De aquí se tiene

$$\Phi = e^{Ah} = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}$$

$$\begin{aligned} \Gamma &= \alpha \int_0^{h/2} e^{A(h-s)} B ds + \beta \int_{h/2}^h e^{A(h-s)} B ds \\ &= \frac{1}{8} \begin{bmatrix} 3\alpha h^2 + \beta h^2 \\ 4\alpha h + 4\beta h \end{bmatrix} \end{aligned}$$

Comportamiento entre instantes de muestreo

Los modelos discretos (3.3) y (3.6) dan los valores de las variables de estado y de las salidas en los instantes de muestreo $\{t(k)\}$. Pero también interesan los valores de

las variables entre los instantes de muestreo. Estos valores vienen dados por (3.2), que puede escribirse en la forma

$$x(t) = \Phi(t, t_k)x(t_k) + \Gamma(t, t_k)u(t_k) \quad (3.22)$$

donde

$$\Phi(t, t_k) = e^{A(t-t_k)} \quad (3.23)$$

$$\Gamma(t, t_k) = \int_{t_k}^t e^{As} ds B \quad (3.24)$$

3.3. Transformación de los modelos del espacio de estados

En la Sección 3.2 se ha descrito un sistema discreto por el modelo de la ecuación (3.2). Lo mismo que en los sistemas continuos, se pueden introducir nuevas coordenadas en el espacio de estados.

Supongamos que T es una matriz no singular y definamos un nuevo vector de estado $z(k) = Tx(k)$. Entonces

$$\begin{aligned} z(k+1) &= Tx(k+1) = T\Phi x(k) + T\Gamma u(k) = T\Phi T^{-1}z(k) + T\Gamma u(k) \\ &= \tilde{\Phi}z(k) + \tilde{\Gamma}u(k) \end{aligned}$$

y

$$y(k) = Cx(k) + Du(k) = CT^{-1}z(k) + Du(k) = \tilde{C}z(k) + Du(k)$$

Las matrices Φ , Γ y C dependen entonces del sistema de coordenadas escogido para representar el estado. Por eso interesa conocer los invariantes frente a la transformación.

Teorema 3.1 La ecuación característica

$$\det[\lambda I - \Phi] = 0$$

es invariante cuando se introducen nuevos estados mediante una matriz de transformación T no singular.

Demostración

$$\begin{aligned} \det[\lambda I - \tilde{\Phi}] &= \det[\lambda TT^{-1} - T\Phi T^{-1}] = \det T \det[\lambda I - \Phi] \det T^{-1} \\ &= \det[\lambda I - \Phi] \end{aligned}$$

□

Las coordenadas pueden escogerse de forma que se obtengan formas sencillas de las ecuaciones del sistema.

Forma diagonal

Supongamos que Φ tiene valores propios distintos. Entonces existe una matriz T tal que

$$T\Phi T^{-1} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

donde λ_i son los valores propios de Φ . El cálculo de T se presenta en la Sección 5.3. En este caso se obtiene un conjunto de ecuaciones en diferencias de primer orden desacopladas.

$$\begin{aligned} z_1(k+1) &= \lambda_1 z_1(k) + \beta_1 u(k) \\ &\vdots \\ z_n(k+1) &= \lambda_n z_n(k) + \beta_n u(k) \\ y(k) &= \gamma_1 z_1(k) + \cdots + \gamma_n z_n(k) \end{aligned}$$

Ahora la solución del sistema de ecuaciones es simple. Cada modo tendrá la solución

$$z_i(k) = \lambda_i^k z_i(0) + \sum_{j=0}^{k-1} \lambda_i^{k-j-1} \beta_j u(j) \quad (3.25)$$

Si Φ tiene valores propios múltiples, entonces no es, en general, posible diagonalizar Φ . Cualquier matriz puede, sin embargo, transformarse a una *forma de Jordan*. En esta forma la matriz transformada, $\tilde{\Phi}$, tiene los valores propios en la diagonal y algunos unos en la línea situada sobre la diagonal.

Forma observable

Supongamos que Φ tiene la ecuación característica

$$\det[\lambda I - \Phi] = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n = 0 \quad (3.26)$$

y que

$$W_o = \begin{bmatrix} C \\ C\Phi \\ \vdots \\ C\Phi^{n-1} \end{bmatrix} \quad (3.27)$$

es no singular. Entonces existe una matriz de transformación tal que el sistema transformado es de la forma

$$\begin{aligned} z(k+1) &= \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ -a_{n-1} & 0 & 0 & \cdots & 1 \\ -a_n & 0 & 0 & \cdots & 0 \end{bmatrix} z(k) + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix} u(k) \\ y(k) &= [1 \ 0 \ \cdots \ 0] z(k) \end{aligned} \quad (3.28)$$

a la que se llama *forma canónica observable*. Esta forma tiene la ventaja de que a partir de ella es fácil encontrar el modelo entrada-salida y determinar un observador adecuado. Una expresión explícita de la matriz de transformación se da en la Sección 5.3.

Forma controlable

Supongamos que la ecuación característica de Φ es (3.26) y que

$$W_c = [\Gamma \ \Phi\Gamma \ \cdots \ \Phi^{n-2}\Gamma \ \Phi^{n-1}\Gamma] \quad (3.29)$$

es no singular. Entonces existe una transformación tal que el sistema transformado es

$$\begin{aligned} z(k+1) &= \begin{bmatrix} -a_1 & -a_2 & \cdots & -a_{n-1} & -a_n \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} z(k) + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u(k) \\ y(k) &= [b_1 \ \cdots \ b_n] z(k) \end{aligned} \quad (3.30)$$

que recibe el nombre de *forma canónica controlable*. La ventaja de esta forma es que mediante ella también es fácil de calcular el modelo entrada-salida y la ley de control de realimentación. En la Sección 5.3 se da una expresión explícita de la matriz de transformación.

Nota. Las formas observables y controlables también reciben el nombre de *formas asociadas*.

3.4. Modelos entrada-salida

Un sistema dinámico puede describirse utilizando modelos internos o externos. Los modelos internos —por ejemplo, los modelos del espacio de estados vistos en la Sección 3.2— describen los acoplamientos internos entre las variables del sistema. Los modelos externos dan sólo la relación entre la entrada y la salida del sistema.

En esta sección se muestra en primer lugar cómo la relación entrada-salida de un sistema lineal de tipo general puede expresarse por una función de respuesta discreta. Se ve entonces que puede emplearse el cálculo del operador de desplazamiento para reducir directamente la relación entrada-salida, lo que lleva a la caracterización del comportamiento entrada-salida en términos de operadores de transferencia discreta.

La respuesta discreta

Consideremos un sistema discreto con una entrada y una salida. Las señales de entrada y salida sobre un intervalo finito pueden representarse como vectores de dimensión finita

$$\begin{aligned} U &= [u(t_0) \dots u(t_{N-1})]^T \\ Y &= [y(t_0) \dots y(t_{N-1})]^T \end{aligned}$$

El modelo lineal general que relaciona Y con U puede expresarse entonces como

$$Y = \tilde{H}U + Y_p$$

donde \tilde{H} es una matriz $N \times N$. Y_p incluye las condiciones iniciales. Si la relación entre U e Y es *causal*, la matriz \tilde{H} debe ser triangular inferior. El km -ésimo elemento, $\tilde{h}(k, m)$, de \tilde{H} es entonces cero si $m > k$. La relación entrada-salida para un sistema lineal general puede escribirse en la forma

$$y(t_k) = \sum_{m=0}^k \tilde{h}(k, m)u(t_m) + y_p(t_k) \quad (3.31)$$

donde el término y_p se introduce para tener en cuenta las condiciones iniciales en el sistema. La función $\tilde{h}(k, m)$ se llama la *respuesta discreta*, o la *función de ponderación*, del sistema. La función de respuesta discreta es una representación muy conveniente, porque puede medirse directamente introduciendo un pulso de magnitud unidad y de anchura la del intervalo de muestreo y registrando la salida. Para condiciones iniciales cero el valor $\tilde{h}(k, m)$ de la respuesta discreta da la salida en el instante t_k para un pulso unidad en el instante t_m . Para sistemas con varias entradas y salidas, la respuesta discreta es simplemente una función matricial.

Para sistemas invariantes, la respuesta discreta es solamente función de $k - m$; esto es,

$$\tilde{h}(k, m) = h(k - m)$$

Es fácil calcular la respuesta discreta de un sistema definido por el modelo en el espacio de estados dado en (3.12). De (3.13) se deduce que

$$y(k) = C\Phi^{k-k_0}x(k_0) + \sum_{j=k_0}^{k-1} C\Phi^{k-j-1}\Gamma u(j)$$

La función de respuesta discreta para un sistema discreto es entonces

$$h(k) = \begin{cases} 0 & k < 1 \\ C\Phi^{k-1}\Gamma & k \geq 1 \end{cases} \quad (3.32)$$

La respuesta discreta es la suma de funciones de la forma

$$\operatorname{Re}\{P(k)\lambda_i^k\}$$

donde P es un polinomio en k y λ_i son los valores propios de la matriz Φ .

La respuesta discreta tiene la siguiente propiedad.

Teorema 3.2. La respuesta discreta de (3.32) es invariante respecto a las transformaciones de coordenadas del modelo en el espacio de estados.

Demostración. Introduzcamos las nuevas coordenadas $z = Tx$. La respuesta discreta del sistema transformado es entonces

$$\begin{aligned} \tilde{h}(k) &= \tilde{C}\tilde{\Phi}^{k-1}\tilde{\Gamma} = (CT^{-1})(T\Phi T^{-1})^{k-1}TT \\ &= (CT^{-1})T\Phi^{k-1}T^{-1}TT = C\Phi^{k-1}\Gamma = h(k) \end{aligned} \quad \square$$

Cálculo del operador de desplazamiento

El cálculo con el operador diferencial es una herramienta muy adecuada para manipular ecuaciones diferenciales lineales con coeficientes constantes. Para los sistemas descritos por ecuaciones lineales en diferencias puede desarrollarse un cálculo con un operador análogo. En el desarrollo del cálculo con operador los sistemas se ven como operadores que aplican las señales de entrada sobre las señales de salida. Para especificar un operador es necesario dar su rango; es decir, definir la clase de señales de entrada y describir cómo actúa el operador sobre las señales. En el cálculo con operador todas las señales se consideran como secuencias doblemente infinitas $\{f(k): k = \dots, -1, 0, 1, \dots\}$. Por conveniencia se toma el período de muestreo como la unidad de tiempo.



El operador de desplazamiento hacia delante se representa por q . Tiene la propiedad

$$qf(k) = f(k + 1)$$

Si la norma de la señal se define como

$$\|f\| = \sup_k |f(k)|$$

o

$$\|f\|^2 = \sum_{k=-\infty}^{\infty} f^2(k)$$

se tiene que el operador de desplazamiento tiene norma unidad. Esto significa que el cálculo con operadores de desplazamiento es más fácil que con operadores diferenciales, ya que el operador diferencial no está acotado. El inverso del operador de desplazamiento hacia delante recibe el nombre de *operador de desplazamiento hacia atrás* y se representa por q^{-1} . Por tanto,

$$q^{-1}f(k) = f(k - 1)$$

Nótese que es importante para el rango del operador que se trate de secuencias doblemente infinitas; en caso contrario el inverso del operador de desplazamiento hacia adelante puede no existir. En la discusión de los problemas relacionados con la ecuación característica de un sistema, como la estabilidad y el orden del sistema, es más conveniente emplear el operador de desplazamiento hacia delante, mientras que para los problemas relacionados con la causalidad, interesa más emplear el operador de desplazamiento hacia atrás.

El cálculo con operadores proporciona descripciones compactas de los sistemas y permite deducir fácilmente relaciones entre las variables del sistema, ya que la manipulación de las ecuaciones en diferencias se reduce a un problema puramente algebraico. En muchos libros de texto se emplea z para el operador de desplazamiento y también para la variable compleja en la transformada z . Sin embargo, es conveniente utilizar diferentes notaciones para los dos conceptos. Es la misma distinción que se hace normalmente entre la variable compleja s en la transformada de Laplace y el operador diferencial $p = d/dt$.

El operador de desplazamiento se emplea para simplificar la manipulación de ecuaciones en diferencias de orden superior. Consideremos la ecuación

$$y(k + na) + a_1y(k + na - 1) + \dots + a_ny(k) = b_0u(k + nb) + \dots + b_nu(k) \quad (3.33)$$

donde $na \geq nb$.

Empleando el operador desplazamiento tenemos

$$(q^{na} + a_1q^{na-1} + \dots + a_n)y(k) = (b_0q^{nb} + \dots + b_n)u(k)$$

Con la introducción de los polinomios

$$A(z) = z^{na} + a_1z^{na-1} + \dots + a_n$$

y

$$B(z) = b_0z^{nb} + b_1z^{nb-1} + \dots + b_n$$

la ecuación en diferencias se puede escribir como

$$A(q)y(k) = B(q)u(k) \quad (3.34)$$

Cuando es necesario, se puede indicar el grado del polinomio por un subíndice; por ejemplo, $A_{na}(q)$.

La ecuación (3.34) puede expresarse también en función del operador de desplazamiento hacia atrás. Nótese que la ecuación (3.33) puede escribirse como

$$y(k) + a_1y(k - 1) + \dots + a_ny(k - na) = b_0u(k - d) + \dots + b_nu(k - d - nb)$$

donde $d = na - nb$ es el *exceso de polos* del sistema.

El polinomio

$$A^*(z) = 1 + a_1z + \dots + a_nz^{na} = z^{na}A(z^{-1})$$

se obtiene a partir del polinomio A invirtiendo el orden de los coeficientes. Este polinomio recibe el nombre de *polinomio recíproco*. La introducción de los polinomios recíprocos permite escribir el sistema (3.33) en la forma

$$A^*(q^{-1})y(k) = B^*(q^{-1})u(k - d) \quad (3.35)$$

Hay que tener cuidado cuando se opera con polinomios recíprocos porque A^{**} no es necesariamente igual que A . El polinomio

$$A(z) = z$$

tiene como reciproco

$$A^*(z) = 1$$

El reciproco de este es

$$A^{**}(z) = 1$$

que es claramente diferente de A .

Una dificultad

El objetivo de la teoría de sistemas algebraicos es convertir las manipulaciones de ecuaciones en diferencias en problemas puramente algebraicos. De la definición del



operador de desplazamiento se sigue que la ecuación en diferencias dada en (3.34) puede multiplicarse por potencias de q , que significa solamente un desplazamiento del tiempo hacia delante. Las ecuaciones de tiempos desplazados pueden también multiplicarse por números reales y sumarse, lo que equivale a multiplicar la ecuación (3.34) por un polinomio en q . Si (3.34) se verifica, también se verificará

$$C(q)A(q)y(k) = C(q)B(q)u(k)$$

Para obtener un álgebra conveniente, es útil poder dividir una ecuación como (3.34) por un polinomio en q . Por ejemplo, si

$$A(q)y(k) = 0$$

se puede concluir que

$$y(k) = 0$$

Si es posible realizar la división, una ecuación como (3.34) puede resolverse con respecto a $y(k)$. Un ejemplo sencillo nos muestra que no es posible realizar la división por un polinomio en q excepto si se hacen algunas hipótesis especiales.

Ejemplo 3.8

Consideremos la ecuación en diferencias

$$y(k+1) - ay(k) = u(k)$$

donde $|a| < 1$. Empleando la notación de operadores la ecuación puede escribirse

$$(q - a)y(k) = u(k)$$

Si $y(k_0) = y_0$, de (3.13) se deduce que la solución puede escribirse en la forma

$$\begin{aligned} y(k) &= a^{k-k_0}y_0 + \sum_{j=k_0}^{k-1} a^{k-j-1}u(j) \\ &= a^{k-k_0}y_0 + \sum_{i=1}^{k-k_0} a^{i-1}u(k-i) \end{aligned} \quad (3.36)$$

Una solución formal de la ecuación de operadores puede obtenerse de la siguiente forma:

$$y(k) = \frac{1}{q-a}u(k) = \frac{q^{-1}}{1-aq^{-1}}u(k)$$

Debido a que q^{-1} tiene norma unidad, la parte derecha puede expresarse como una serie convergente.

$$\begin{aligned} y(k) &= q^{-1}(1 + aq^{-1} + a^2q^{-2} + \dots)u(k) \\ &= \sum_{i=1}^{\infty} a^{i-1}u(k-i) \end{aligned} \quad (3.37)$$

Es evidente que las soluciones presentadas en (3.36) y (3.37) sólo serán iguales si se hacen hipótesis adicionales.

Es posible desarrollar un álgebra de operadores que permita la división por un polinomio arbitrario en q si se supone que existe algún k_0 tal que todas las secuencias

son cero para $k \leq k_0$. Este álgebra permite entonces las manipulaciones normales de multiplicación y división de ecuaciones por polinomios del operador desplazamiento, así como la suma y resta de ecuaciones. Sin embargo, la hipótesis no implica que todas las condiciones iniciales para las ecuaciones en diferencias sean cero, que es el convenio empleado en este libro. (Compárese con el ejemplo.)

Si no se realiza ninguna suposición sobre la secuencia de entrada, es posible desarrollar un álgebra del operador de desplazamiento ligeramente diferente que permite la división sólo por polinomios con ceros dentro del círculo unidad. Esto se corresponde con el hecho de que los efectos de las condiciones iniciales sobre los modos estables terminan desapareciendo. Este álgebra es un poco más complicada debido a que no admite la división habitual.

El operador de transferencia discreta

El empleo del cálculo de operador permite expresar de forma conveniente la relación entrada-salida como una función racional tanto del operador de desplazamiento hacia delante como hacia atrás. Esta función recibe el nombre de *operador de transferencia discreta* y se puede obtener fácilmente a partir de cualquier descripción del sistema eliminando las variables internas mediante puras manipulaciones algebraicas.

Consideremos, por ejemplo, el modelo del espacio de estados dado en (3.6). Para obtener la relación entrada-salida debemos eliminar el vector de estado. De (3.6) se deduce que

$$x(k+1) = qx(k) = \Phi x(k) + \Gamma u(k)$$

Por tanto,

$$(qI - \Phi)x(k) = \Gamma u(k)$$

Esto da

$$y(k) = Cx(k) + Du(k) = [C(qI - \Phi)^{-1}\Gamma + D]u(k)$$

El operador de transferencia discreta para el sistema (3.6) viene dado entonces por

$$H(q) = C(qI - \Phi)^{-1}\Gamma + D$$

El operador de transferencia discreta puede expresarse también en términos del operador de desplazamiento hacia atrás.

$$H^*(q^{-1}) = C(I - q^{-1}\Phi)^{-1}q^{-1}\Gamma + D = H(q)$$

El operador de transferencia discreta del sistema (3.6) es entonces una matriz cuyos elementos son funciones racionales de q . Para un sistema con una entrada y una salida,

$$H(q) = C(qI - \Phi)^{-1}\Gamma + D = B(q)/A(q)$$

Si el vector de estado es de dimensión n y si los polinomios $A(q)$ y $B(q)$ no tienen factores comunes, entonces el polinomio A es de grado n . De (3.38) se deduce que el



polinomio A es también el polinomio característico de la matriz Φ , lo que significa que el modelo entrada-salida se puede escribir en la forma

$$y(k) + a_1 y(k-1) + \cdots + a_n y(k-n) = b_0 u(k) + \cdots + b_n u(k-n)$$

El caso más corriente en los sistemas controlados por computador es que $b_0=0$; esto es, no hay término independiente en el modelo discreto. Normalmente se mide primero $y(k)$, y luego se determina $u(k)$. En este caso $y(k)$ no puede ser influenciado por $y(k)$ incluso aunque el sistema continuo tenga un término independiente.

Ejemplo 3.9

Consideremos el integrador doble del Ejemplo 3.1 con $h=1$. De (3.38) se tiene

$$\begin{aligned} H(q) &= [1 \ 0] \begin{bmatrix} q-1 & -1 \\ 0 & q-1 \end{bmatrix}^{-1} \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \frac{0.5(q+1)}{(q-1)^2} \\ &= \frac{0.5(q^{-1} + q^{-2})}{1 - 2q^{-1} + q^{-2}} \end{aligned}$$

Ejemplo 3.10

Utilicemos $h=1$ para el integrador doble e introduzcamos un retardo temporal de 0,5 s. De (3.17) y del Ejemplo 3.5 se tiene

$$\begin{aligned} H(q) &= C(qI - \Phi)^{-1}(\Gamma_0 + \Gamma_1 q^{-1}) \\ &= [1 \ 0] \begin{bmatrix} q-1 & 1 \\ 0 & q-1 \end{bmatrix}^{-1} \begin{bmatrix} 0.125 + 0.375q^{-1} \\ 0.5 + 0.5q^{-1} \end{bmatrix} \\ &= \frac{0.125(q^2 + 6q + 1)}{q(q^2 - 2q + 1)} = \frac{0.125(q^{-1} + 6q^{-2} + q^{-3})}{1 - 2q^{-1} + q^{-2}} \end{aligned}$$

Ejemplo 3.11

Consideremos el siguiente sistema, que está dado en forma canónica observable:

$$\begin{aligned} x(k+1) &= \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix} x(k) + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(k) \\ y(k) &= [1 \ 0] x(k) \end{aligned}$$

El operador de transferencia discreta es

$$\begin{aligned} H(q) &= [1 \ 0] \begin{bmatrix} q + a_1 & -1 \\ a_2 & q \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \\ &= \frac{1}{q^2 + a_1 q + a_2} [1 \ 0] \begin{bmatrix} q & 1 \\ -a_2 & q + a_1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \\ &= \frac{b_1 q + b_2}{q^2 + a_1 q + a_2} = \frac{b_1 q^{-1} + b_2 q^{-2}}{1 + a_1 q^{-1} + a_2 q^{-2}} \end{aligned}$$

Entonces los coeficientes a_i y b_i de la forma canónica definen los polinomios A y B , respectivamente. Esto es cierto también para sistemas de orden n , tanto en forma observable como controlable.

La Sección 3.3 muestra que pueden emplearse diferentes representaciones en el espacio de estados. Por supuesto, sin cambiar esto en absoluto el modelo entrada-salida.

Teorema 3.3. El operador de transferencia discreta $H(q)$ para el modelo en el espacio de estados (3.6) es independiente de la representación en el espacio de estados.

Demostración. Consideremos el operador de transferencia discreta

$$H(q) = C(qI - \Phi)^{-1}\Gamma$$

y una matriz de transformación T . En las nuevas coordenadas

$$\begin{aligned} \tilde{H}(q) &= \tilde{C}(qI - \tilde{\Phi})^{-1}\tilde{\Gamma} = CT^{-1}(qTT^{-1} - T\Phi T^{-1})^{-1}TT \\ &= CT^{-1}[T(qI - \Phi)T^{-1}]^{-1}T\Gamma = CT^{-1}T(qI - \Phi)^{-1}T^{-1}T\Gamma \\ &= C(qI - \Phi)^{-1}\Gamma = H(q) \end{aligned} \quad \square$$

Los modelos entrada-salida de un sistema con un mantenedor de orden cero pueden obtenerse utilizando las expresiones (3.7) y (3.38). Para simplificar el cálculo del operador de transferencia discreta $H(q)$ es conveniente emplear la Tabla 3.1, que nos da $H(q)$ para algunos sistemas típicos.

Polos y ceros

Los *polos* de un sistema son los ceros del denominador de $H(q)$, el polinomio característico. Los *ceros* se obtienen a partir de $B(q)=0$, los polos del sistema inverso. Por ejemplo, el sistema del Ejemplo 3.9 tiene un cero en -1 ; el sistema tiene dos polos en 1 .

El retardo en un sistema da lugar a polos en el origen. El sistema del Ejemplo 3.10 tiene tres polos: dos en 1 y uno en el origen. Hay dos ceros: $-3 \pm \sqrt{8}$.

El orden de un sistema

El *orden* de un sistema es lo mismo que la dimensión de una representación en el espacio de estados o, lo que es equivalente, el número de polos del sistema.

Nótese que, debido a los retardos, es importante emplear la forma del operador de desplazamiento hacia delante para determinar el orden. La determinación de los polos, ceros y del orden de un sistema son casos en que es importante emplear la forma correspondiente al operador de desplazamiento hacia adelante.

3.5. La transformada Z

La transformada discreta análoga a la de Laplace es la transformada Z ; una herramienta muy apropiada para estudiar las ecuaciones en diferencias lineales con o

TABLA 3.1 Muestreo de un sistema continuo $G(s)$.

La tabla da el equivalente con mantenedor de orden cero del sistema continuo, $G(s)$, precedido por el mantenedor de orden cero. El sistema muestreado se describe por el operador de transferencia discreta. Para sistemas de segundo orden el operador de transferencia discreta viene dado en términos de los coeficientes de

$$H(q) = \frac{b_1 q + b_2}{q^2 + a_1 q + a_2}$$

$G(s)$	$H(q)$ o los coeficientes de $H(q)$
$\frac{1}{s}$	$\frac{h}{q - 1}$
$\frac{1}{s^2}$	$\frac{h^2(q + 1)}{2(q - 1)^2}$
e^{-sh}	q^{-1}
$\frac{a}{s + a}$	$\frac{1 - \exp(-ah)}{q - \exp(-ah)}$
$\frac{a}{s(s + a)}$	$b_1 = \frac{1}{a}(ah - 1 + e^{-ah}) \quad b_2 = \frac{1}{a}(1 - e^{-ah} - ah e^{-ah})$ $a_1 = -(1 + e^{-ah}) \quad a_2 = e^{-ah}$
$\frac{a^2}{(s + a)^2}$	$b_1 = 1 - e^{-ah}(1 + ah) \quad b_2 = e^{-ah}(e^{-ah} + ah - 1)$ $a_1 = -2e^{-ah} \quad a_2 = e^{-2ah}$
$\frac{ab}{(s + a)(s + b)}$	$b_1 = \frac{b(1 - e^{-ah}) - a(1 - e^{-bh})}{b - a}$ $b_2 = \frac{a(1 - e^{-bh})e^{-ah} - b(1 - e^{-ah})e^{-bh}}{b - a}$ $a_1 = -(e^{-ah} + e^{-bh})$ $a_2 = e^{-(a+b)h}$
$\frac{(s + c)}{(s + a)(s + b)}$	$b_1 = \frac{e^{-bh} - e^{-ah} + (1 - e^{-bh})c/b - (1 - e^{-ah})c/a}{b - a}$ $b_2 = \frac{c}{ab}e^{-(a+b)h} + \frac{b - c}{b(a - b)}e^{-ah} + \frac{c - a}{a(a - b)}e^{-bh}$ $a_1 = -e^{-ah} - e^{-bh} \quad a_2 = e^{-(a+b)h}$
$\frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2}$	$b_1 = 1 - \alpha(\beta + \frac{\zeta\omega_0}{\omega}\gamma) \quad \omega = \omega_0\sqrt{1 - \zeta^2} \quad \zeta < 1$ $b_2 = \alpha^2 + \alpha(\frac{\zeta\omega_0}{\omega}\gamma - \beta) \quad \alpha = e^{-\zeta\omega_0 h}$ $a_1 = -2\alpha\beta \quad \beta = \cos(\omega h)$ $a_2 = \alpha^2 \quad \gamma = \sin(\omega h)$
$\frac{s}{s^2 + 2\zeta\omega_0 s + \omega_0^2}$	$b_1 = \frac{1}{\omega}e^{-\zeta\omega_0 h} \operatorname{sen}(\omega h) \quad b_2 = -b_1 \quad \omega = \omega_0\sqrt{1 - \zeta^2}$ $a_1 = -2e^{-\zeta\omega_0 h} \cos(\omega h) \quad a_2 = e^{-2\zeta\omega_0 h}$

sin condiciones iniciales. La transformada z aplica una *serie de secuencias temporales semiinfinitas* sobre el plano complejo. Nótese la diferencia en rango para la transformada z y para el cálculo con operador. La variable z es una variable compleja y debe distinguirse del operador q .

Definición 3.1 – Transformada z . Considérese la señal discreta $\{f(kh): k=0, 1, \dots\}$. La transformada z de $f(kh)$ se define como

$$F(z) = \sum_{k=0}^{\infty} f(kh)z^{-k} \quad (3.39)$$

donde z es una variable compleja. La transformada z de f se representa por $\mathcal{Z}f$ o F . La transformada inversa viene dada por

$$f(kh) = \frac{1}{2\pi i} \oint F(z)z^{k-1} dz \quad (3.40)$$

donde el contorno de integración encierra todas las singularidades de $F(z)$.

Ejemplo 3.12

Sea $y(kh) = kh$ para $k \geq 0$. Entonces

$$Y(z) = 0 + hz^{-1} + 2hz^{-2} + \dots = h(z^{-1} + 2z^{-2} + \dots) = \frac{hz}{(z - 1)^2}$$

En la Tabla 3.2 se dan algunas propiedades de la transformada z . Nótese que las fórmulas para el desplazamiento hacia delante y hacia atrás no son las mismas. Esto es consecuencia de la hipótesis de que las secuencias temporales son semiinfinitas.

TABLA 3.2 Algunas propiedades de la transformada z .

1. Definición.

$$F(z) = \sum_{k=0}^{\infty} f(kh)z^{-k}$$

2. Inversión

$$f(kh) = \frac{1}{2\pi i} \oint F(z)z^{k-1} dz$$

3. Linealidad.

$$\mathcal{Z}(\alpha f + \beta g) = \alpha \mathcal{Z}f + \beta \mathcal{Z}g$$

4. Desplazamiento temporal.

$$\mathcal{Z}q^{-n}f = z^{-n}F$$

$$\mathcal{Z}\{q^n f\} = z^n(F - F_1)$$

$$\text{donde } F_1(z) = \sum_{j=0}^{n-1} f(jh)z^{-j}$$

5. Teorema del valor inicial.

$$\lim_{k \rightarrow 0} f(kh) = \lim_{z \rightarrow \infty} F(z)$$

6. Teorema del valor final.

Si $(1 - z^{-1})F(z)$ no tiene polos sobre la frontera o fuera del círculo unidad, entonces

$$\lim_{k \rightarrow \infty} f(kh) = \lim_{z \rightarrow 1} (1 - z^{-1})F(z)$$

7. Convolución

$$\mathcal{Z}f * g = \mathcal{Z} \sum_{n=0}^{\infty} f(n)g(k - n) = \mathcal{Z}f \mathcal{Z}g$$



La transformada z puede emplearse para resolver ecuaciones en diferencias; por ejemplo,

$$\begin{aligned}x(k+1) &= \Phi x(k) + \Gamma u(k) \\y(k) &= Cx(k)\end{aligned}$$

Si tomamos la transformada z de ambos lados,

$$\begin{aligned}\sum_{k=0}^{\infty} z^{-k} x(k+1) &= z \left[\sum_{k=0}^{\infty} z^{-k} x(k) - x(0) \right] \\&= \sum_0^{\infty} \Phi z^{-k} x(k) + \sum_0^{\infty} \Gamma z^{-k} u(k)\end{aligned}$$

Por tanto,

$$\begin{aligned}z[X(z) - x(0)] &= \Phi X(z) + \Gamma U(z) \\X(z) &= (zI - \Phi)^{-1}[zx(0) + \Gamma U(z)]\end{aligned}$$

y

$$Y(z) = C(zI - \Phi)^{-1}zx(0) + C(zI - \Phi)^{-1}\Gamma U(z)$$

Podemos ya introducir la *función de transferencia discreta*.

$$H(z) = C(zI - \Phi)^{-1}\Gamma \quad (3.41)$$

que es la misma que (3.38) con q sustituido por z . La secuencia temporal $y(k)$ puede ya obtenerse mediante la transformación inversa. El siguiente teorema es análogo al de los sistemas continuos.

Teorema 3.4. La respuesta discreta de (3.32) y la función de transferencia discreta son un par de transformadas z .

Cálculo de la función de transferencia discreta

Es posible determinar la función de transferencia discreta directamente a partir de la función de transferencia continua. Sea el sistema descrito por la función de transferencia $G(s)$ precedido por un mantenedor de orden cero (véase la Fig. 3.3). La función de transferencia discreta queda determinada de forma única por la respuesta a una señal dada. Consideremos, por ejemplo, una entrada salto unidad. La secuencia $\{u(kh)\}$ es entonces una secuencia de unos, y la señal $u(t)$ es también un salto unidad. Representemos por $Y(s)$ la transformada de Laplace de $y(t)$; es decir,

$$Y(s) = \frac{G(s)}{s}$$

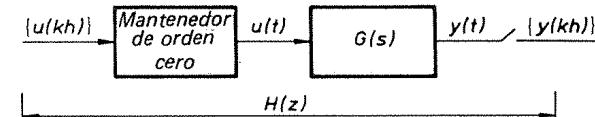


Figura 3.3 Muestreo de un sistema continuo.

La salida muestreada $\{y(kh)\}$ tiene como transformada z $\tilde{Y} = \mathcal{Z}y = \mathcal{Z}\mathcal{L}^{-1}Y$. La división de \tilde{Y} por la función de transferencia discreta de la entrada, que es $z/(z-1)$, nos da

$$H(z) = (1 - z^{-1})\tilde{Y}(z) \quad (3.42)$$

La función de transferencia discreta se obtiene entonces de la forma siguiente:

1. Determinar la función temporal que corresponde a $G(s)/s$.
2. Determinar la transformada z correspondiente (normalmente a partir de una tabla).
3. Multiplicar por $1 - z^{-1}$ para obtener la función de transferencia discreta con un mantenedor de orden cero.

La Tabla 3.3 muestra algunas funciones temporales y sus correspondientes transformadas de Laplace y z . La tabla puede entonces usarse para combinar los pasos 1 y 2. Las tablas de los libros de texto tienen normalmente esta misma forma.

TABLA 3.3 Algunas funciones del tiempo y sus correspondientes transformadas de Laplace y z . Aviso: ¡Usese la tabla sólo como se indica!

f	$\mathcal{L}f$	zf
$1 \quad k \geq 0$ (salto)	$\frac{1}{s}$	$\frac{z}{z-1}$
kh	$\frac{1}{s^2}$	$\frac{hz}{(z-1)^2}$
$\frac{1}{2}(kh)^2$	$\frac{1}{s^3}$	$\frac{h^2z(z+1)}{2(z-1)^3}$
$e^{-kh/T}$	$\frac{T}{1+sT}$	$\frac{z}{z-e^{-h/T}}$
$1 - e^{-kh/T}$	$\frac{1}{s(1+sT)}$	$\frac{z(1-e^{-h/T})}{(z-1)(z-e^{-h/T})}$
$\operatorname{sen} \omega kh$	$\frac{\omega}{s^2 + \omega^2}$	$\frac{z \operatorname{sen} \omega h}{z^2 - 2z \cos \omega h + 1}$

Aviso. Téngase en cuenta que $\mathcal{L}f$ en la Tabla 3.3 no da el muestreo con mantenedor de orden cero de un sistema que tiene una función de transferencia $\mathcal{L}f$. Es

una equivocación muy corriente creer que lo incluye. La función de transferencia discreta se obtiene mediante el procedimiento señalado.

Si no hay término independiente ni polos múltiples en el sistema continuo, entonces se puede reescribir (3.42) y $H(z)$ puede obtenerse empleando

$$H(z) = \sum_{s=s_i} \frac{1}{z - e^{sh}} \operatorname{Res} \left\{ \frac{e^{sh} - 1}{s} G(s) \right\} \quad (3.43)$$

donde s_i son los polos de $G(s)$. Una demostración de esta fórmula puede encontrarse en la Sección 4.5.

3.6. Polos y ceros

Los sistemas con función de transferencia racional quedan caracterizados de forma adecuada mediante los polos y ceros de la función de transferencia. Por tanto es muy útil comprender correctamente la relación entre los polos y los ceros de un sistema continuo y los del sistema discreto correspondiente.

Polos

Consideremos un sistema continuo descrito por el modelo en el espacio de estados de orden n

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \quad (3.44)$$

Los polos del sistema son los valores propios de A , que designamos por $\lambda_i(A)$, $i = 1, \dots, n$. El muestreo con mantenedor de orden cero de (3.44) da el sistema discreto

$$\begin{aligned} x(kh + h) &= \Phi x(kh) + \Gamma u(kh) \\ y(kh) &= Cx(kh) \end{aligned} \quad (3.45)$$

Sus polos son los valores propios de Φ , $\lambda_i(\Phi)$, $i = 1, \dots, n$. Ya que $\Phi = \exp(Ah)$, de las propiedades de las funciones matriciales (véase el Apéndice B) se deduce que

$$\lambda_i(\Phi) = e^{\lambda_i(A)h} \quad (3.46)$$

La ecuación (3.46) da la relación entre los polos continuos y los discretos. La Figura 3.4 ilustra la conversión del plano complejo s en el plano z , cuando $z = \exp(sh)$. Por ejemplo, la mitad izquierda del plano s se corresponde con el círculo unidad del plano z . La aplicación no es biyectiva; varios puntos del plano s se aplican sobre el mismo

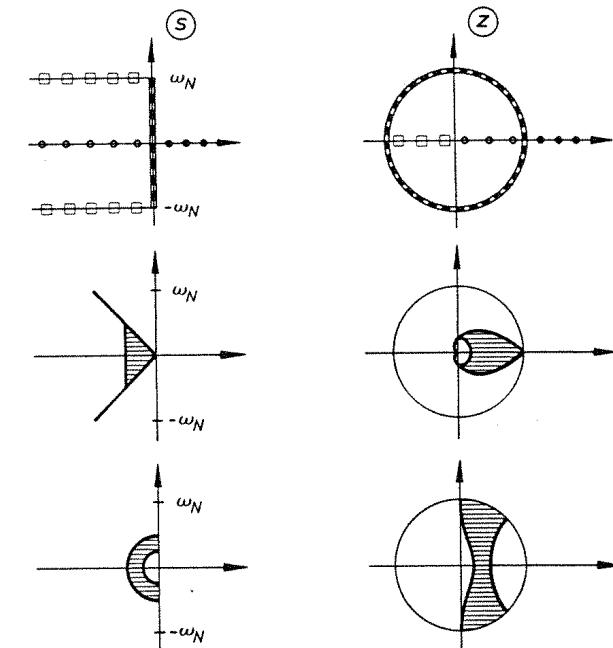


Figura 3.4 La transformación conforme $z = \exp(sh)$.

punto del plano z (véase la Fig. 3.5). Esta es otra ilustración del efecto de enmascaramiento presentado en la Sección 2.5. Para los polos que están dentro de la banda fundamental S_0 en la Figura 3.5 hay una relación sencilla entre los polos continuos y discretos. (Compárese también con el Ejemplo 3.4).

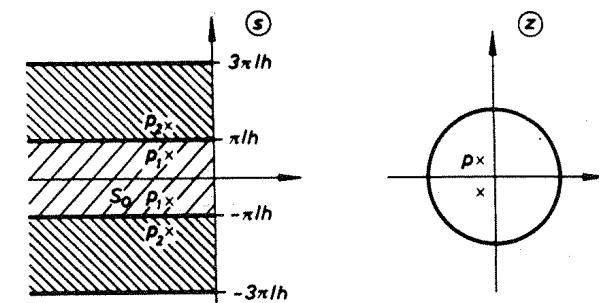


Figura 3.5 Cada banda de la mitad izquierda del plano s se aplica sobre el disco unidad. Esto significa que los pares de polos p_1 y p_2 se aplican ambos sobre el par p .

Ejemplo 3.13

Consideremos el sistema continuo

$$\frac{\omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \quad (3.47)$$

Los polos del sistema discreto correspondiente vienen dados por la ecuación característica

$$z^2 + a_1 z + a_2 = 0$$

donde

$$a_1 = -2e^{-\zeta\omega_0 h} \cos(\sqrt{1-\zeta^2}\omega_0 h)$$

$$a_2 = e^{-2\zeta\omega_0 h}$$

(Compárese con la Tabla 3.1.) La Figura 3.6 muestra las respuestas a un salto escalón de un sistema discreto para diferentes intervalos de muestreo cuando $\omega_0 = 1,83$ y $\zeta = 0,5$. La Figura 3.7 da una visión más detallada de cómo se aplican los polos continuos de (3.47) sobre el círculo unidad para diferentes valores de ζ y $\omega_0 h$ cuando se muestrea el sistema.

Ceros

No es posible dar una fórmula sencilla sobre la correspondencia de los ceros. Si una función de transferencia continua se ve como una función racional, tiene sus ceros en los ceros del polinomio numerador y además $d=r-1$ ceros en el infinito, siendo d el exceso de polos de la función de transferencia continua; es decir, la diferencia entre el número de polos y el número de ceros.

Para períodos de muestreo cortos, un sistema discreto tendrá ceros en

$$z_i \approx e^{s_i h}$$

donde los s_i son los ceros del sistema continuo. Los $r-1$ ceros que se corresponden con los ceros en el infinito del sistema continuo cuando el período de muestreo tiende a

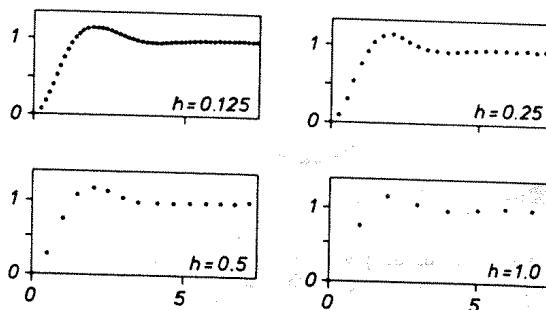


Figura 3.6 Respuestas salto del sistema discreto del Ejemplo 3.13 para diferentes valores de h cuando $\zeta = 0,5$ y $\omega_0 = 1,83$, lo que da un tiempo de subida $T_r = 1$.

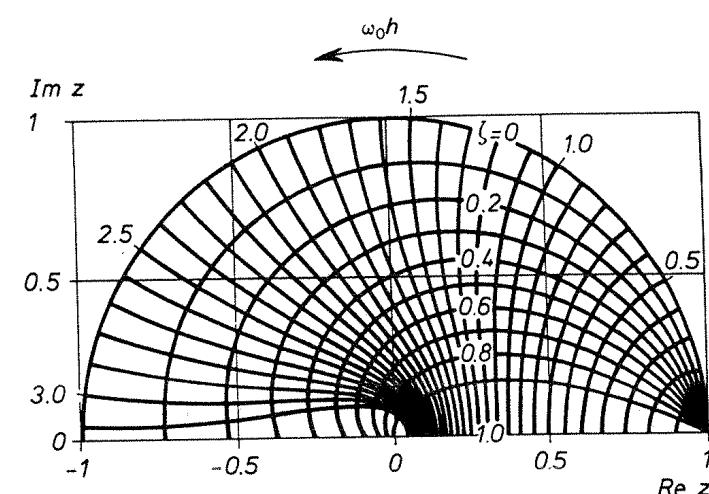


Figura 3.7 Lugares de ζ y $\omega_0 h$ constantes cuando se muestrea (3.47).

cero irán a los ceros de los polinomios Z_r de la Tabla 3.4 debido a que para s grande, la función de transferencia del sistema continuo viene dada aproximadamente por $G(s) \approx s^{-r}$.

TABLA 3.4 Polinomios numerador, Z_r , cuando se muestrea s^{-r} .

r	Z_r
1	1
2	$z + 1$
3	$z^2 + 4z + 1$
4	$z^3 + 11z^2 + 11z + 1$
5	$z^4 + 26z^3 + 66z^2 + 26z + 1$

Ejemplo 3.14

Consideremos la función de transferencia continua

$$\frac{2}{(s+1)(s+2)}$$

Empleando la Tabla 3.1 se obtiene el cero de la función de transferencia discreta

$$z = -\frac{(1-e^{-2h})e^{-h} - 2(1-e^{-h})e^{-2h}}{2(1-e^{-h}) - (1-e^{-2h})}$$

Cuando h es pequeña

$$z \approx -1 + 3h$$

y cuando h se aproxima a cero el cero se desplaza hacia -1 . El cero se mueve hacia el origen cuando h aumenta.

Para pequeños valores de h el cero puede también obtenerse en la Tabla 3.4. El polo excedente del sistema continuo es $r=2$. El sistema discreto tendrá un cero en $z_r(z)=z+1=0$ cuando h tiende a cero.

Los ceros de un sistema discreto obtenido muestreando un sistema continuo dependen del circuito usado como mantenedor. Esto se ve en el Ejemplo 3.15.

Ejemplo 3.15

Consideremos el integrador doble del Ejemplo 3.7, en el que el mantenedor es tal que la señal de control vale $\alpha u(k)$ durante la primera mitad del intervalo de muestreo y $\beta u(k)$ en la segunda mitad. La función de transferencia discreta del sistema es

$$\begin{aligned} H(z) &= [1 \ 0](zI - \Phi)^{-1}\Gamma \\ &= \frac{h^2[z(3\alpha + \beta) + (\alpha + 3\beta)]}{8(z - 1)^2} \end{aligned}$$

La ganancia a baja frecuencia viene dada por $\alpha + \beta$. Tomemos

$$\alpha + \beta = 2$$

La función de transferencia discreta tiene un cero

$$z = -\frac{\alpha + 3\beta}{3\alpha + \beta} = -\frac{3 - \alpha}{1 + \alpha}$$

El cero de la función de transferencia discreta puede entonces posicionarse en un valor arbitrario sin más que escoger α de forma apropiada. El valor $\alpha=1$, que corresponde a un mantenedor de orden cero, da $z=-1$ y $\alpha=3$ da $z=0$.

Los resultados del ejemplo pueden extenderse a sistemas de orden n . Se observa que los ceros pueden colocarse arbitrariamente si la señal de control es constante sobre cada n -ésima parte del intervalo de muestreo. Un circuito de tales características puede realizarse muestreando la señal de control a mayor frecuencia; sin embargo, esta idea no ha sido muy utilizada debido a que la señal de control puede hacerse muy irregular.

Sistemas con inversos inestables

Un sistema continuo con una función de transferencia racional se dice que es de fase no mínima si tiene ceros en la mitad derecha del plano o retardos temporales. De forma análoga, se define que un sistema discreto es de fase no mínima si tiene ceros fuera del círculo unidad. Esta definición implica que un retardo temporal no hace que

el sistema sea de fase no mínima. Por otra parte, los retardos no presentan los mismos problemas que producen en los sistemas continuos. En los sistemas discretos es más importante hablar acerca de sistemas con o sin inversos estables, que se definen de la siguiente forma.

Definición 3.2 – Inverso inestable. Un sistema discreto tiene un inverso inestable si tiene ceros fuera del círculo unidad.

Un sistema continuo con un inverso estable puede convertirse en un sistema discreto con inverso inestable cuando se le muestrea. De la Tabla 3.4 se deduce que el sistema inverso será siempre inestable si el exceso de polos del sistema continuo es mayor que dos, y si el período de muestreo es suficientemente corto. Además un sistema continuo de fase no mínima no siempre se convertirá en un sistema discreto con inverso inestable, como puede verse en el siguiente ejemplo.

Ejemplo 3.16

La función de transferencia

$$G(s) = \frac{6(1-s)}{(s+2)(s+3)}$$

tiene un cero inestable en $s=1$. Al muestrear el sistema se obtiene una función de transferencia discreta con un cero en

$$z_1 = -\frac{8e^{-2h} - 9e^{-3h} + e^{-5h}}{1 - 9e^{-2h} + 8e^{-3h}}$$

Para $h \approx 1,25$, $z_1 = -1$; para valores mayores de h el cero está siempre dentro del círculo unidad y el sistema muestreado tiene un inverso estable.

3.7. Selección de la frecuencia de muestreo

En la Sección 2.6 se ha expuesto la importancia de una selección adecuada de la frecuencia de muestreo. Un período de muestreo demasiado grande hace imposible la reconstrucción de la señal continua. Un período de muestreo demasiado corto aumentará la carga del computador. El problema de la selección de la frecuencia de muestreo se volverá a tratar en el libro más adelante. Esta sección analiza la relación entre la frecuencia de muestreo y los polos del sistema continuo.

Es útil caracterizar el período de muestreo por una variable sin dimensiones y esto tiene una fácil interpretación física. En los sistemas oscilatorios es natural normalizar respecto al período de oscilación; en los sistemas no oscilatorios, el tiempo de subida es un factor de normalización natural.

Introduzcamos N_r como el número de períodos de muestreo por tiempo de subida,

$$N_r = \frac{T_r}{h}$$



donde T_r es el tiempo de subida. El teorema de muestreo de Shannon nos da el límite inferior. Para señales sinusoidales puras, esto da $N_r \approx 0,32$. La reconstrucción de Shannon es, sin embargo, un tanto complicada.

Para sistemas de primer orden, el tiempo de subida es igual a la constante de tiempo. Por tanto, es razonable escoger N_r entre 2 y 4. Para un sistema de segundo orden con un amortiguamiento ζ y una frecuencia natural ω_0 , el tiempo de subida viene dado por

$$T_r = \omega_0^{-1} e^{\varphi/\lg \varphi}$$

donde $\zeta = \cos \varphi$. Para un amortiguamiento alrededor de $\zeta = 0,7$ esto da

$$\omega_0 h \approx 0.5 - 1$$

donde ω_0 está medido en radianes por segundo.

Las Figuras 3.6 y 3.8 ilustran la elección del período de muestreo para diferentes señales. Es razonable escoger el período de muestreo de forma que

$$N_r = \frac{T_r}{h} \approx 2-4$$

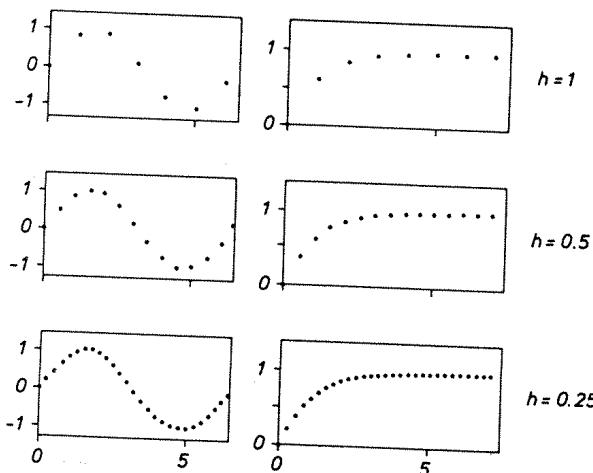


Figura 3.8 Ilustración del muestreo de una señal sinusoidal y de otra exponencial. Los tiempos de subida de las señales son $T_r = 1$.

3.8. Problemas

3.1. Considérese el sistema

$$\begin{aligned}\dot{x} &= -ax + bu \\ y &= cx\end{aligned}$$

Sea la entrada constante durante períodos de longitud h . Muestrear el sistema y explicar cómo varían los polos del sistema discreto con el intervalo de muestreo h .

3.2. Deducir el sistema discreto correspondiente a los siguientes sistemas continuos cuando se emplea un circuito mantenedor de orden cero:

$$(a) \frac{dx}{dt} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}x + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u$$

$$(b) \frac{d^2y}{dt^2} + 3 \frac{dy}{dt} + 2y = \frac{du}{dt} + 3u$$

$$(c) \frac{d^3y}{dt^3} = u$$

3.3. Se supone que las siguientes ecuaciones en diferencias describen sistemas continuos muestreados empleando un mantenedor de orden cero y con un período de muestreo h . Determinar, si es posible, los sistemas continuos correspondientes.

$$(a) y(kh) - 0.5y(kh - h) = 6u(kh - h)$$

$$(b) x(kh + h) = \begin{bmatrix} -0.5 & 1 \\ 0 & -0.3 \end{bmatrix}x(kh) + \begin{bmatrix} 0.5 \\ 0.7 \end{bmatrix}u(kh)$$

$$y(kh) = [1 \ 1]x(kh)$$

$$(c) y(kh) + 0.5y(kh - h) = 6u(kh - h).$$

3.4. Considérese el oscilador armónico [véanse el Ejemplo A.3 o el Problema 3.2 a)]. Calcular la respuesta a un salto escalón en 0, h , $2h$, ..., cuando el período de muestreo es a) $h = \pi/2$, b) $h = \pi/4$. Comparar con la respuesta escalón continua.

3.5. Muestrear el sistema $G(s) = 1/s$ empleando un mantenedor de orden uno.

3.6. Considérese el sistema dado en (3.1). Supóngase que la entrada es una suma de impulsos en los instantes de muestreo; es decir,

$$u(t) = \sum \delta(t - kh)u(kh)$$

Determinese la representación discreta.

3.7. Encontrar la matriz de transformación, T , que transforma la representación en el espacio de estados del integrador doble (3.11) en la forma canónica controlable.

3.8. Determinar la función de transferencia discreta del sistema

$$\begin{aligned}x(kh + h) &= \begin{bmatrix} 0.5 & -0.2 \\ 0 & 0 \end{bmatrix}x(kh) + \begin{bmatrix} 2 \\ 1 \end{bmatrix}u(kh) \\ y(kh) &= [1 \ 0]x(kh)\end{aligned}$$

3.9. Muchos sistemas físicos pueden describirse por la expresión

$$\frac{dx}{dt} = \begin{bmatrix} -a & b \\ c & -d \end{bmatrix}x + \begin{bmatrix} f \\ g \end{bmatrix}u$$

donde a, b, c y d son no negativos. Deducir una fórmula para los sistemas muestreados cuando emplean un mantenedor de orden cero.
(Sugerencia: Demostrar primero que los polos del sistema son reales.)

- 3.10. La Figura 3.9 presenta un sistema con dos tanques, donde la señal de entrada es el flujo al primer tanque y la salida es el nivel del segundo tanque. Empleando los niveles como variables de estado se obtiene el sistema

$$\frac{dx}{dt} = \begin{bmatrix} -0.0197 & 0 \\ 0.0178 & -0.0129 \end{bmatrix}x + \begin{bmatrix} 0.0263 \\ 0 \end{bmatrix}u$$

$$y = [0 \ 1]x$$

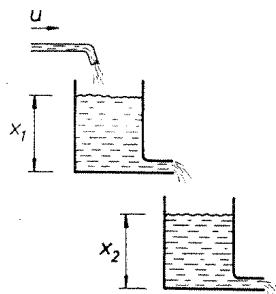


Figura 3.9 Proceso de dos tanques.

- a) Muestrear el sistema con un período de muestreo $h=12$.
b) Comprobar que el operador discreto para el sistema es

$$H_0(q) = \frac{0.030q + 0.026}{q^2 - 1.65q + 0.68}$$

- 3.11. En el Ejemplo A.2 se describe el motor normalizado. Demostrar que el sistema muestreado está descrito por (A.6). Determinar lo siguiente:

- a) La función de transferencia discreta.
b) La respuesta discreta.
c) Una ecuación en diferencias que relacione la entrada y la salida.
d) La variación de los polos y ceros de la función de transferencia discreta con el período de muestreo.

- 3.12. Se muestrea con $h=1$ un sistema continuo con función de transferencia

$$G(s) = \frac{1}{s} e^{-\tau s}$$

cuando $\tau=0.5$.

- a) Determinar una representación en el espacio de estados del sistema muestreado. ¿Cuál es el orden del sistema muestreado?
b) Determinar la función de transferencia discreta y la respuesta discreta del sistema muestreado.
c) Determinar los polos y ceros del sistema muestreado.

- 3.13. Resolver el Problema 3.12 con

$$G(s) = \frac{1}{s+1} e^{-\tau s}$$

y $\tau=1.5$ y $h=1$.

- 3.14. Considérese el sistema muestreado

$$y(k+1) = ay(k) + b_3u(k-3) + b_4u(k-4)$$

con un período de muestreo de 1 s. Demostrar que este sistema se puede obtener muestreando el sistema

$$\frac{dy(t)}{dt} = -\alpha y(t) + bu(t-\tau)$$

donde

$$\tau = 4 - \frac{\ln(ab_3 + b_4)/(b_3 + b_4)}{\ln a}$$

- 3.15. Considérese el sistema

$$y(k) - 0.5y(k-1) = u(k-9) + 0.2u(k-10)$$

Determinar los polinomios $A(q)$, $B(q)$, $A^*(q^{-1})$ y $B^*(q^{-1})$ en las representaciones

$$A(q)y(k) = B(q)u(k)$$

y

$$A^*(q^{-1})y(k) = B^*(q^{-1})u(k-d)$$

¿Cuánto valen d y el orden del sistema?

- 3.16. Un filtro con un operador de transferencia discreta

$$H^*(q^{-1}) = b_0 + b_1q^{-1} + \cdots + b_nq^{-n}$$

recibe el nombre de un filtro de respuesta impulsional finita (FIR).

- a) Determinar el orden del sistema.
b) Dar una representación en el espacio de estados en la forma canónica observable.

- 3.17. Utilizar la transformada z para determinar la secuencia de salida de la ecuación en diferencias

$$y(k+2) - 1.5y(k+1) + 0.5y(k) = u(k+1)$$

cuando $u(k)$ es un salto escalón en $k=0$ y cuando $y(0)=0.5$ e $y(-1)=1$.

- 3.18. Verificar que

$$z \left\{ \frac{1}{2}(kh)^2 \right\} = \frac{h^2 z(z+1)}{2(z-1)^3}$$

Comparar con la Tabla 3.3 y emplearla para determinar la función de transferencia discreta del integrador doble (véase el Ejemplo A.2).

- 3.19. Emplear (3.43) para determinar la función de transferencia discreta de: a) el sistema en el Problema 3.1 y b) el motor normalizado (véase el Ejemplo A.2).

- 3.20. Demostrar que una curva de amortiguamiento constante ζ en el plano s es una espiral logarítmica en el plano z cuando se usa la aplicación $z = \exp(sh)$.
- 3.21. Si $\beta < \alpha$ a

$$\frac{s + \beta}{s + \alpha}$$

se le llama *red de adelanto* (es decir, produce un adelanto de fase). Considérese el sistema discreto

$$\frac{z + b}{z + a}$$

- a) Determinar cuándo se trata de una red de adelanto.
 b) Simular la respuesta a un salto escalón para diferentes posiciones de los polos y ceros.

- 3.22. Considérese el sistema

$$\frac{z + b}{(1 + b)(z^2 - 1.1z + 0.4)}$$

La situación de los polos corresponde a un sistema continuo con amortiguamiento $\zeta = 0.7$. Simular el sistema y determinar el rebalse para diferentes valores de b en el intervalo $(-1, 1)$.

- 3.23. Considérese el sistema continuo estable

$$G(s) = \frac{s + b}{s + a}$$

donde $a \neq b$. Muestrear el sistema con un período de muestreo h . Determinar las condiciones en las que el sistema muestreado tendrá un inverso estable.

3.9. Referencias

Los textos primitivos sobre sistemas muestreados

- JURY, E. I. (1958): *Sampled-Data Control Systems*. New York: John Wiley.
 RAGAZZINI, J. R. y G. F. FRANKLIN (1958): *Sampled-Data Control Systems*. New York: McGraw-Hill.
 TSYPKIN, Y. Z. (1958): *Theory of Impulse Systems*. Moscow: Editora Estatal de Literatura Física y Matemática.

tratan exclusivamente los modelos entrada-salida y la teoría de la transformada. La aproximación del espacio de estados empleada en este capítulo ofrece simplificaciones significativas. Con un mantenedor de orden cero, la señal de control es constante durante el período de muestreo y el modelo discreto se obtiene fácilmente integrando la ecuación de estado durante un período de muestreo. Esta formulación del problema se introdujo en

- KALMAN, R. E. y J. E. BERTRAM (1958): «General Synthesis Procedure for Computer Control of Single and Multiloop Linear Systems», *AIEE Trans.* 77, 602-09.

Fue necesario que pasara un cierto tiempo para que este planteamiento encontrara sitio en los libros de texto. Debido a su sencillez es hoy día el planteamiento predominante.

La transformación de las variables de estado y las formas canónicas es un material normal en la teoría del espacio de estados. Estos resultados son muy similares a los resultados correspondientes para los sistemas continuos. Un tratamiento más detallado puede encontrarse en

KAILATH, T. (1980): *Linear Systems*. Englewood Cliffs, N. J.: Prentice Hall, Inc.

Históricamente, la formulación entrada-salida es anterior a la del espacio de estados. En los textos clásicos mencionados anteriormente se hace un tratamiento directo desde este punto de vista. El caso multivariable se trata en

ROSEN BROCK, H. H. (1970): *State-Space and Multivariable Theory*. London: Nelson.

KUČERA, V. (1979): *Discrete Linear Control*. Prague: Academia.

La transformada z se trata ampliamente en

JURY, E. I. (1958): *Sampled-Data Control Systems*. New York: John Wiley. Second printing: Krieger, 1977.

JURY, E. I. (1964): *Theory and Application of the Z-Transform Method*. New York: John Wiley. Second printing: Krieger, 1973.

DOETSCH, G. (1971): *Guide to the Applications of the Laplace and Z-Transforms*. New York: Van Nostrand Reinhold.

En estas referencias pueden encontrarse tablas más amplias de pares de transformadas z . Una tabla de las funciones de transferencia equivalentes con mantenedor de orden cero (comárese con la Tabla 3.1) puede encontrarse en

NEUMAN, C. P. y C. S. BARADELLO (1979): «Digital Transfer Functions for Microcomputer Control», *IEEE Trans. Sys., Man. and Cybernetics*, SMC-9, 856-60.

La relación entre los ceros de los sistemas continuos y muestreados se expone en

ÅSTRÖM, K. J., P. HAGANDER, y J. STERNBY (1980): «Zeros of Sampled Systems», *Proc. 19th IEEE Conf. on Decision and Control*, Alburquerque, 1077-81.

En este trabajo se dan los teoremas para los ceros en los límites para períodos de muestreo grandes y pequeños.

La observación de que los ceros de un sistema muestreado pueden posicionarse arbitrariamente eligiendo de forma adecuada el mantenedor creemos que es nueva.

Modelos orientados al proceso

Objetivo:

Desarrollar los modelos matemáticos que proporcionan las relaciones entre las señales continuas en un sistema muestreado.

4.1. Introducción

En el Capítulo 3 se desarrollaron los modelos matemáticos para un sistema muestreado *desde el punto de vista del computador*. Estos modelos son bastante simples. Las variables que representan la señal medida y la señal de control se consideran únicamente en los instantes de muestreo. Estas variables cambian en unos instantes de tiempo indicados por el reloj del sistema. Las señales se presentan normalmente al computador como una secuencia de números. Por tanto, el modo en que los sistemas muestreados varían con el tiempo puede ignorarse, ya que las señales se consideran únicamente en los instantes de tiempo indicados por el reloj de sistema. Los sistemas muestreados pueden describirse, por tanto, como *sistemas invariantes y discretos en el tiempo*. El modelo obtenido se denomina modelo estroboscópico.

El modelo estroboscópico tiene la ventaja de ser muy simple. La mayoría de los problemas de análisis y diseño de sistemas muestreados puede, afortunadamente, tratarse mediante este modelo. El modelo también da una completa descripción del sistema mientras que sea observado desde el computador, pero en algunas ocasiones esto no es suficiente. Por tanto, es útil disponer de otros modelos que den una descripción detallada. Estos modelos se necesitan cuando el sistema controlado por computador se observa desde el proceso, por ejemplo, si se analiza la respuesta en frecuencia cortando el lazo de la parte analógica. Los modelos requeridos son necesariamente más complicados que los desarrollados en el Capítulo 3, ya que debe tratarse explícitamente la naturaleza periódica del sistema.

En la Sección 4.2 se da una descripción detallada de los acontecimientos que

ocurren en un sistema controlado por computador. La clave del problema del modelado está en la descripción del proceso de muestreo. Esto se describe en la Sección 4.3 utilizando el *modelo de modulación*. La Sección 4.4 trata de la respuesta en frecuencia de los sistemas muestreados —pueden suceder varias situaciones inesperadas—. Los resultados dan mayor claridad al problema del enmascaramiento. En la Sección 4.5 se perfila una teoría algebraica para sistemas muestreados. Los sistemas multiperiódicos se discuten en la Sección 4.6.

4.2. Sistema controlado con computador

El diagrama de un sistema controlado con computador aparece en la Figura 4.1. En el Capítulo 3, el lazo de realimentación se cortaba dentro del computador entre los convertidores A-D y D-A; por ejemplo, en el punto C de la figura. En este capítulo, el lazo se corta en la parte analógica; por ejemplo, en el punto A de la figura.

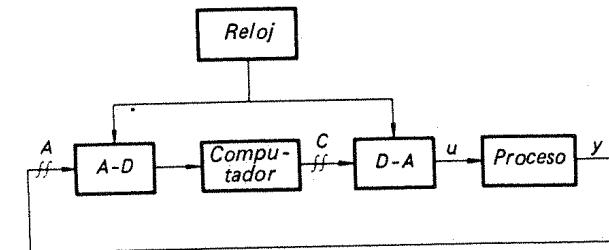


Figura 4.1 Diagrama esquemático de un sistema controlado con computador.

El estudio de este capítulo requiere una descripción más detallada de la secuencia de operaciones en un sistema controlado por computador. El algoritmo del computador es el siguiente:

1. Esperar hasta que ocurra un pulso de reloj.
2. Realizar una conversión analógica a digital.
3. Calcular las variables de control.
4. Realizar una conversión digital a analógica.
5. Actualizar el estado del regulador.
6. Volver de nuevo al punto 1.

Ya que las operaciones en un computador tardan algún tiempo en realizarse, existe un retardo entre los pasos 2 y 4. La relación entre las diferentes señales en el sistema se muestran en la Figura 4.2. Cuando la ley de control se realiza con un computador es muy importante estructurar el programa para minimizar el tiempo de cálculo que requiere el paso 3 (véase el Capítulo 15).

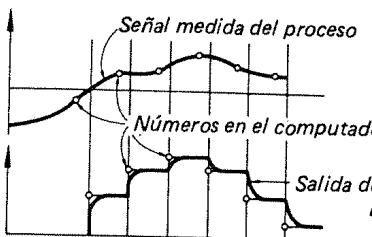


Figura 4.2 Relación entre la señal medida, la señal de control y su representación en el computador.

También es importante expresar perfectamente la sincronización entre señales. Para el análisis se han elegido arbitrariamente los instantes de muestreo en los mismos instantes en que se completa la conversión D-A. Ya que la señal de control es discontinua, es importante precisar los puntos límite. Se ha adoptado el convenio de continuidad por la derecha. Debe notarse que la señal real de entrada al sistema es continua debido al tiempo de asentamiento no nulo del convertidor D-A y de los actuadores.

4.3. El modelo de modulación

Una de las características típicas de los sistemas controlados con computador con mantenedores de orden cero es que la señal de control es constante durante cada período de muestreo. Esta característica se utilizó en el Capítulo 3 para describir la evolución de los sistemas de un instante de muestreo al siguiente, integrando la ecuación del sistema durante un período de muestreo; esta sección intenta describir lo que ocurre entre los instantes de muestreo. Se necesitan otros modelos matemáticos, puesto que no resulta suficiente modelar las señales como secuencias (funciones que aplican Z en R); sino que tienen que ser modeladas como funciones continuas (funciones que aplican R en R).

El tema central consiste en desarrollar el *modelo de modulación*. Este modelo es más complicado que el modelo estroboscópico expuesto en el Capítulo 3. La principal dificultad radica en que debe considerarse la naturaleza periódica de los sistemas muestreados. El sistema puede describirse como un modulador de amplitud seguido de un sistema lineal. La señal de modulación es un tren de pulsos. Una idealización adicional se obtiene approximando los pulsos por impulsos. El modelo tiene su origen en los primeros trabajos sobre sistemas muestreados de MacColl (1945) y Linvil (1951), entre otros.

En el caso especial de un control con computador que emplee un algoritmo con ganancia unidad y retardo de tiempo insignificante, la acción combinada del convertidor A-D, el computador, y el convertidor D-A puede describirse como un sistema que muestra la señal analógica y produce otra señal analógica que es constante durante cada período de muestreo. A tal circuito se le denominó *muestreador y mantenedor*. Un convertidor A-D también puede describirse como un

muestreador y mantenedor. El circuito mantenedor conserva constante el voltaje analógico durante la conversión a una representación digital. En primer lugar, desarrollaremos un modelo detallado para el circuito muestreador y mantenedor.

Un modelo del muestreador y mantenedor

La Figura 4.3 muestra el esquema de un circuito analógico muestreador y mantenedor. Se supone que el circuito continua con un amplificador de muy alta impedancia de entrada. El circuito funciona del siguiente modo: cuando se cierra el interruptor de muestreo, el condensador se carga al voltaje de entrada a través de la resistencia R . Cuando el interruptor de muestreo se abre, el condensador mantiene su voltaje hasta el siguiente período de muestreo.

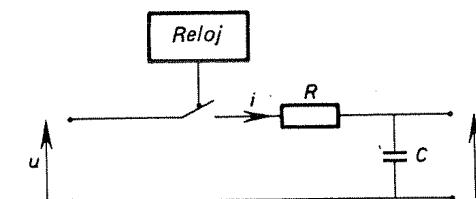


Figura 4.3 Diagrama esquemático de un circuito muestreador y mantenedor.

Para describir el sistema se introduce una función m , la cual describe la apertura y cierre del interruptor de muestreo. Esta función está definida por

$$m(t) = \begin{cases} 1 & \text{si el interruptor está abierto} \\ 0 & \text{si el interruptor está cerrado} \end{cases}$$

Entonces, la corriente viene dada por

$$i = \frac{u - y}{R} m$$

De este modo, la corriente se *modula* por la función m , la cual se denomina *función de modulación*. Si la impedancia de entrada del circuito que sigue al muestreador y mantenedor es grande, el voltaje del condensador está definido por

$$C \frac{dy(t)}{dt} = i(t) = \frac{u(t) - y(t)}{R} m(t) \quad (4.1)$$

La ecuación diferencial (4.1) es un sistema lineal *que varía con el tiempo*. La variación en el tiempo está causada por la modulación. Si el período de muestreo es h y el interruptor permanece cerrado τ segundos en cada muestreo, la función m tiene la forma mostrada en la Figura 4.4. Como m es una función periódica, el sistema resulta ser un *sistema periódico*.

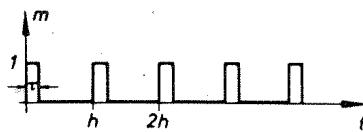


Figura 4.4 Gráfico de la función de modulación m con periodo h y anchura de pulso τ .

Una vez que se ha obtenido el modelo matemático del circuito, puede estudiarse su respuesta a una señal de entrada u . De la ecuación (4.1) se obtiene que el voltaje del condensador es constante cuando el interruptor está abierto; es decir, cuando $m(t)=0$. Cuando el interruptor se cierra, el voltaje y alcanza la señal de entrada u como un sistema dinámico de primer orden con una constante de tiempo RC . La constante de tiempo RC del circuito debe ser considerablemente menor que la anchura del pulso; de otro modo, el condensador no dispone de tiempo para cargarse al voltaje de entrada cuando el circuito se cierra.

La Figura 4.5 muestra los resultados de una simulación del circuito muestreador y mantenedor. Con los parámetros elegidos, la anchura del pulso τ es tal que la señal de entrada cambia significativamente cuando se cierra el interruptor.

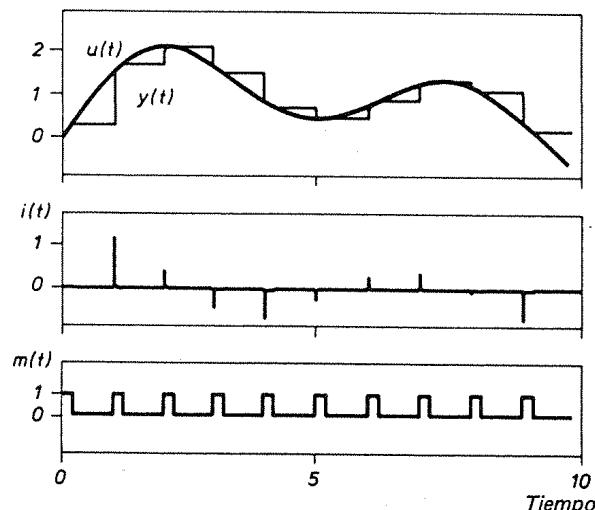


Figura 4.5 Simulación de un circuito muestreador y mantenedor. La anchura τ es 0,2 s y la constante de tiempo es $RC=0,01$ s.

La Figura 4.6 muestra lo que ocurre cuando la anchura del pulso es menor. El resultado de la Figura 4.6 representa una elección razonable del valor de los parámetros. El circuito muestreador y mantenedor alcanza rápidamente el valor de la señal de entrada y la mantiene constante durante el período de muestreo.

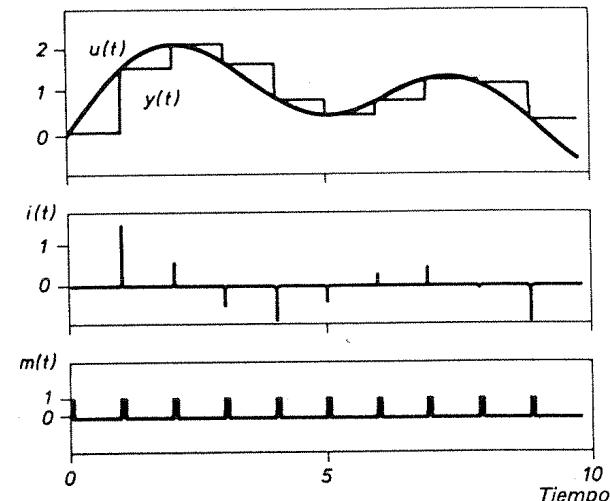


Figura 4.6 Simulación de un circuito muestreador y mantenedor. La anchura τ es 0,05 s y la constante de tiempo es $RC=0,01$ s.

Muestreadores prácticos

En la práctica, un muestreador no se realiza como se representa en la Figura 4.3. Una realización común se muestra en la Figura 4.7. Para mantener la carga en el condensador se utiliza un amplificador operacional. El circuito de la Figura 4.7 también se puede describir por la ecuación (4.1).

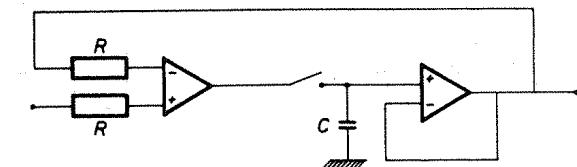


Figura 4.7 Una realización práctica de un circuito muestreador y mantenedor.

Para evitar problemas de ruidos eléctricos y derivaciones a masa es importante aislar galvánicamente el computador de las señales del proceso. Esto puede llevarse a cabo con el circuito de la Figura 4.8, el cual se denomina *circuito de condensador*

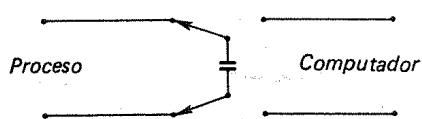


Figura 4.8 Circuito muestreador y mantenedor basado en la realización del condensador flotante.



flotante. El circuito combina el aislamiento eléctrico con la acción de muestrear y mantener de un modo elegante. El condensador se carga al voltaje de entrada cuando éste se conecta a la línea de entrada. Cuando el condensador se conecta al convertidor D-A mantiene su voltaje. El aislamiento eléctrico se obtiene porque el condensador está conectado al proceso o al convertidor D-A del computador de control. En la práctica es muy común cargar el condensador a través de un amplificador operacional. El circuito de condensador flotante también se puede describir con la ecuación (4.1).

Una idealización matemática

El esquema de modulación de pulso es fácil de simular pero difícil de analizar. Por tanto, introduciremos una idealización matemática más fácil de utilizar. Parece razonable diseñar el circuito muestreador y mantenedor para que la anchura del pulso τ sea menor que el período de muestreo. También parece razonable elegir la constante de tiempo RC más corta que la anchura del pulso. Entonces, la corriente a través del condensador estará formada por pulsos cortos. Tanto la altura como la integral en el tiempo de un pulso es proporcional a la diferencia $u - y$ entre el voltaje de entrada u y el voltaje del condensador y en los instantes de muestreo.

En la idealización, los pulsos de corriente se reemplazan por impulsos. Por simplicidad, la integral de los impulsos se elige proporcional a la señal de entrada u en los instantes de muestreo. Entonces, el condensador se reemplaza por un integrador. Ya que los pulsos se eligieron proporcionales a u y no a $u - y$, es necesario anular la integral cada vez que llega un nuevo pulso. Por consiguiente, la corriente se representa por

$$u^* = um \quad (4.2)$$

donde

$$m(t) = \sum_{k=-\infty}^{\infty} \delta(t - kh) \quad (4.3)$$

y δ es la función delta [compárese con (4.1)]. La señal u^* se denomina *representación muestreada* de la señal continua u . Resulta útil recordar que u^* está relacionada con la corriente a través del condensador del circuito muestreador y mantenedor de la Figura 4.3.

La señal u^* puede considerarse como una modulación de u por una señal portadora de la forma de un tren de impulsos. El modelo, por tanto, se denomina modelo de *modulación por tren de impulsos*. La señal u^* contiene la misma información que la secuencia $\{u(hk), k = \dots, -1, 0, 1, \dots\}$. Debe notarse, sin embargo, que u^* es una función (generalizada) del tiempo. La señal u^* se introduce para representar una señal muestreada de modo que pueda procesarse con un filtro lineal.

El circuito mantenedor

El circuito mantenedor puede representarse como un integrador que se anula automáticamente después de un período de muestreo. Tal sistema tiene la función de transferencia

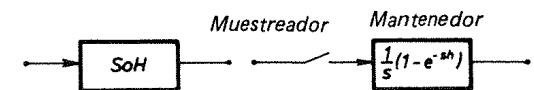
$$G(s) = \frac{1}{s}(1 - e^{-sh}) \quad (4.4)$$

La respuesta a un impulso de la función de transferencia $1/s$ es un escalón unidad y la respuesta a un impulso de $(1/s)\exp(-sh)$ es un escalón unidad retardado h unidades de tiempo. La diferencia de ambas respuestas da la respuesta a un pulso de altura unidad y duración h .

Obsérvese que la ganancia en estado estacionario del circuito mantenedor es $G(0) = h$. En la Sección 2.3 se dice que el muestreo ideal tiene una ganancia $1/h$. Por tanto, la combinación de un muestreador y un mantenedor deben tener una ganancia unidad en estado estacionario. De este modo, para muestreos muy rápidos, el circuito muestreador y mantenedor actúa como un sistema continuo con función de transferencia unidad.

En este caso, el modelo idealizado para el circuito muestreador y mantenedor se obtiene combinando un muestreador con modulación de impulso definido por (4.2) y (4.3) con un circuito mantenedor definido por (4.4). Una representación en diagrama de bloques para este sistema se muestra en la Figura 4.9. Ya que el modulador de impulsos es un sistema periódico, se deduce que el muestreador y mantenedor también es un sistema periódico.

Figura 4.9 Diagrama de bloques de un circuito muestreador y mantenedor y su representación idealizada.



Relaciones entre entrada y salida

Una vez obtenida una representación adecuada para un circuito muestreador y mantenedor, puede calcularse la respuesta de un sistema muestreador para una entrada arbitraria. Considérese el sistema representado en la Figura 4.10a), el cual está formado por un circuito muestreador y mantenedor conectado a un sistema dinámico lineal, invariable en el tiempo de función de transferencia G . Esta es una representación típica de un muestreador y un convertidor D-A conectado a un proceso. El uso del modelo de modulación de impulsos para el muestreador y mantenedor permite representar al sistema como se indica en la Figura 4.10b).

Sea u la entrada e y la salida. Del modelo de modulación de impulso de (4.2) se obtiene

$$u^* = um$$

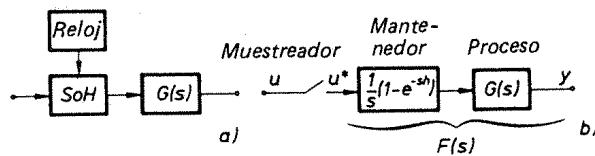


Figura 4.10 Diagrama esquemático de un circuito muestreador y mantenedor conectado a un sistema lineal y su representación mediante el modelo idealizado de un muestreador y mantenedor.

Sea F la función de transferencia de la combinación del mantenedor de orden cero y el proceso; es decir,

$$F(s) = \frac{1}{s}(1 - e^{-sh})G(s) \quad (4.5)$$

La relación entre entrada y salida puede determinarse fácilmente utilizando la teoría de transformadas. La transformada de Laplace de u^* es

$$U^*(s) = \int_0^\infty e^{-st} u^*(t) dt = \sum_{k=0}^{\infty} e^{-skh} u(kh) \quad (4.6)$$

La transformada de Laplace de la señal de salida vendrá definida por

$$Y(s) = F(s) \sum_{k=0}^{\infty} e^{-skh} u(kh) \quad (4.7)$$

Por tanto, la transformada de Laplace de la salida puede obtenerse directamente. Obsérvese, sin embargo, que no es posible extraer el factor de la transformada de Laplace de la señal u del término de la derecha de (4.7). Esto significa que la relación entre la entrada y la salida *no puede* caracterizarse por una función de transferencia normal. Esto se debe a que el sistema no es invariante en el tiempo.

4.4. Respuesta en frecuencia

En muchos casos, las especificaciones de un sistema de control se dan en el dominio de la frecuencia. Un ejemplo típico lo constituyen los pilotos automáticos para aviones. Para garantizar que los modos de oscilación de alta frecuencia no son excitados por el sistema de control, resulta muy crítico asegurar que la respuesta en frecuencia en lazo abierto es menor que 1 en magnitud para altas frecuencias. Un modo de asegurar que se cumplen las especificaciones consiste en realizar un examen de respuesta en frecuencia del sistema completo. Si el regulador se lleva a cabo utilizando un computador digital, es muy importante comprender el fenómeno que ocurre en la respuesta en frecuencia de los sistemas muestreados.

Un caso especial

Cuando se realiza el ensayo de respuesta en frecuencia es normal cortar el lazo de realimentación de la parte analógica, por ejemplo, en el punto A de la Figura 4.1. Para simplificar el análisis consideremos el caso especial en el cual la salida del convertidor D-A es igual a la entrada al convertidor A-D. La labor del computador sobre las señales puede describirse, en este caso, como un circuito muestreador y mantenedor. En la Figura 4.9 se comprueba que un circuito muestreador y mantenedor puede representarse como un muestreador seguido de un circuito mantenedor. El problema se reduce al cálculo de la respuesta de un muestreador seguido de un sistema lineal invariante en el tiempo.

La ecuación (4.2) proporciona la representación muestreada u^* de la señal de entrada u . Una representación formal en serie de Fourier de una secuencia de funciones delta da

$$m(t) = \sum_{k=-\infty}^{\infty} \delta(t - kh) = \frac{1}{h} \left[1 + 2 \sum_{k=1}^{\infty} \cos k\omega_s t \right] \quad (4.8)$$

donde h es el período de muestreo y ω_s es la frecuencia de muestreo correspondiente en radianes por segundo.

Supongamos que la entrada al sistema es

$$u(t) = \sin(\omega t + \phi) = \operatorname{Im}[\exp i(\omega t + \phi)]$$

El desarrollo en serie de la salida $u^* = um$ del muestreador es

$$\begin{aligned} u^*(t) &= \frac{1}{h} \left[\sin(\omega t + \phi) + 2 \sum_{k=1}^{\infty} \cos(k\omega_s t) \sin(\omega t + \phi) \right] \\ &= \frac{1}{h} \left\{ \sin(\omega t + \phi) + \sum_{k=1}^{\infty} [\sin(k\omega_s t + \omega t + \phi) - \sin(k\omega_s t - \omega t - \phi)] \right\} \end{aligned}$$

La señal u^* tiene una componente con la frecuencia ω de la señal de entrada. Esta componente está multiplicada por $1/h$ que es la ganancia del muestreador en estado estacionario. La señal también tiene componentes correspondientes a las *bandas laterales* $k\omega_s \pm \omega$. El contenido de frecuencia de la salida u^* del muestreador se representa en la Figura 4.11. La señal de salida y se obtiene por filtrado lineal de la señal u^* mediante un sistema de función de transferencia $F(s)$. De este modo, la salida tiene componentes a la frecuencia fundamental ω y en las bandas laterales $k\omega_s \pm \omega$.

Para $\omega \neq k\omega_N$, donde ω_N es la frecuencia de Nyquist, la componente fundamental de la salida es

$$y(t) = \frac{1}{h} \operatorname{Im}[F(i\omega) e^{i(\omega t + \phi)}]$$

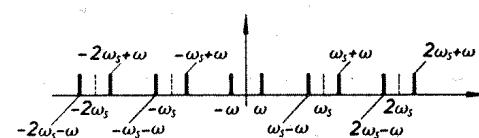


Figura 4.11 Contenido de frecuencia de la señal de entrada muestreada u^* .

Para $\omega = k\omega_N$ la frecuencia de una de las bandas laterales coincide con la frecuencia fundamental. De este modo, a la componente de frecuencia ω contribuyen dos términos. Esta componente es

$$\begin{aligned} y(t) &= \frac{1}{h} \operatorname{Im} \{F(i\omega)e^{i(\omega t + \varphi)} - F(i\omega)e^{i(\omega t - \varphi)}\} \\ &= \frac{1}{h} \operatorname{Im} \{(1 - e^{-2i\varphi})F(i\omega)e^{i(\omega t + \varphi)}\} \\ &= \frac{1}{h} \operatorname{Im} \{2e^{i(\pi/2 - \varphi)} \operatorname{sen} \varphi F(i\omega)e^{i(\omega t + \varphi)}\} \end{aligned}$$

Si la señal de entrada es senoidal de frecuencia ω , se encuentra que la salida contiene la frecuencia fundamental ω y las bandas laterales $k\omega_N \pm \omega$, $k = 1, 2, \dots$ (Compárese con la discusión sobre enmascaramiento de frecuencias de 2.5). La transmisión de la frecuencia fundamental se caracteriza por

$$\hat{F}(i\omega) = \begin{cases} \frac{1}{h} F(i\omega) & \omega \neq k\omega_N \\ \frac{2}{h} F(i\omega)e^{i(\pi/2 - \varphi)} \operatorname{sen} \varphi & \omega = k\omega_N \end{cases} \quad (4.9)$$

Para $\omega \neq k\omega_N$, la transmisión se caracteriza simplemente por una combinación de la función de transferencia del circuito muestreador y mantenedor y el sistema G . El factor $1/h$ se debe a la ganancia del muestreador en estado estacionario.

El hecho de que la transmisión de la señal a la frecuencia de Nyquist ω_N dependa críticamente de φ —es decir, de cómo se sincroniza la señal senoidal de entrada con los instantes de muestreo— se ilustra en la Figura 4.12.

Pueden existir interferencias entre las bandas laterales y la frecuencia fundamental, las cuales pueden producir irregularidades en la salida del sistema. Una ilustración

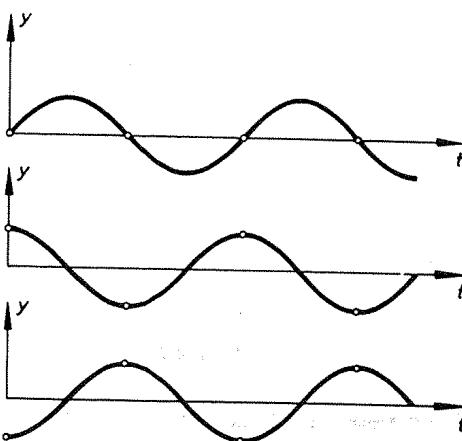


Figura 4.12 Muestreo de una señal sinusoidal a una frecuencia que corresponde a la frecuencia de Nyquist. Nótese que la señal muestreada depende estrechamente de cómo se sincroniza la onda sinusoidal con los instantes de muestreo.

típica de este fenómeno se mostró en el Ejemplo 1.2. En este caso, la componente fundamental tiene una frecuencia de 4,9 Hz y la frecuencia de Nyquist es de 5 Hz. La interacción entre la componente fundamental y la banda lateral baja, que tiene una frecuencia de 5,1 Hz, producirá pulsaciones de 0,1 Hz de frecuencia. Esto se aprecia claramente en la Figura 1.4.

Si se filtran las frecuencias de las bandas laterales, el sistema muestreado se presenta como lineal e invariable en el tiempo excepto en las frecuencias que son múltiplos de la frecuencia de Nyquist, $\omega_N/2$. A esta frecuencia, la amplitud y el retardo de fase dependen del desplazamiento de fase de la entrada respecto a los instantes de muestreo.

Si se intenta determinar la respuesta en frecuencia de un sistema muestreado utilizando las propias técnicas de la respuesta en frecuencia, es importante filtrar eficazmente las bandas laterales. Incluso con un filtrado perfecto se presentarán problemas en la frecuencia de Nyquist. El resultado depende críticamente del modo en que se sincronice la entrada con el reloj del sistema.

El caso general

Es fácil extender el análisis al caso general del sistema mostrado en la Figura 4.1. El correspondiente sistema en lazo abierto se muestra en la Figura 4.13. Está formado por un convertidor A-D, el computador, un convertidor D-A, y el proceso. Se supone que el convertidor D-A mantiene la señal constante durante un intervalo de muestreo. También se supone que los cálculos efectuados por el computador pueden expresarse por la función de transferencia discreta $H(z)$ y que el proceso se describe por la función de transferencia $G(s)$.

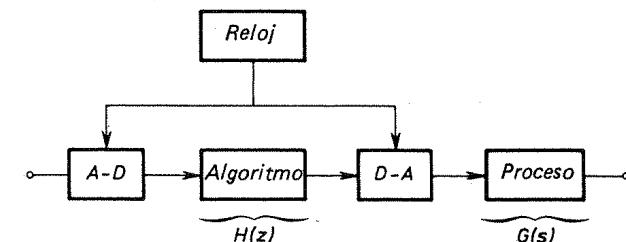


Figura 4.13 Sistema controlado con computador en lazo abierto.

Si se aplica una sinusode

$$v(t) = \operatorname{sen}(\omega t + \varphi) = \operatorname{Im}[\exp i(\omega t + \varphi)]$$

a un convertidor A-D, el computador generará una secuencia de pulsos que en el estado estacionario puede describirse por

$$w(kh) = \operatorname{Im} \{H(e^{i\omega h})e^{i(\omega kh + \varphi)}\}, \quad k = \dots, -1, 0, 1, \dots$$

Esta secuencia se aplica al convertidor D-A. Como el convertidor D-A mantiene la señal constante durante un período de muestreo, la salida es similar a la obtenida cuando se aplica una señal ω directamente a un circuito mantenedor. De este modo, se puede aplicar la discusión de la sección anterior: la salida contiene la componente fundamental de frecuencia ω y las bandas laterales $k\omega_s \pm \omega$. La transmisión de señal de la componente fundamental puede describirse por la función de transferencia

$$K(i\omega) = \begin{cases} \frac{1}{h} H(e^{i\omega h}) F(i\omega) & \omega \neq k\omega_N \\ \frac{2}{h} H(e^{i\omega h}) F(i\omega) e^{i(\pi/2 - \varphi)} \operatorname{sen} \varphi & \omega = k\omega_N \end{cases} \quad (4.10)$$

donde ω_N es la frecuencia de Nyquist y

$$F(s) = \frac{1}{s} (1 - e^{-sh}) G(s) \quad (4.11)$$

Cuando ω no es múltiplo de la frecuencia de Nyquist, la transmisión de señal de la componente fundamental puede caracterizarse por una función de transferencia que es el producto de cuatro términos: la ganancia $1/h$ del muestreador, la función de transferencia $[1 - \exp(-sh)]/s$ del circuito mantenedor, la función de transferencia discreta $H[\exp(sh)]$ del algoritmo del computador, y la función de transferencia $G(s)$ del proceso. Obsérvese, sin embargo, que hay otras frecuencias en la salida del sistema debidas al muestreo. A la frecuencia de Nyquist la componente fundamental y la banda lateral inferior coinciden.

Se deduce de la exposición que el circuito mantenedor puede interpretarse como un filtro. Las funciones del circuito mantenedor de orden cero y de primer orden se muestran en la Figura 4.14. Se observa claramente en la figura que ambos mantenedores permiten una transmisión significativa de la señal para frecuencias superiores a la de Nyquist $\omega_N = \pi/h$. Obsérvese que la curva de fase es discontinua

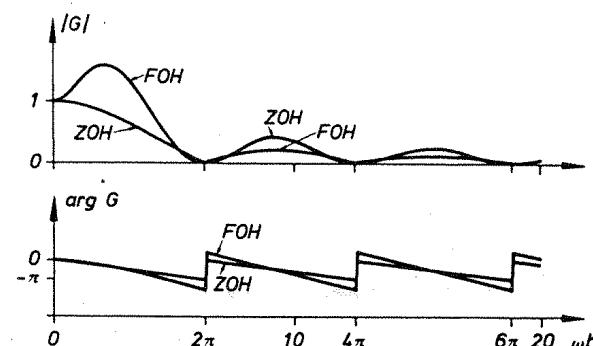


Figura 4.14 Curvas de módulo y argumento de la función de transferencia de los circuitos de mantenedores de orden cero y de primer orden.

para los argumentos $\omega h = 2k\pi$, $k = 1, 2, \dots$. Ya que la fase tiene módulo 2π , las discontinuidades pueden ser π o $-\pi$. En la figura se muestran únicamente en π por conveniencia.

El siguiente ejemplo ilustra el cálculo e interpretación de la respuesta en frecuencia de un sistema muestreado.

Ejemplo 4.1

Considérese un sistema formado por un muestreador y un mantenedor de orden cero definido por (4.4) seguidos por un sistema lineal con la función de transferencia

$$G(s) = \frac{1}{s + 1}$$

El período de muestreo es $h = 0,05$ s. La frecuencia de Nyquist es, por tanto, $\pi/0,05 = 61,8$ rad/s. La Figura 4.15 muestra el diagrama de Bode del sistema. Para poder comparar se representa en la misma figura el diagrama de Bode de la función de transferencia G . Las

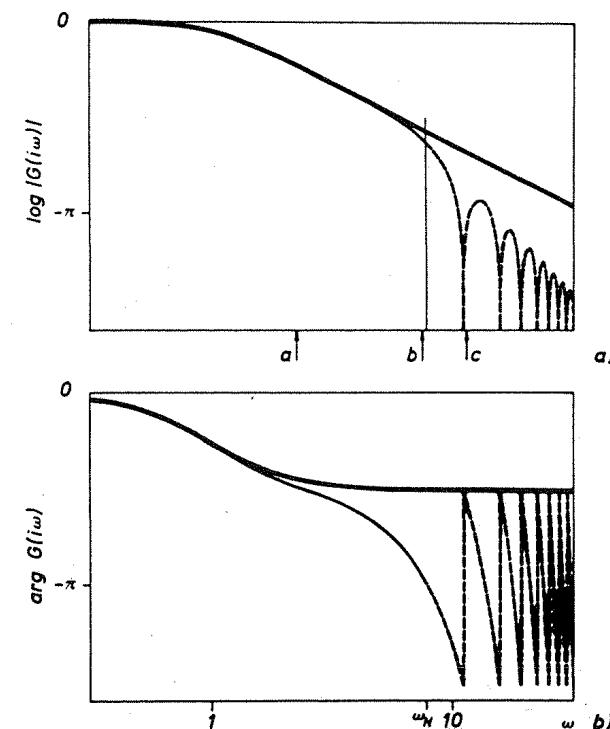


Figura 4.15 Diagrama de bloques de un muestreador y mantenedor de orden cero seguido por un retardo de primer orden. El período de muestreo es 0,05 s.

curvas son similares para frecuencias mucho menores que la frecuencia de Nyquist. Las desviaciones aparecen primero en las curvas de fase. Para $\omega = 0,1 \omega_N$ las curvas de fase difieren aproximadamente en 10° . Para frecuencias múltiplos de la frecuencia de muestreo ω_s no hay transmisión de señal porque la función de transferencia del mantenedor de orden cero se anula para estas frecuencias. La curva de fase también es discontinua en estas frecuencias. (Compárese con la Fig. 4.14.) Obsérvese también la ambigüedad de la función de transferencia a frecuencias múltiplos de la frecuencia de Nyquist. El conjunto de valores a ω_N se indica en la Figura 4.15 con una línea.

La interpretación del diagrama de Bode requiere algunas consideraciones debido a la modulación que introduce el muestreo. Si se introduce una señal senoidal de frecuencia ω , la señal de salida es la suma de la señal senoidal y todas sus frecuencias enmascaradas. Esto se ilustra en la Figura 4.16, la cual muestra el estado estacionario de la salida para diferentes frecuencias. Para frecuencias menores que la frecuencia de Nyquist, es dominante la contribución de la frecuencia fundamental. A frecuencias cercanas a la frecuencia de Nyquist existe una interacción sustancial con el primer enmascaramiento, $\omega_s - \omega$. De este modo se obtienen oscilaciones típicas. A la frecuencia de Nyquist, la señal y su primer enmascaramiento tienen la misma frecuencia y magnitud. La señal resultante depende entonces del desplazamiento de fase entre las señales. Para frecuencias superiores a la frecuencia de Nyquist, domina la contribución de las frecuencias enmascaradas del intervalo $(0, \omega_N)$.

Esto muestra la importancia de filtrar una señal antes de muestrearla, de modo que la transmisión de señal por encima de la frecuencia de Nyquist sea insignificante. Compárese esta conclusión con las de la Sección 2.8.

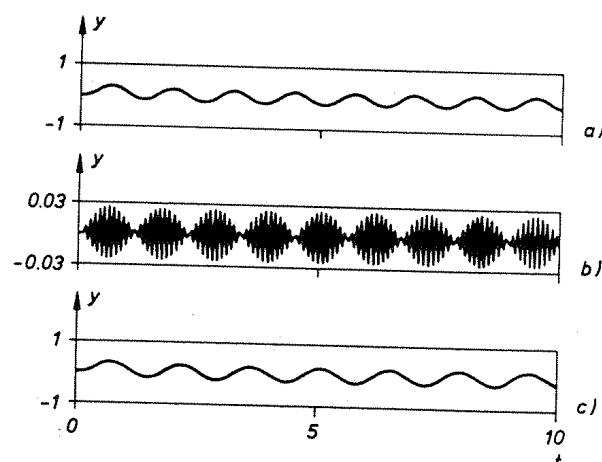


Figura 4.16. Respuestas en estado estacionario a sinusoides con diferentes frecuencias de un mantenedor de orden cero seguido de un sistema de primer orden con constante de tiempo unidad. El periodo de muestreo es 0,05 s. Las frecuencias son 5 rad/s en a), 60 rad/s en b) y 130 rad/s en c). Estas están indicadas por flechas en la Figura 4.15.

Una aplicación

Un motor de combustión interna es un ejemplo típico de un sistema que es inherentemente muestreado. El muestreado está causado por el mecanismo de ignición y su frecuencia es el número de cilindros independientes dividido por el tiempo empleado en un ciclo completo.

Cuando se intentó investigar la respuesta dinámica de los motores se obtuvieron resultados fácilmente reproducibles para frecuencias inferiores a la de muestreo. Sin embargo, durante mucho tiempo los resultados para frecuencias superiores fueron irregulares: se obtuvieron diferentes resultados en las medidas y los resultados de los experimentos no pudieron verificarse al repetir éstos. Esto se debe a la naturaleza muestreada del proceso. Para señales de entrada de frecuencia próxima a la frecuencia de Nyquist hay interferencia de las bandas laterales. A la frecuencia de Nyquist, el resultado depende de cómo se sincronice la sinusoida con los pulsos de ignición.

Cuando se comprendió el origen del problema fue fácil encontrar la solución. Simplemente se sincronizó la sinusoida con los pulsos de ignición, entonces fue posible medir la respuesta en frecuencia a altas frecuencias. Un resultado típico se muestra en la Figura 4.17. Obsérvese, en particular, que la medida se efectúa en un intervalo de frecuencias que incluye la frecuencia de Nyquist.

4.5. Formalismo de la función de transferencia discreta

Los sistemas lineales y continuos pueden describirse, analizarse y sintetizarse convenientemente utilizando métodos algebraicos. Cuando se desarrolló la teoría de sistemas muestreados, fue natural el intento de desarrollar herramientas algebraicas similares. Muchos de los primeros desarrollos de la teoría de sistemas muestreados se realizaron en esta dirección.

Este enfoque es útil, simple y seguro si se considera el sistema desde el punto de vista del computador o si el proceso se observa en los instantes de sincronización con el reloj del computador, puesto que entonces el sistema es *invariable en el tiempo*. (Para un análisis apropiado, véase el Capítulo 3.) Sin embargo, cuando el sistema se analiza desde el punto de vista del proceso, como se hace en este capítulo, el sistema es variable en el tiempo. De este modo, el método algebraico pierde su simplicidad, porque la multiplicación de funciones temporales no commuta con operadores diferenciales y operadores en diferencias.

Para completar se da una breve descripción de la teoría de sistemas algebraicos en los casos más complicados. La razón principal es histórica. Gran parte de la teoría de sistemas muestreados se desarrolló originalmente utilizando estos métodos, que también se emplean en muchas publicaciones.

Objetivos

Antes de entrar en detalles, resulta útil establecer los objetivos. El propósito principal consiste en desarrollar un formalismo para manipular la descripción de

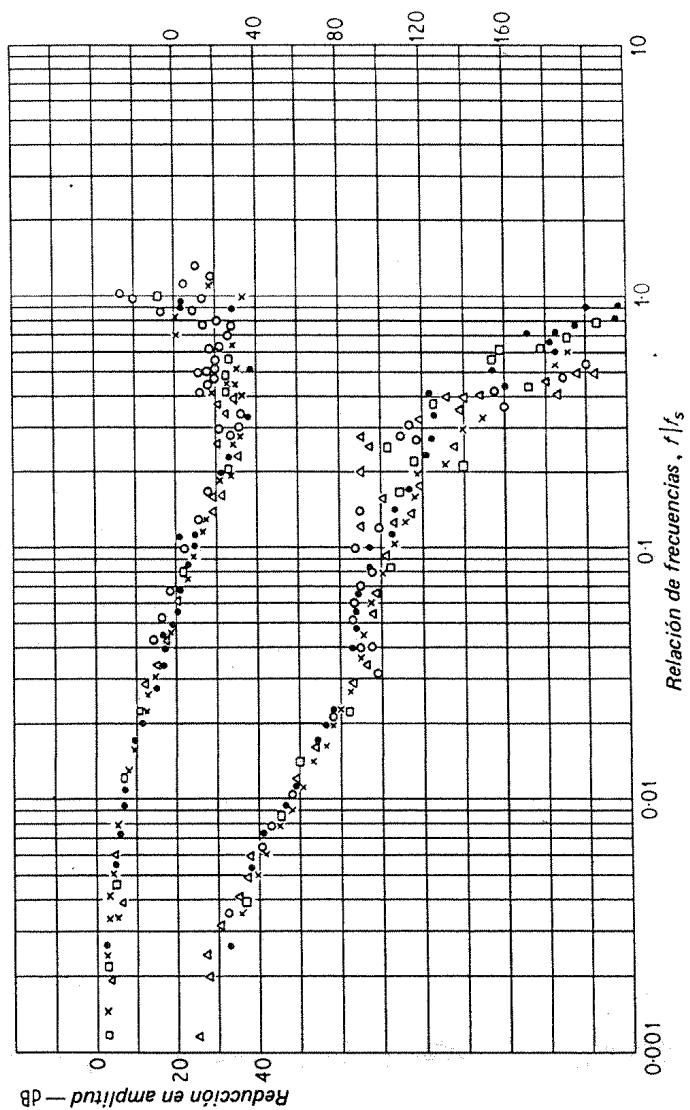


Figura 4.17 Medida de la respuesta en frecuencia de un motor diesel. Las frecuencias se han normalizado respecto a la frecuencia de muestreo. [Figura tomada de D. E. Bowns, «The Dynamic Transfer Characteristics of Reciprocating Engines», *Proc. Mech. Engrs.*, 185 (1971) con autorización.]

sistemas. El formalismo tendrá algunas propiedades comunes con los métodos de transformadas para sistemas lineales invariantes en el tiempo. Cada convertidor A-D y D-A se asocia con una operación de muestreo. Ya que el muestreo puede definirse como una modulación en amplitud, las partes que varían en el tiempo se asociarán con estas operaciones. De este modo, el sistema puede separarse en diferentes partes: algunas partes son sistemas lineales invariantes en tiempo que pueden manejarse con los métodos normales de transformadas; las otras partes están formadas por muestreadores que son intrínsecamente variables en el tiempo.

La transformada Z

La Sección 3.5 introduce la transformada z como una aplicación de secuencias sobre funciones de variable compleja. Una transformada z diferente cuyo dominio son funciones continuas puede definirse del siguiente modo:

Definición 4.1 – La transformada z. La transformada z de una función continua se define como

$$\tilde{F}(z) = \sum_{k=0}^{\infty} z^{-k} f(kh) \quad (4.12)$$

La transformada inversa viene dada por

$$f(kh) = \frac{1}{2\pi i} \oint_{\Gamma} z^{k-1} \tilde{F}(z) dz \quad (4.13)$$

donde el contorno de integración Γ encierra todas las singularidades del término integral.

La transformada z de una señal continua se obtiene muestreando la señal y aplicando la transformada z a la secuencia muestreada. Como la transformada depende sólo de los valores en los instantes de muestreo, todas las funciones temporales que concuerden en los instantes de muestreo tienen la misma transformada.

Obsérvese que la transformada está relacionada con el reloj del sistema, el cual define los instantes de muestreo. Obsérvese también que la transformada inversa define la función únicamente en los instantes de muestreo.

Estas propiedades de la transformada z de una función continua no se comprenden fácilmente y conducen a muchas confusiones y a algunos errores.

Dos teoremas básicos

Para desarrollar un álgebra que permita una manipulación formal del sistema se necesitan dos teoremas. El primer teorema relaciona la transformada z de una función continua con su transformada de Laplace.



Teorema 4.1. Sea la función f con transformada de Laplace F y transformada M , y sea F^* la transformada de Laplace de la representación muestreada f^* de f . Supongamos que para algún $\varepsilon > 0$, $|F(s)| \leq |s|^{-1-\varepsilon}$ para $|s|$ elevados, entonces

$$F^*(s) = \tilde{F}(e^{sh}) = \frac{1}{h} \sum_{k=-\infty}^{\infty} F(s + ik\omega_s) \quad (4.14)$$

donde $\omega_s = 2\pi/h$ es la frecuencia de muestreo.

Demuestração. La definición de F^* da

$$\begin{aligned} F^*(s) &= \int_0^\infty e^{-st} f^*(t) dt \\ &= \int_0^\infty e^{-st} f(t) m(t) dt \\ &= \int_0^\infty e^{-st} f(t) \left\{ \sum_{k=-\infty}^{\infty} \delta(t - kh) \right\} dt \end{aligned}$$

donde la última igualdad se obtiene de (4.3). Intercambiando el orden de integración y suma se obtiene

$$\begin{aligned} F^*(s) &= \sum_{k=0}^{\infty} \int_0^\infty e^{-st} f(t) \delta(t - kh) dt \\ &= \sum_{k=0}^{\infty} (e^{sh})^{-k} f(kh) \\ &= \tilde{F}(e^{sh}) \end{aligned} \quad (4.15)$$

La última igualdad se obtiene de (4.12).

Como la transformada de Laplace de un producto de dos funciones es una convolución de sus transformadas, se obtiene que

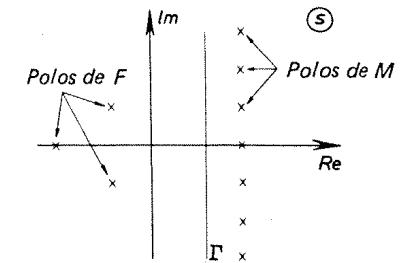
$$\begin{aligned} F^*(s) &= F(s) * M(s) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} F(v) M(s-v) dv \\ &= \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} F(v) \frac{1}{1 - e^{-hv}} dv \end{aligned} \quad (4.16)$$

El contorno de integración debería estar a la derecha de todos los polos de F y a la izquierda de todos los polos de M (véase la Fig. 4.18).

Si F se anula más rápidamente que $|s|^{-1-\varepsilon}$ cuando $|s| \rightarrow \infty$, la integral de FM se anula en un semicírculo de gran radio. En lugar de la integración completa sobre un semicírculo en la derecha del plano complejo, se puede realizar el cálculo de residuos para resolver la integral. En el dominio encerrado por el contorno el término integral tiene polos simples en los ceros de

$$\exp(h(s-v)) = 1$$

Figura 4.18 Singularidades de F y M y el contorno de integración.



es decir, en

$$v_k = s + \frac{2\pi ik}{h} = s + ik\omega_s, \quad k = \dots, -1, 0, 1, \dots$$

Los residuos de estos polos son

$$\frac{1}{h} F\left(s + \frac{2\pi ik}{h}\right)$$

La suma de los residuos proporciona (4.14). \square

Nota 1. Obsérvese que la ecuación (4.14) también se puede escribir

$$F^*(s) = \frac{1}{h} [F(s) + F(i\omega_s + s) + F(s - i\omega_s) + \dots] \quad (4.17)$$

Nota 2. Obsérvese que si F es analítica para $\operatorname{Re} s < -\gamma_0$, el contorno de integración en (4.16) puede ser un semicírculo de gran radio situado en la parte izquierda. Se obtiene la siguiente fórmula:

$$\tilde{F}(z) = \sum_{\text{Polos de } F} \operatorname{Res} \left\{ F(s) \frac{z}{z - \exp(sh)} \right\} \quad (4.18)$$

Esto demuestra la fórmula (3.43).

Nota 3. El teorema puede extenderse al caso en que la función F tiende a cero de la misma forma que $1/|s|$ para $|s|$ elevados. La ecuación (4.14) se reemplaza entonces por

$$F^*(s) = \frac{1}{h} \sum_{k=-\infty}^{\infty} F(s + ik\omega_s) + \frac{1}{2} f(0+) \quad (4.19)$$

Nota 4. En la literatura algunas veces se utiliza la misma notación para las funciones F^* y \tilde{F} . Debería evitarse esta confusión.

Nota 5. Obsérvese que la fórmula (4.14) está relacionada con la fórmula (2.4) por la transformada de Fourier de una señal muestreada.

Funciones de transferencia discreta. La Sección 4.3 muestra que la relación entre la entrada y la salida de un muestreador seguido de una función de transferencia lineal

está definida por la ecuación (4.7). Esta ecuación no puede describirse mediante una función de transferencia. Si se añade un muestreador ficticio a la salida se obtiene la configuración mostrada en la Figura 4.19. Para este sistema es posible definir una función de transferencia. La relación entre la entrada y la salida viene dada por

$$y^*(t) = [f(t)*u^*(t)]^* \quad (4.20)$$

El siguiente teorema es útil para obtener las transformadas correspondientes.

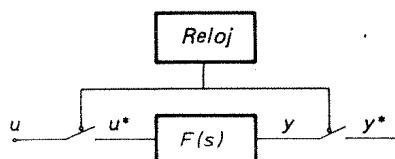


Figura 4.19 Diagrama de bloques de un sistema con dos muestreadores.

Teorema 4.2. Sean f y g funciones que tienen transformadas de Laplace y sea m la función de modulación correspondiente a un tren de impulsos. Entonces

$$m(t)[f(t)*[m(t) \cdot g(t)]] = [m(t) \cdot f(t)]*[m(t) \cdot g(t)] \quad (4.21)$$

o, lo que es equivalente,

$$[f(t)*g^*(t)]^* = f^*(t)*g^*(t) \quad (4.22)$$

Demostración. La definición de convolución permite escribir el término de la izquierda de (4.21) como

$$\begin{aligned} (f*g^*)^*(t) &= m(t) \int_{-\infty}^{\infty} f(t - \tau) g^*(\tau) d\tau \\ &= \int_{-\infty}^{\infty} m(t) f(t - \tau) m(\tau) g(\tau) d\tau \end{aligned}$$

Análogamente, el lado derecho de la ecuación (4.22) puede escribirse como

$$\begin{aligned} (f^* * g^*)(t) &= \int_{-\infty}^{\infty} m(t - \tau) f(t - \tau) m(\tau) g(\tau) d\tau \\ &= \int_{-\infty}^{\infty} m(t) f(t - \tau) m(\tau) g(\tau) d\tau \end{aligned}$$

La última igualdad se mantiene porque $m(\tau) \neq 0$ solamente para $\tau = nh$ y $m(t - nh) = m(t)$. \square

Nota 1. La transformada de Laplace de (4.21) da

$$[F(s)G^*(s)]^* = F^*(s)G^*(s) \quad (4.23)$$

Nota 2. Obsérvese que la multiplicación por m fuera de las llaves de (4.21) puede interpretarse como la introducción de un muestreador ficticio.

Formalismo

Es inmediato desarrollar un formalismo para tratar sistemas muestreados. En primer lugar, un sistema se representa por un diagrama de bloques. Cada convertidor A-D se representa por un muestreador ideal. Cada convertidor D-A se representa por un circuito mantenedor de función de transferencia (4.4). Los bloques lineales y continuos se representan por su función de transferencia, y los cálculos lineales en el computador, por su función de transferencia discreta. La unión entre muestreadores puede reducirse utilizando las reglas normales para sistemas lineales invariantes en el tiempo. Entonces pueden escribirse las ecuaciones que definen el sistema. Los Teoremas 4.1 y 4.2 se emplean para reescribir estas ecuaciones. El procedimiento se ilustra con un ejemplo.

Ejemplo 4.2

Considérese la configuración típica de un sistema controlado por computador como el de la Figura 4.20a). El proceso se caracteriza por la función de transferencia lineal G , y los

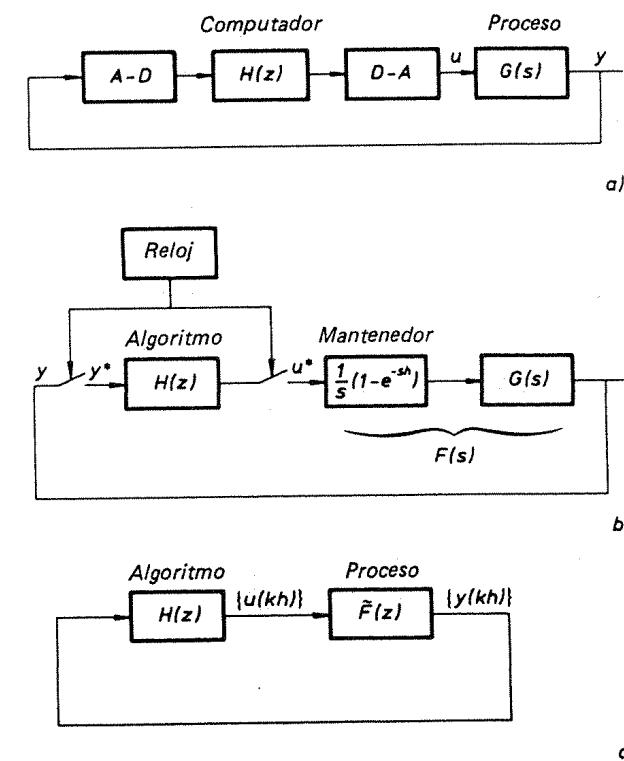


Figura 4.20 Configuración típica de un sistema controlado con computador.



cálculos efectuados en el computador se representa por la función de transferencia discreta H . Las partes analógicas y digitales del sistema están conectadas, como es usual, a través de los convertidores D-A y A-D. Para aplicar el formalismo, el convertidor A-D se representa por un muestreador ideal. El computador se representa como un sistema que transforma una señal modulada en impulsos en otra del mismo tipo. El convertidor D-A se representa por un muestreador seguido de un mantenedor de orden cero. Se supone que los muestreadores están perfectamente sincronizados. En este caso se obtiene el diagrama de bloques de la Figura 4.20b). De este modo, las partes analógicas son el mantenedor y el proceso. Su función de transferencia combinada es

$$F(s) = \frac{1}{s}(1 - e^{-sh})G(s)$$

La transformada de Laplace Y de la salida es

$$Y(s) = F(s)U^*(s)$$

La salida muestreada tiene la forma

$$Y^*(s) = [F(s)U^*(s)]^* = F^*(s)U^*(s)$$

donde se utiliza el teorema 4.2 para obtener la última igualdad. La relación entre y^* y u^* puede representarse por la función de transferencia discreta

$$\tilde{F}(z) = F^*(s)|_{s=(\ln z)/h}$$

Los cálculos en el computador pueden, por tanto, representarse por la función de transferencia discreta $H(z)$. Si se corta el lazo en el computador la función de transferencia discreta es

$$H(z)\tilde{F}(z)$$

La Figura 4.20c) muestra un diagrama de bloques de las propiedades del sistema que pueden observarse desde el computador. Considerando todas las señales como secuencias del tipo $\{y(kh), k = \dots, -1, 0, 1, \dots\}$ e introduciendo funciones de transferencia discretas apropiadas para el algoritmo y el proceso con el muestreador y mantenedor, puede obtenerse una representación que es equivalente a la representación del diagrama de bloques usual.

Un ejemplo ligeramente más complicado ilustrará mejor esto.

Ejemplo 4.3

El sistema ilustrado en la Figura 4.21a) tiene dos realimentaciones analógicas, y_1 e y_2 , y una consigna analógica, u_c . La señal analógica es examinada por un multiplexor y convertida a digital. El computador calcula la señal de control, la cual alimenta al proceso a través de un convertidor D-A. La Figura 4.21b) se obtiene por el procedimiento mostrado en el Ejemplo 4.2.

Introduciendo

$$F_1(s) = G_1(s)\frac{1}{s}(1 - e^{-sh})$$

$$F_2(s) = G_2(s)F_1(s)$$

Entonces las transformadas de Laplace de las salidas son

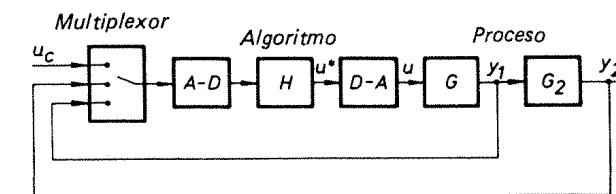
$$Y_1(s) = F_1(s)U^*(s)$$

$$Y_2(s) = F_2(s)U^*(s)$$

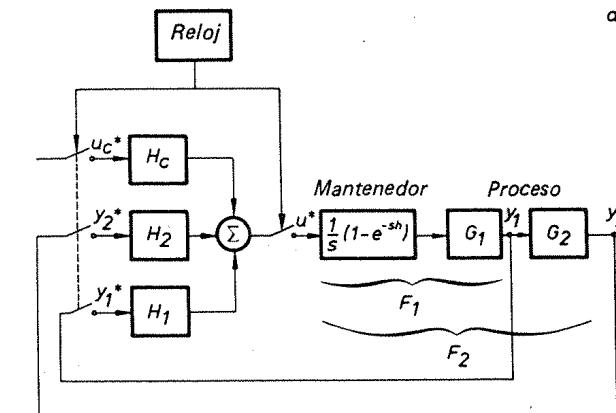
Por tanto,

$$Y_1^*(s) = [F_1(s)U^*(s)]^* = F_1^*(s)U^*(s)$$

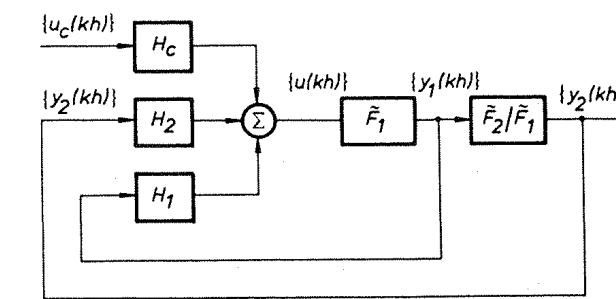
$$Y_2^*(s) = [F_2(s)U^*(s)]^* = F_2^*(s)U^*(s)$$



a)



b)



c)

Figura 4.21 Sistema controlado con computador con multiplexor y dos lazos de realimentación y los diagramas de bloques equivalentes.

Se obtiene de (4.14) y (4.23) que

$$Y_1(z) = \tilde{F}_1(z)U(z)$$

$$Y_2(z) = \tilde{F}_2(z)U(z)$$

Representemos los cálculos realizados por el computador por

$$U(z) = H_c(z)U_c(z) - H_1(z)Y_1(z) - H_2(z)Y_2(z)$$

La relación entre la salida, Y_2 , y la señal de consigna muestreada, U_c , se obtiene eliminando Y_1 en la ecuación precedente. Esto proporciona

$$Y_2(z) = \frac{H_c(z)\tilde{F}_2(z)}{1 + H_1(z)\tilde{F}_1(z) + H_2(z)\tilde{F}_2(z)} U_c(z)$$

Obsérvese, sin embargo, que la relación entre la señal analógica y_1 y u_c no puede representarse por una simple función de transferencia discreta debido a la naturaleza periódica del sistema muestreado.

Con la introducción de las señales muestreadas como secuencias y de las funciones de transferencia discreta, el sistema se puede representar como en la Figura 4.21c).

La transformada Z modificada

El problema de muestrear un sistema con un retardo puede realizarse con la transformada z modificada, que se define de la siguiente manera.

Definición 4.2—La transformada z modificada. La transformada z modificada de una función continua se define como

$$\tilde{F}(z, m) = \sum_{k=0}^{\infty} z^{-k} f(kh - h + mh), \quad 0 \leq m \leq 1$$

La transformada inversa es

$$f(nh - h + mh) = \frac{1}{2\pi i} \int_{\Gamma} \tilde{F}(z, m) z^{n-1} dz$$

donde el contorno Γ encierra todas las singularidades del término integral.

La transformada z modificada resulta útil en muchos casos; por ejemplo, puede investigarse fácilmente el comportamiento entre instantes de muestreo utilizando esta transformada. Existen tablas extensas de transformadas z modificadas y algunos teoremas sobre sus propiedades (véanse las Referencias).

4.6. Muestreo multiperiodico

Hasta ahora sólo se han estudiado sistemas en los cuales las conversiones A-D y D-A se efectúan a la misma frecuencia. El Ejemplo 3.15 indica que esto puede ser

ventajoso para realizar la conversión D-A más rápidamente. También hay situaciones donde puede ocurrir lo contrario. Por ejemplo, es difícil realizar filtros anti-enmascaramiento con constantes de tiempo grandes utilizando técnicas analógicas. En estos casos es mucho más fácil muestrear las señales rápidamente con un filtro anti-enmascaramiento y después filtrar digitalmente. En ambos casos, el sistema tiene dos muestreadores que operan a diferentes frecuencias. Esto se denomina *muestreo multiperiodico*. Este esquema de muestreo puede ser necesario en sistemas con transmisión especial de datos o sensores y actuadores especiales y son útiles para mejorar la respuesta de sistemas en los que las muestras se toman a baja frecuencia, por ejemplo, cuando se usa instrumentación de laboratorio. Los sistemas multiperiodicos pueden permitir un mejor control de las variables entre los instantes de muestreo. En sistemas multivariables también puede ser ventajoso disponer de diferentes frecuencias de muestreo en diferentes lazos para reducir la carga de cálculo del computador y mejorar la precisión. El uso de muestreo multiperiodico también es normal en los sistemas multiprocesadores.

Un tratamiento detallado de los sistemas multiperiodicos está fuera de las pretensiones de este libro; sin embargo, se realizará una breve exposición de las ideas más importantes, para mostrar cómo pueden extenderse los métodos presentados en este libro para cubrir los sistemas multiperiodicos.

Descripción en el espacio de estado

Consideremos un sistema compuesto de dos subsistemas dinámicos, continuos y de coeficientes constantes. Supongamos que hay dos muestreadores periódicos con período h_1 y h_2 . Sea la razón entre los períodos un número racional $h_1/h_2 = m_1/m_2$ donde m_1 y m_2 no tienen factores comunes. Entonces existen un número entero mínimo m y un número real h tal que

$$h_1 = \frac{hm_1}{m} \quad (4.24)$$

$$h_2 = \frac{hm_2}{m} \quad (4.25)$$

y

$$m = m_1 m_2 \quad (4.26)$$

Si los muestreadores están sincronizados se deduce que las señales de control serán constantes sobre períodos de muestreo de longitud h/m . Muestreando con este período se obtiene un sistema discreto periódico de período h . Entonces el sistema se puede describir como un sistema discreto constante si los valores de las variables del sistema se consideran únicamente en múltiplos enteros de h . En este caso se puede aplicar la teoría usual de sistemas discretos. Un ejemplo ilustra esta idea.

Ejemplo 4.4

Consideremos el sistema mostrado en la Figura 4.22, el cual tiene dos subsistemas y dos muestreadores con períodos 0,5 y 1. Se supone que los muestreadores están sincronizados. También se supone que los circuitos mantenedores están incluidos en los subsistemas. Si los subsistemas se muestran con período 0,5 y este valor se elige como unidad de tiempo, entonces

$$\begin{cases} x_1(k+1) = \Phi_1 x_1(k) + \Gamma_1 u_1(k) \\ y_1(k) = C_1 x_1(k) \end{cases}$$

$$\begin{cases} x_2(k+1) = \Phi_2 x_2(k) + \Gamma_2 u_2(k) \\ y_2(k) = C_2 x_2(k) \end{cases}$$

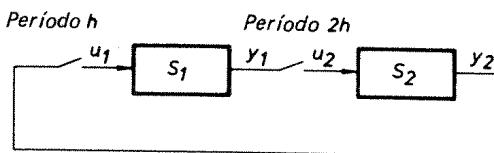


Figura 4.22 Diagrama de bloques de un sistema multiperiodico sencillo.

Las interconexiones se describen por

$$u_1(k) = y_2(k), \quad k = \dots -1, 0, 1, 2, \dots$$

$$u_2(k) = y_1(k), \quad k = \dots -2, 0, 2, 4, \dots$$

El sistema es periódico, con un período igual a dos intervalos de muestreo. Se puede obtener una descripción invariable en el tiempo considerando las variables del sistema solamente en períodos de muestreo pares. Mediante cálculo directo se obtiene

$$\begin{bmatrix} x_1(2k+2) \\ x_2(2k+2) \end{bmatrix} = \begin{bmatrix} \Phi_1^2 + \Gamma_1 C_2 \Gamma_2 C_1 & \Phi_1 \Gamma_1 C_2 + \Gamma_1 C_2 \Phi_2 \\ (\Phi_2 \Gamma_2 + \Gamma_2) C_1 & \Phi_2^2 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} \quad (4.27)$$

Esta ecuación puede utilizarse para obtener la respuesta de un sistema multiperiodico. Por ejemplo, la condición de estabilidad exige que la matriz del término de la derecha de (4.27) tenga todos sus valores propios dentro del círculo unidad. Los valores de las variables de estado en los períodos de muestreo impares están definidas por

$$\begin{bmatrix} x_1(2k+1) \\ x_2(2k+1) \end{bmatrix} = \begin{bmatrix} \Phi_1 & \Gamma_1 C_2 \\ \Gamma_2 C_1 & \Phi_2 \end{bmatrix} \begin{bmatrix} x_1(2k) \\ x_2(2k) \end{bmatrix}$$

El análisis mostrado en el ejemplo puede extenderse a un número arbitrario de muestreadores a condición de que las relaciones entre los períodos de muestreo sean números racionales. También se puede tratar el muestreo retardado con los métodos descritos en la Sección 3.2.

Modelos de entrada-salida

Los sistemas multiperiodicos también pueden estudiarse mediante análisis de entrada-salida. En primer lugar, observemos como en el caso anterior que el sistema es periódico de período h si la relación entre los períodos de muestreo son números racionales. En este caso, los valores de las variables del sistema en los instantes en que están sincronizadas con el período, pueden describirse como un sistema dinámico invariable. Entonces, pueden utilizarse los métodos usuales de operadores o funciones de transferencia para sistemas lineales.

A continuación se describe el procedimiento para analizar un sistema: en primer lugar se dibuja un diagrama de bloques del sistema incluyendo todos los subsistemas y todos los muestreadores. Se determina el período h . Todos los muestreadores que aparecen en el sistema tienen períodos h/m , donde m es entero. Se utiliza un truco denominado *descomposición de interruptores* para convertir muestreadores de relación h/m en una combinación de muestreadores de período h . Entonces el sistema se puede analizar utilizando los métodos descritos en la Sección 4.5.

Descomposición de interruptores. Para comprender el concepto de la descomposición de interruptores, consideremos en primer lugar un muestreador con período $h/2$. Tal muestreador se puede obtener combinando un muestreador de período h con otro muestreador de período h retrasado $h/2$. El esquema se ilustra en la Figura 4.23a). La idea puede extenderse fácilmente para muestreadores de relación h/m , donde m es un entero arbitrario [véase la Fig. 4.23b)].

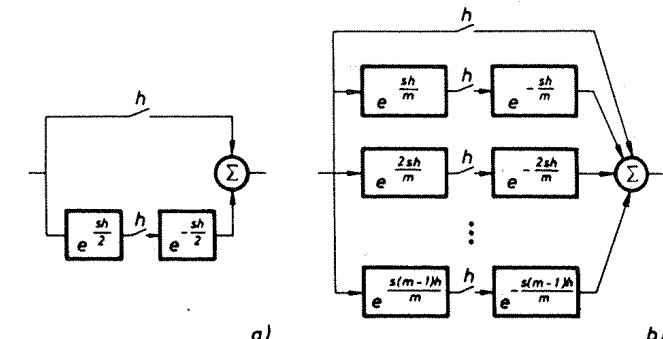


Figura 4.23 Representación de muestreadores con períodos: a) $h/2$ y b) h/m empleando una descomposición de los interruptores.

Sistemas multiperiodicos con períodos no racionales

El método descrito hasta ahora funcionará solamente cuando las relaciones entre períodos de muestreo sean números racionales. Si esto no ocurre, no es posible

obtener un sistema periódico; en este caso se pueden utilizar técnicas diferentes. Las técnicas multiperíodicas también conducen a análisis complicados si existen muchos muestreadores con una amplia variedad de períodos.

4.7. Problemas

- 4.1. Encontrar Y^* para los sistemas de la Figura 4.24.

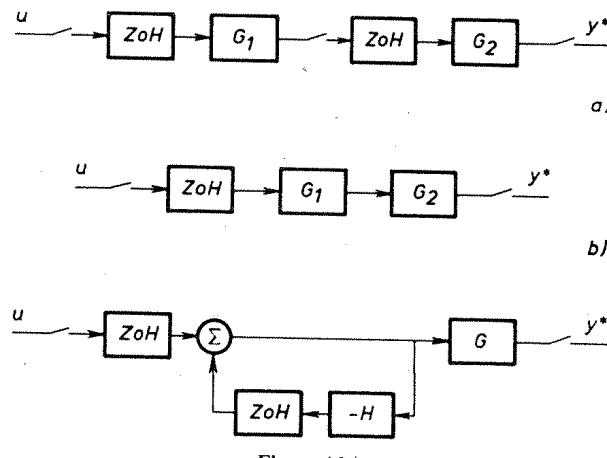


Figura 4.24

- 4.2. Escribir un programa para calcular la respuesta en frecuencia de un sistema muestreado. Considerar los siguientes datos como entradas al programa:

- Los polinomios de la función de transferencia discreta $H(z)$.
- El intervalo de muestreo.
- Las frecuencias máxima y mínima.

Utilizar el programa para dibujar $H[\exp(i\omega h)]$ para el motor normalizado muestreado con un mantenedor de orden cero y comparar con el sistema continuo.

4.8. Referencias

El desarrollo de este capítulo corresponde al tratamiento clásico de los sistemas muestreados. El modelo de modulación fue propuesto por MacColl (1945) y elaborado posteriormente por Linvil (1951). Un tratamiento más detallado se da en los textos clásicos de Ragazzini y Franklin (1958) y Jury (1958), los cuales se mencionan en la Sección 1.6. La aproximación al muestreador ideal se trata en

- LI, Y. T., J. L. MEIRY, y R. E. CURRY (1972): «On the Ideal-Sampler Approximation», *IEEE Trans.*, AC-17, 167-68.

La respuesta en frecuencia es importante desde el punto de vista del análisis y diseño. Un tratamiento completo de este problema se encuentra en
LINDORFF, D. P. (1965): *Theory of Sampled-Data Control Systems*. New York: John Wiley.

Aplicaciones prácticas del análisis de respuesta en frecuencia se estudian en
FLOWER, J. O., G. P. WINDETT, y S. C. FORGE (1971): «Aspects of the Frequency Response Testing of Simple Sampled Systems», *Int. Journal Control*, 14, 881-96.

Más información sobre la transformada z se encuentra en
JURY, E. I. (1964): *Theory and Applications of the Z-Transform Method*. New York: John Wiley.
La transformada z modificada se trata en
JURY, E. I. (1958): *Sample-Data control systems*. New York: John Wiley.

En este libro también se pueden encontrar tablas de transformadas z modificadas. Los sistemas con muestreo multiperíodico se analizaron inicialmente en
KRANC, G. M. (1957): «Input-Output Analysis of Multirate Feedback Systems», *IEEE Trans.*, AC-3, 21-28.

Resultados adicionales se presentan en
JURY, E. I. (1967): «A Note on Multirate Sampled-Data Systems», *IEEE Trans.*, AC-12, 319-20.
JURY, E. I. (1967): «A General z-Transform Formula for Sampled-Data Systems», *IEEE Trans.*, AC-12, 606-08.
KONAR, A. F., y J. K. MAHESH (1978): «Analysis Method for Multirate Digital Control Systems», *Honeywell Report No. F0636-TR1*, Honeywell Systems and Research Center, Minneapolis, Minn.
WHITBECK, R. F. (1980): «Multirate Digital Control Systems with Simulations Applications», *Report AFW AL-TR-80-3101*, vols. I, II y III, Flight Dynamics Laboratory, Air Force Wright Aeronautical Lab, Wright-Patterson Air Force Base, Ohio.

Análisis de sistemas discretos

Objetivo:

Introducir los conceptos de estabilidad, alcanzabilidad y observabilidad. Trabajar sobre el análisis de lazos de realimentación sencillos y su simulación.

5.1. Introducción

Los capítulos anteriores muestran cómo se transforman los sistemas continuos cuando se muestrean. En las Secciones 5.2 y 5.3 se exponen los conceptos de controlabilidad, alcanzabilidad y observabilidad, que son de gran utilidad para comprender los sistemas discretos. En la Sección 5.4 se tratan los lazos de control sencillos y sus propiedades.

La simulación es una herramienta muy importante para el análisis de los sistemas muestreados; por ejemplo, para investigar el comportamiento entre muestras.

En el Capítulo 3 se ha presentado el muestreo de los sistemas continuos. La matriz del sistema continuo A da la matriz del sistema discreto $\Phi = \exp(Ah)$. Este capítulo trata principalmente de los sistemas discretos y emplea el modelo en el espacio de estados

$$\begin{aligned} x(k+1) &= \Phi x(k) + Tu(k) \\ y(k) &= Cx(k) \end{aligned} \quad (5.1)$$

y el modelo entrada-salida

$$A(q)y(k) = B(q)u(k) \quad (5.2)$$

donde

$$\begin{aligned} A(q) &= q^n + a_1q^{n-1} + \dots + a_n \\ B(q) &= b_1q^{n-1} + \dots + b_n \end{aligned}$$

5.2. Estabilidad

Definiciones

Se supone que se conoce suficientemente la noción de estabilidad por los textos básicos de teoría de control. Aquí sólo se van a dar las definiciones fundamentales. Consideraremos la ecuación discreta en el espacio de estados (que puede ser no lineal y de coeficientes variables con el tiempo)

$$x(k+1) = f(x(k), k) \quad (5.3)$$

Sean $x^0(k)$ y $x(k)$ soluciones de (5.3) cuando las condiciones iniciales son, respectivamente, $x^0(k_0)$ y $x(k_0)$. Además $\|\cdot\|$ representa una norma.

Definición 5.1—Estabilidad. La solución $x^0(k)$ de (5.3) es *estable* si, para un $\varepsilon > 0$ dado, existe un $\delta(\varepsilon, k_0) > 0$ tal que todas las soluciones que cumplen $\|x(k_0) - x^0(k_0)\| < \delta$ se verifica que $\|x(k) - x^0(k)\| < \varepsilon$ para todo $k \geq k_0$.

Definición 5.2—Estabilidad asintótica. La solución $x^0(k)$ de (5.3) es *asintóticamente estable* si es estable y si $\|x(k) - x^0(k)\| \rightarrow 0$ cuando $k \rightarrow \infty$ siempre que $\|x(k_0) - x^0(k_0)\|$ sea suficientemente pequeño.

De las definiciones se sigue que la noción de estabilidad en general está definida para una *solución particular* y no para el sistema.

Estabilidad de los sistemas discretos lineales

Consideremos el sistema lineal

$$x^0(k+1) = \Phi x^0(k), \quad x^0(0) = a^0 \quad (5.4)$$

Para investigar la estabilidad de la solución de (5.4), perturbamos el valor inicial. Por tanto,

$$x(k+1) = \Phi x(k), \quad x(0) = a$$

La diferencia $\tilde{x} = x - x^0$ satisface la ecuación

$$\tilde{x}(k+1) = \Phi \tilde{x}(k), \quad \tilde{x}(0) = a - a^0 \quad (5.5)$$

Esto implica que si la solución x^0 es estable, entonces cualquier otra solución también es estable. Por tanto, para sistemas lineales invariantes, la estabilidad es una propiedad del sistema y no sólo de una solución en particular.

El sistema (5.5) tiene la solución

$$\tilde{x}(k) = \Phi^k \tilde{x}(0)$$

[Véase (3.13).] Si es posible diagonalizar Φ , entonces la solución es una combinación de términos λ_i^k , donde λ_i , $i = 1, \dots, n$, son los valores propios de Φ [véase (3.25)]. En el caso general, cuando Φ no puede diagonalizarse, la solución es una combinación lineal de los términos $p_i(k)\lambda_i^k$, donde $p_i(k)$ son polinomios de la variable k de orden inferior en una unidad a la multiplicidad del valor propio correspondiente. Para lograr la estabilidad asintótica, todas las soluciones deben tender a cero cuando k tiende a infinito. Entonces los valores propios de Φ tienen la propiedad

$$|\lambda_i| < 1, \quad i = 1, \dots, n$$

que queda formulada mediante el siguiente teorema.

Teorema 5.1 – Estabilidad asintótica de sistemas lineales. Un sistema discreto, lineal e invariante (5.4) es asintóticamente estable si y sólo si todos los valores propios de Φ están, de manera estricta, dentro del círculo unidad. \square

Ya se ha definido la estabilidad respecto a las perturbaciones en los valores iniciales. Pero también interesan otros tipos de estabilidad.

Definición 5.3 – Estabilidad BIBO. Un sistema lineal invariante se dice que es estable *entrada-acotada-salida-acotada* (BIBO, del inglés bounded-input-bounded-output) si una entrada acotada produce una salida acotada para cualquier valor inicial.

De la definición se deduce que la estabilidad asintótica es el concepto más fuerte. Una consecuencia de esto es el teorema siguiente.

Teorema 5.2. La estabilidad asintótica implica la estabilidad y la estabilidad entrada-acotada-salida-acotada. \square

Cuando en este texto se emplea la palabra *estable* sin ningún calificativo posterior, significa normalmente estabilidad asintótica.

Es fácil presentar ejemplos que muestran que la estabilidad no implica la estabilidad BIBO, y viceversa.

Ejemplo 5.1

Consideremos el oscilador armónico muestreado (véase el Ejemplo A.3)

$$\begin{aligned} x(kh + h) &= \begin{bmatrix} \cos \omega h & \operatorname{sen} \omega h \\ -\operatorname{sen} \omega h & \cos \omega h \end{bmatrix} x(kh) + \begin{bmatrix} 1 - \cos \omega h \\ \operatorname{sen} \omega h \end{bmatrix} u(kh) \\ y(kh) &= [1 \ 0] x(kh) \end{aligned}$$

El módulo de los valores propios es la unidad. El sistema es estable, ya que $\|x(kh)\| = \|x(0)\|$ si $u(kh)=0$. Tomemos como entrada una onda cuadrada de frecuencia ω rad/s. La Figura 5.1 presenta la entrada y la salida del sistema. La amplitud de salida crece indefinidamente y, por tanto, el sistema no es estable BIBO.

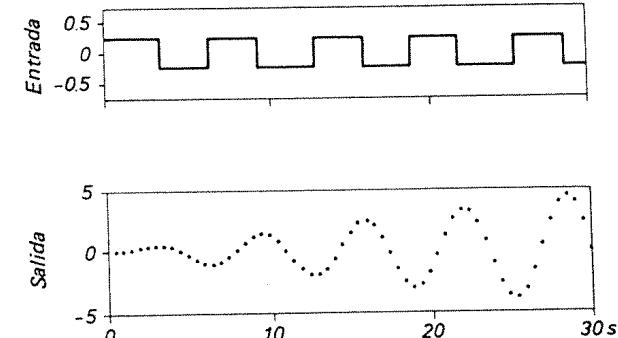


Figura 5.1 Entrada y salida del sistema del Ejemplo 5.1 cuando $\omega=1$, $h=0.5$ y el estado inicial es cero.

Pruebas de estabilidad

Estas son algunas de las formas de determinar la estabilidad de un sistema discreto:

Cálculo directo de los valores propios de Φ .

Métodos basados en las propiedades de los polinomios característicos.

Método del lugar de las raíces.

Criterio de Nyquist.

Método de Lyapunov.

Del Teorema 5.1 se deduce que una forma directa de comprobar la estabilidad de un sistema dado es calcular los valores propios de la matriz Φ . Existen buenos algoritmos numéricos para hacerlo. Se disponen de métodos muy probados, por ejemplo, en el paquete EISPACK, a los que se puede acceder fácilmente en la mayor parte de los centros de cálculo.

No es conveniente realizar de forma manual el cálculo directo de los valores propios para sistemas de orden superior a dos, y tampoco se puede emplear el método de los valores propios cuando la matriz tiene parámetros en sus coeficientes.

En algunos casos es más fácil calcular el polinomio característico

$$A(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_n \quad (5.6)$$

e investigar la ecuación característica

$$A(z) = a_0 z^n + a_1 z^{n-1} + \dots + a_n = 0 \quad (5.7)$$

Conviene recordar de la Sección 3.4 que el polinomio característico es el polinomio denominador de la función de transferencia discreta. Las pruebas de



estabilidad pueden obtenerse investigando las condiciones para que los ceros del polinomio estén dentro del círculo unidad.

El criterio de Routh-Hurwitz es uno de los métodos existentes para determinar si un polinomio tiene todos sus ceros en el semiplano izquierdo. La transformación de Möbius

$$w = \frac{z+1}{z-1}$$

aplica el círculo unidad del plano z en el semiplano izquierdo del plano w . Se puede aplicar la transformación a (5.7) y entonces utilizar el criterio de Routh-Hurwitz.

Sin embargo, es útil disponer de condiciones que digan directamente si un polinomio tiene todos sus ceros dentro del círculo unidad. Schur, Cohn y Jury desarrollaron un criterio de este tipo, equivalente al criterio de Routh-Hurwitz. El siguiente apartado está dedicado a describirlo en detalle. El cálculo de los coeficientes del polinomio característico a partir de los elementos de la matriz es una solución de mala naturaleza. Si se da la matriz, es preferible calcular los valores propios directamente en lugar de calcular la ecuación característica.

También puede emplearse el conocido método del lugar de las raíces para los sistemas continuos. Por otra parte las reglas empíricas para dibujar el lugar de las raíces son las mismas. El método del lugar de las raíces y el criterio de Nyquist se emplean para determinar la estabilidad de los sistemas en lazo cerrado cuando se conoce el sistema en lazo abierto.

Criterio de Jury

Para determinar si un polinomio como (5.6) tiene todos sus ceros dentro del círculo unidad se puede emplear la siguiente prueba.

Se forma la tabla

a_0	a_1	\dots	a_{n-1}	a_n	
a_n	a_{n-1}	\dots	a_1	a_0	$\alpha_n = \frac{a_n}{a_0}$
a_0^{n-1}	a_1^{n-1}	\dots	a_{n-1}^{n-1}		
a_{n-1}^{n-1}	a_{n-2}^{n-1}	\dots	a_0^{n-1}		$\alpha_{n-1} = \frac{a_{n-1}^{n-1}}{a_0^{n-1}}$
\vdots	\vdots				
a_0^n					

donde

$$a_i^{k-1} = a_i^k - \alpha_k a_{k-i}^k$$

$$\alpha_k = a_k^k / a_0^k$$

La primera y segunda filas son los coeficientes de (5.6) en orden directo e inverso, respectivamente. La tercera fila se obtiene multiplicando la segunda por $\alpha_n = a_n/a_0$ y restando esto de la primera fila. Entonces el último elemento de la tercera fila será cero. La cuarta fila es la tercera en orden inverso. Se repite este esquema hasta que haya $2n+1$ filas. La última fila sólo tiene un elemento. De aquí resulta el siguiente teorema.

Teorema 5.3 – Prueba de estabilidad de Jury. Si $a_0 > 0$, la ecuación (5.7) tiene todas sus raíces dentro del círculo unidad si y sólo si todos los a_0^k , $k=0, 1, \dots, n-1$ son positivos. Si ningún a_0^k es nulo, entonces el número de a_0^k negativos es igual al número de raíces fuera del círculo unidad. \square

Nota. Si todos los a_0^k son positivos para $k=1, 2, \dots, n-1$ entonces puede demostrarse que la condición $a_0^0 > 0$ es equivalente a las condiciones

$$A(1) > 0 \\ (-1)^n A(-1) > 0$$

Estas condiciones constituyen las condiciones necesarias para la estabilidad y, por tanto, pueden utilizarse antes de formar la tabla.

Ejemplo 5.2

Sea la ecuación característica

$$A(z) = z^2 + a_1 z + a_2 = 0 \quad (5.8)$$

El esquema de Jury es

1 a_2	a_1 a_1	a_2 1 $\alpha_2 = a_2$
$1 - a_2^2$ $a_1(1 - a_2)$	$a_1(1 - a_2)$ $1 - a_2^2$	$\alpha_1 = \frac{a_1}{1 + a_2}$
$1 - a_2^2 - \frac{a_1^2(1 - a_2)}{1 + a_2}$		

Todas las raíces de la ecuación (5.8) están dentro del círculo unidad si

$$1 - a_2^2 > 0 \\ \frac{1 - a_2}{1 + a_2} [(1 + a_2)^2 - a_1^2] > 0$$

Esto da las condiciones

$$a_2 < 1 \\ a_2 > -1 + a_1 \\ a_2 > -1 - a_1$$

En la Figura 5.2 se da el área de estabilidad para la ecuación de segundo orden.

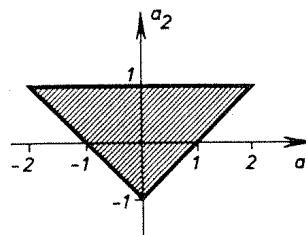


Figura 5.2 El área de estabilidad de la ecuación de segundo orden (5.8) como una función de los coeficientes a_1 y a_2 .

El proceso con retardo de los Ejemplos A.4 y 3.6 tiene la ecuación característica

$$z^2 - z \exp(-h) = 0$$

cuando el período de muestreo es h . El sistema es estable, ya que $\exp(-h) < 1$. El modelo de inventario del Ejemplo A.5 tiene la ecuación característica

$$z^2 - z = 0$$

Este sistema es estable pero no asintóticamente estable, porque la ecuación característica tiene la raíz $z=1$.

El criterio de Nyquist

El criterio de Nyquist es una prueba de estabilidad muy conocida para los sistemas continuos. Está basado en el principio del argumento. La prueba puede formularse fácilmente para tratar los sistemas discretos. El criterio de Nyquist es especialmente útil para determinar la estabilidad del sistema en lazo cerrado cuando se da el sistema en lazo abierto.

Considérese el sistema discreto de la Figura 5.3. El sistema en lazo cerrado tiene como función de transferencia discreta

$$H(z) = \frac{Y(z)}{U_1(z)} = \frac{H_0(z)}{1 + H_0(z)}$$

La ecuación característica del sistema en lazo cerrado es

$$1 + H_a(z) \equiv 0 \quad (5.9)$$

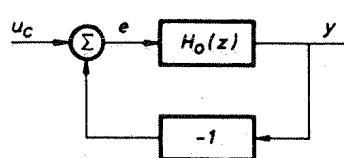


Figura 5.3 Sistema sencillo con realimentación unidad.

La estabilidad del sistema en lazo cerrado puede estudiarse a partir del trazo de Nyquist de $H_0(z)$. Para los sistemas discretos, el área de estabilidad es el círculo unidad en lugar del semiplano izquierdo. La Figura 5.4 presenta el trazo Γ_c rodeando el área exterior al círculo unidad. La pequeña desviación en $z = 1$ es para excluir a los integradores del sistema en lazo abierto. El semicírculo infinitesimal en $z = 1$ con argumentos decrecientes desde $\pi/2$ hasta $-\pi/2$ se aplica sobre el plano $H_0(z)$ como un círculo de radio infinito desde $-n\pi/2$ hasta $n\pi/2$, donde n es el número de integradores en el sistema en lazo abierto. Si hay polos sobre la circunferencia unidad distintos del $z = 1$, es preciso excluirlos mediante pequeños semicírculos de la misma forma que se ha hecho con $z = 1$. La aplicación de la circunferencia unidad es $H_0(e^{j\omega})$ con $\omega \in (0, 2\pi)$.

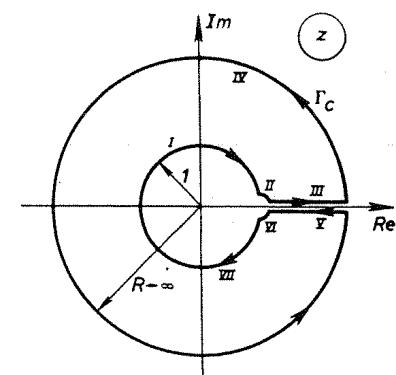


Figura 5.4 El trazo Γ_c rodeando el área fuera del círculo unidad.

Ahora se puede determinar la estabilidad del sistema en lazo cerrado estudiando la imagen de la curva Γ_c por la aplicación $H_0(z)$. La curva $H_0(e^{j\omega})$ con argumentos desde 0 a π recibe el nombre de *curva de respuesta en frecuencia*, o *curva de Nyquist* del sistema. El principio del argumento establece que el número de vueltas N en sentido positivo alrededor del punto $(-1, 0)$ que realiza la curva en que se transforma Γ_c por la aplicación es igual a

$$N = Z - P$$

donde Z y P son el número de ceros y polos, respectivamente, de $1 + H_0(z)$ fuera del círculo unidad. Hay que notar que si el sistema en lazo abierto es estable, entonces $P=0$ y, por tanto, $N=Z$. Entonces la estabilidad del sistema en lazo cerrado está asegurada si la curva de la aplicación de Γ_c no rodea al punto $(-1, 0)$. Si $H(z) \rightarrow 0$ cuando $z \rightarrow \infty$ las líneas paralelas III y V no influyen en la prueba de la estabilidad, y es suficiente encontrar la aplicación del círculo unidad y del pequeño semicírculo para $z=1$. El criterio de Nyquist puede simplificarse todavía más si el sistema en lazo abierto y su inverso son estables. Entonces, la estabilidad del sistema en lazo cerrado está asegurada si el punto $(-1, 0)$ del plano $H_0(z)$ está a la izquierda de la curva

aplicación de $H_0(e^{i\omega})$ para ω desde 0 a π ; es decir, a la izquierda de la curva de Nyquist.

Ejemplo 5.3

Consideremos un sistema con período de muestreo $h=1$ y la función de transferencia discreta

$$H_0(z) = \frac{0.25K}{(z-1)(z-0.5)}$$

Entonces

$$H_0(e^{i\omega}) = \frac{0.25K[1.5(1 - \cos \omega) - 2 \sin^2 \omega - i \sin \omega(2 \cos \omega - 1.5)]}{(2 - 2 \cos \omega)(1.25 + \cos \omega)}$$

En la Figura 5.5 se presenta la curva aplicación Γ_c . La línea completa es la curva de Nyquist, es decir, la curva aplicación de $H_0(e^{i\omega})$ para ω desde 0 a π . En la figura se puede ver que la curva de Nyquist cruza el eje real negativo en -0.5 . Entonces el sistema en lazo cerrado es estable si $K < 2$.

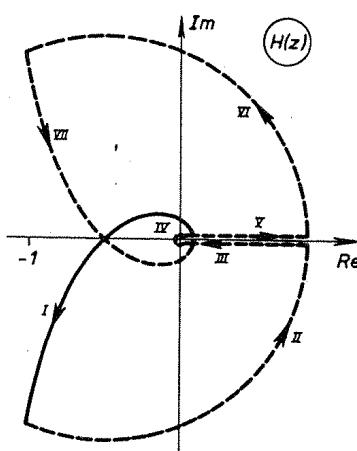


Figura 5.5 La aplicación de Γ_c sobre el plano $H_0(z)$ del sistema del Ejemplo 5.3, cuando $K=1$. La línea completa es la curva de Nyquist.

Hasta ahora el criterio de Nyquist sólo se ha considerado para los sistemas discretos en general. Ahora supondremos que la función de transferencia discreta se obtiene muestreando un sistema continuo $G(s)$, precedida por un mantenedor de orden cero con un período de muestreo h . Un punto de vista alternativo se obtiene a partir de (4.14) y (4.17), ya que

$$H_0(e^{i\omega}) = \frac{1}{h} \sum_{k=-\infty}^{\infty} F(i\omega + ik\omega_s)$$

donde

$$F(s) = \frac{1 - \exp(-hs)}{s} G(s)$$

Si se dispone de los valores numéricos de $G(i\omega)$, entonces se puede determinar los valores de la función de transferencia discreta empleando el desarrollo en serie indicado. En la mayor parte de las aplicaciones, $|G(i\omega)|$ disminuye al aumentar ω . Entonces dominará el término para $k=0$. Esto es especialmente cierto si se incluye en $G(s)$ un filtro antienmascaramiento efectivo. Esto implica que es suficiente con conocer bien $G(s)$ por encima de la frecuencia de Nyquist.

El inconveniente del criterio de Nyquist es la dificultad para dibujar la curva de Nyquist, pero se pueden emplear programas de computador para dibujar la curva y facilitar la evaluación de la estabilidad de los sistemas en lazo cerrado.

Robustez

Es muy interesante ver la sensibilidad de un sistema a las perturbaciones, que pueden introducirse por las tolerancias de los componentes. Debido a que los diseños de los sistemas de control se basan en modelos simplificados, también es interesante conocer qué grado de exactitud debe tener el modelo para que el diseño sea correcto. El teorema de Nyquist puede permitir hacerse una idea correcta de estos problemas.

Consideremos el sistema sencillo en lazo cerrado de la Figura 5.3. Sea $H^0(z)$ la verdadera función de transferencia discreta en lazo abierto y $H(z)$ el valor nominal de H^0 . De lo que se trata es de estudiar qué proximidad de H a H^0 se necesita para hacer que el sistema en lazo cerrado sea estable. La función de transferencia discreta del sistema en lazo cerrado es

$$H_c(z) = \frac{H^0(z)}{1 + H^0(z)} \quad (5.10)$$

Los polos del sistema en lazo cerrado son entonces los ceros de la función

$$f(z) = 1 + H^0(z) = 1 + H(z) + H^0(z) - H(z)$$

Si

$$|H^0(z) - H(z)| < |1 + H(z)| \quad (5.11)$$

respecto al círculo unidad, entonces del principio de variación del argumento se deduce que las diferencias entre el número de polos y el de ceros fuera del círculo unidad para las funciones $1 + H$ y $1 + H^0$ son las mismas. De aquí resulta el siguiente teorema:

Teorema 5.4. Considerense los sistemas en lazo cerrado S y S^0 obtenidos aplicando una realimentación unidad negativa a los sistemas con funciones de

transferencia discreta H y H^0 , respectivamente. El sistema S^0 es estable si se verifican las siguientes condiciones:

1. S es estable.
2. H y H^0 tienen el mismo número de polos fuera del círculo unidad.
3. La desigualdad (5.11) se verifica para $|z|=1$. \square

El resultado muestra que es importante conocer el número de modos inestables para diseñar un regulador para el sistema. La desigualdad también da el rango de frecuencia en el que es importante tener una buena descripción del proceso. Nótese en particular que los requisitos de precisión son muy pequeños para las frecuencias en las que la ganancia del lazo es grande. Se requiere una buena precisión para las frecuencias en las que $H^0(z) \approx -1$.

A continuación se presenta un resultado íntimamente relacionado con el anterior que proporciona una mayor comprensión del problema. La función de transferencia discreta del sistema en lazo cerrado dada en (5.10) puede escribirse también como

$$H_c = \frac{1}{1 + 1/H^0}$$

Los polos del sistema en lazo cerrado vienen dados entonces por los ceros de la función

$$f_c(z) = 1 + \frac{1}{H^0(z)} = 1 + \frac{1}{H(z)} + \left[\frac{1}{H^0(z)} - \frac{1}{H(z)} \right]$$

Del principio de variación del argumento se deduce que las diferencias entre los ceros y polos fuera del círculo unidad de las funciones $1+1/H^0$ y $1+1/H$ son las mismas si

$$\left| \frac{1}{H^0(z)} - \frac{1}{H(z)} \right| < \left| 1 + \frac{1}{H(z)} \right| \quad (5.12)$$

respecto al círculo unidad. Se obtiene entonces el siguiente resultado.

Teorema 5.5. Considerense los sistemas en lazo cerrado S y S^0 obtenidos aplicando una realimentación negativa unitaria a los sistemas con funciones de transferencia discreta H y H^0 , respectivamente. El sistema S^0 es estable si se cumplen las siguientes condiciones:

1. S es estable.
2. H y H^0 tienen el mismo número de ceros fuera del círculo unidad.
3. Se verifica la desigualdad (5.12) para $|z|=1$. \square

El teorema indica la importancia de conocer el número de ceros fuera del círculo unidad. También muestra que la estabilidad se puede mantener a pesar de que existan grandes diferencias entre H y H^0 siempre que la ganancia del lazo sea grande.

De las conclusiones de los Teoremas 5.4 y 5.5, se obtienen las siguientes reglas para el diseño de sistemas realimentados a partir de modelos aproximados o con imprecisiones.

Es importante conocer el número de polos y ceros inestables.

No es importante conocer con precisión el modelo para aquellas frecuencias en las que se puede hacer que la ganancia del lazo sea grande.

Es necesario hacer pequeña la ganancia del lazo para aquellas frecuencias en las que el modelo es impreciso.

Es necesario tener un modelo que describa el sistema con precisión para aquellas frecuencias en las que $H^0(z) \approx -1$.

Segundo método de Lyapunov

El *segundo método de Lyapunov* es una herramienta muy útil para determinar la estabilidad de los sistemas dinámicos no lineales. Lyapunov desarrolló la teoría para ecuaciones diferenciales, pero se puede obtener también una teoría equivalente para ecuaciones en diferencias.

Definición 5.4 – Función de Lyapunov. $V(x)$ es una *función de Lyapunov* para el sistema

$$x(k+1) = f(x(k)), \quad f(0) = 0 \quad (5.13)$$

si:

1. $V(x)$ es continua en x y $V(0)=0$.
2. $V(x)$ es definida positiva.
3. $\Delta V(x) = V(f(x)) - V(x)$ es definida negativa.

Una ilustración geométrica sencilla de la definición puede verse en la Figura 5.6. Las curvas de nivel de una función continua definida positiva V son curvas cerradas en un entorno del origen. Identifiquemos cada curva por el valor de la función. La condición 3 implica que la dinámica del sistema es tal que la solución siempre se mueve hacia las curvas con valores más bajos. Todas las curvas de nivel rodean al

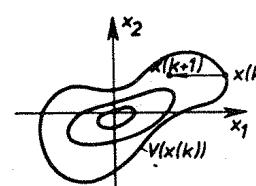


Figura 5.6 Ilustración geométrica del teorema de Lyapunov.



origen y no se cortan entre sí. Entonces parece razonable que la existencia de una función de Lyapunov asegure la estabilidad asintótica. El siguiente teorema formula de forma precisa esta afirmación.

Teorema 5.6 – Teorema de estabilidad de Lyapunov. La solución $x(k)=0$ es asintóticamente estable si existe una función de Lyapunov para el sistema (5.13). Además, si

$$0 < \varphi(\|x\|) < V(x)$$

donde $\varphi(\|x\|) \rightarrow \infty$ cuando $\|x\| \rightarrow \infty$, entonces la solución es asintóticamente estable para todas las condiciones iniciales. \square

El principal obstáculo para el empleo de la teoría de Lyapunov es encontrar una función de Lyapunov apropiada. En general se trata de un problema difícil; sin embargo, para el sistema lineal dado en (5.4), la determinación de funciones cuadráticas de Lyapunov es inmediato. Tomemos $V(x)=x^T Px$ como una posible función de Lyapunov. El incremento de V viene dado entonces por

$$\begin{aligned}\Delta V(x) &= V(\Phi x) - V(x) = x^T \Phi^T P \Phi x - x^T P x \\ &= x^T [\Phi^T P \Phi - P] x = -x^T Q x\end{aligned}$$

Para que V sea una función de Lyapunov es necesario y suficiente que exista una matriz definida positiva P que verifique la ecuación

$$\Phi^T P \Phi - P = -Q \quad (5.14)$$

donde Q es definida positiva. La ecuación (5.14) se conoce como *ecuación de Lyapunov*. Se demuestra que siempre hay una solución de la ecuación de Lyapunov cuando el sistema lineal es estable. La matriz P es definida positiva si Q lo es.

5.3. Controlabilidad, alcanzabilidad y observabilidad

En esta sección se discutirán dos cuestiones fundamentales para los sistemas dinámicos. La primera es si es posible o no llevar un sistema desde un estado inicial dado a cualquier otro estado. La segunda es cómo determinar el estado de un sistema dinámico a partir de las observaciones de las entradas y salidas. Estas cuestiones fueron propuestas y resueltas por Kalman (1960), quien introdujo también los conceptos de controlabilidad y observabilidad. Los sistemas pueden ser de varias entradas y varias salidas.

Controlabilidad y alcanzabilidad

Consideremos el sistema dado en (5.1). Supongamos que tenemos un estado inicial $x(0)$ dado. El estado en el instante n , donde n es el orden del sistema, viene dado por la expresión

$$\begin{aligned}x(n) &= \Phi^n x(0) + \Phi^{n-1} \Gamma u(0) + \dots + \Gamma u(n-1) \\ &= \Phi^n x(0) + W_c U\end{aligned} \quad (5.15)$$

donde

$$\begin{aligned}W_c &= [\Gamma \quad \Phi\Gamma \quad \dots \quad \Phi^{n-1}\Gamma] \\ U &= [u^T(n-1) \quad \dots \quad u^T(0)]^T\end{aligned}$$

[Compárese con la ecuación (3.13).] Si W_c tiene rango n , entonces es posible encontrar n ecuaciones a partir de las cuales pueden calcularse las señales de control tales que transfieren el estado inicial al estado final deseado $x(n)$. Nótese que la solución no es única si hay más de una señal de entrada. En la literatura, la controlabilidad se define de diferentes formas; en este texto se usará la definición siguiente:

Definición 5.5 – Controlabilidad. El sistema (5.1) es controlable si es posible encontrar una secuencia de control tal que puede alcanzarse el origen desde cualquier estado inicial en un tiempo finito.

Un concepto relacionado con el de controlabilidad es el de *alcanzabilidad*, que se define de la siguiente forma.

Definición 5.6 – Alcanzabilidad. El sistema (5.1) es alcanzable si es posible encontrar una secuencia de control tal que puede alcanzarse un estado arbitrario desde cualquier estado inicial en un tiempo finito.

La controlabilidad no implica la alcanzabilidad, como puede verse fácilmente en (5.15). Si $\Phi^n x(0)=0$, entonces el origen se alcanzará con entrada cero pero el sistema no es necesariamente alcanzable. Sin embargo, los dos conceptos son equivalentes si Φ es invertible. Aquí trataremos, principalmente, la alcanzabilidad.

De la anterior definición y realizando algunos cálculos se obtiene el siguiente teorema.

Teorema 5.7. El sistema (5.1) es alcanzable si y sólo si la matriz W_c tiene rango n . \square

Nota. A la matriz W_c se la suele llamar matriz de controlabilidad por su analogía con los sistemas continuos.

Ejemplo 5.4

El sistema

$$x(k+1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

no es alcanzable porque

$$W_c = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Si en lugar de ser así el sistema tuviera dos entradas con una matriz Γ_c no singular, entonces el sistema sería alcanzable.

Mediante el teorema de Cayley-Hamilton se deduce de (5.15) que todos los estados que pueden alcanzarse desde el origen están cubiertos por las columnas de la matriz de controlabilidad W_c . Esto implica que todos los estados alcanzables pertenecen al subespacio lineal cubierto por las columnas de W_c .

Ejemplo 5.5

Dado el sistema

$$x(k+1) = \begin{bmatrix} 1 & 1 \\ -0.25 & 0 \end{bmatrix}x(k) + \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}u(k); \quad x(0) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

¿es posible encontrar una secuencia de control tal que $x^T(2) = [-0.5 \ 1]^T$? De (5.15),

$$x(2) = \Phi^2 x(0) + \Phi \Gamma u(0) + \Gamma u(1)$$

o

$$\begin{bmatrix} -0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.5 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}[0.5u(0) + u(1)]$$

que da la condición

$$0.5u(0) + u(1) = -4$$

Una posible secuencia de control es $u(0) = -2$ y $u(1) = -3$. Supongamos en su lugar que $x^T(2) = [0.5 \ 1]^T$. Esto da el sistema de ecuaciones

$$\begin{bmatrix} -3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}[0.5u(0) + u(1)]$$

que no tiene solución. La razón es que el sistema no es alcanzable. La matriz de controlabilidad es

$$W_c = \begin{bmatrix} 1 & 0.5 \\ -0.5 & -0.25 \end{bmatrix}$$

Partiendo del origen es posible alcanzar sólo aquellos puntos del espacio de estados que pertenecen al subespacio cubierto por el vector $[1 \ -0.5]^T$. En el ejemplo se pueden alcanzar otros puntos debido al efecto del valor inicial.

Supongamos que se introducen unas nuevas coordenadas por una matriz de transformación T no singular (compárese con la Sección 3.3). En las nuevas coordenadas

$$\begin{aligned} \tilde{W}_c &= [\tilde{\Gamma} \ \tilde{\Phi}\tilde{\Gamma} \ \dots \ \tilde{\Phi}^{n-1}\tilde{\Gamma}] \\ &= [T\Gamma \ T\Phi T^{-1}T\Gamma \ \dots \ T\Phi^{n-1}T^{-1}T\Gamma] \\ &= TW_c \end{aligned} \quad (5.16)$$

Si W_c tiene rango n , entonces \tilde{W}_c también tiene rango n . Esto significa que la alcanzabilidad de un sistema es independiente de las coordenadas.

Nótese que la fórmula que relaciona \tilde{W}_c con W_c es útil para calcular la matriz de transformación de una a otra.

La forma controlable, introducida en la Sección 3.3, puede obtenerse si la matriz de (3.29) tiene inversa, lo que explica el nombre de *forma controlable canónica*.

Ejemplo 5.6

Consideremos el sistema de tercer orden

$$x(k+1) = \begin{bmatrix} -a_1 & -a_2 & -a_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}x + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}u$$

que está en forma controlable. La matriz de controlabilidad es

$$W_c = [\Gamma \ \Phi\Gamma \ \Phi^2\Gamma] = \begin{bmatrix} 1 & -a_1 & a_1^2 - a_2 \\ 0 & 1 & -a_1 \\ 0 & 0 & 1 \end{bmatrix}$$

La inversa viene dada por

$$W_c^{-1} = \begin{bmatrix} 1 & a_1 & a_2 \\ 0 & 1 & a_1 \\ 0 & 0 & 1 \end{bmatrix}$$

El ejemplo se puede generalizar al caso de orden n , donde

$$W_c^{-1} = \begin{bmatrix} 1 & a_1 & a_2 & \dots & a_{n-2} & a_{n-1} \\ 0 & 1 & a_1 & \dots & a_{n-3} & a_{n-2} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & a_1 & 0 \\ 0 & 0 & \dots & 0 & 1 & \end{bmatrix}$$

Seguimiento de trayectoria

Mediante las definiciones y cálculos anteriores es posible determinar una secuencia de control tal que se puede alcanzar el estado que se desee después de, como máximo, n períodos de tiempo. ¿La alcanzabilidad también implica que es posible seguir una trayectoria dada en el espacio de estados? Supongamos que se da un $x(k)$ cualquiera y que es preciso llegar a $x(k+1)$. De (5.15) puede verse que esto es posible si Γ tiene rango n ; es decir, tener n señales de entrada es una condición necesaria pero no suficiente. Para un sistema de una entrada y una salida, en general, es posible alcanzar los estados deseados sólo cada n -ésimo punto de muestreo, suponiendo que los puntos deseados se conocen con n períodos de antelación.



Es más fácil hacer que la salida siga una trayectoria determinada. Supongamos que la trayectoria viene dada por $u_c(k)$. La señal de control u debe cumplir entonces

$$\begin{aligned} y(k) &= \frac{B(q)}{A(q)} u(k) = u_c(k) \\ \text{o} \\ u(k) &= \frac{A(q)}{B(q)} u_c(k) \end{aligned} \quad (5.17)$$

Supongamos que hay d períodos de retardo en el sistema. La generación de $u(k)$ sólo será causal si se conoce la trayectoria deseada con d períodos de antelación. La señal de control puede entonces generarse en tiempo real. La señal de control se obtiene entonces enviando la trayectoria de salida deseada a través del sistema inverso A/B . La ecuación (5.17) tiene una solución única si la señal $u_c(k)$ es tal que existe un k_0 tal que $u(k)=0$ para todo $k < k_0$ (compárese con la Sección 3.4). La señal u está acotada si u_c está acotada y si el sistema tiene un inverso estable.

Observabilidad

El concepto de estados inobservables se introduce para resolver el problema de encontrar el estado de un sistema a partir de las observaciones de la salida.

Definición 5.7. El estado $x^0 \neq 0$ es inobservable si existe un número finito $k_1 \geq n-1$ tal que $y(k)=0$ para $0 \leq k \leq k_1$ cuando $x(0)=x^0$ y $u(k)=0$ para $0 \leq k \leq k_1$.

El sistema dado en (5.1) es *observable* si hay un k finito tal que basta con conocer las entradas $u(0), \dots, u(k-1)$ y las salidas $y(0), \dots, y(k-1)$ para determinar el estado inicial del sistema.

Consideremos el sistema dado en (5.1). Siempre es posible determinar el efecto de la señal de entrada conocida, y por tanto no hay pérdida de generalidad al suponer que $u(k)=0$. Supongamos que se dan $y(0), y(1), \dots, y(n-1)$. Esto conduce al siguiente sistema de ecuaciones:

$$y(0) = Cx(0)$$

$$y(1) = Cx(1) = C\Phi x(0)$$

⋮

$$y(n-1) = C\Phi^{n-1}x(0)$$

Empleando notación vectorial tenemos

$$\begin{bmatrix} C \\ C\Phi \\ \vdots \\ C\Phi^{n-1} \end{bmatrix} x(0) = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(n-1) \end{bmatrix} \quad (5.18)$$

El estado $x(0)$ puede obtenerse a partir de (5.18) si y sólo si la matriz de observabilidad

$$W_o = \begin{bmatrix} C \\ C\Phi \\ \vdots \\ C\Phi^{n-1} \end{bmatrix} \quad (5.19)$$

tiene rango n . El estado $x(0)$ es inobservable si $x(0)$ está en el espacio nulo de W_o . Si dos estados son inobservables, también lo es cualquier combinación lineal de ellos; esto es, los estados inobservables forman un subespacio lineal.

Teorema 5.8. El sistema (5.1) es observable si y sólo si W_o tiene rango n . \square

La prueba de observabilidad dada en el Teorema 5.8 es equivalente a la de los sistemas continuos. La demostración de que la matriz de observabilidad es independiente del sistema de coordenadas de la misma forma que lo es la matriz de controlabilidad, es inmediata.

Ejemplo 5.7

Consideremos el sistema

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 1.1 & -0.3 \\ 1 & 0 \end{bmatrix} x(k) \\ y(k) &= [1 \ -0.5]x(k) \end{aligned}$$

La matriz de observabilidad es

$$W_o = \begin{bmatrix} C \\ C\Phi \end{bmatrix} = \begin{bmatrix} 1 & -0.5 \\ 0.6 & -0.3 \end{bmatrix}$$

El rango de W_o es 1, y los estados inobservables pertenecen al espacio nulo de W_o ; es decir, a $[0.5, \ 1]$. La Figura 5.7 muestra la salida para cuatro estados iniciales diferentes. Todos los estados iniciales que están sobre una línea paralela a $[0.5, \ 1]$ dan la misma salida [véanse las Figs. 5.7b) y d)].

Descomposición de Kalman

La parte alcanzable y la inobservable de un sistema son dos subespacios lineales del espacio de estados. Ambos subespacios son independientes de las coordenadas en el espacio de estados. Kalman demostró que se pueden emplear unas coordenadas tales que permiten descomponer un sistema de la siguiente forma:

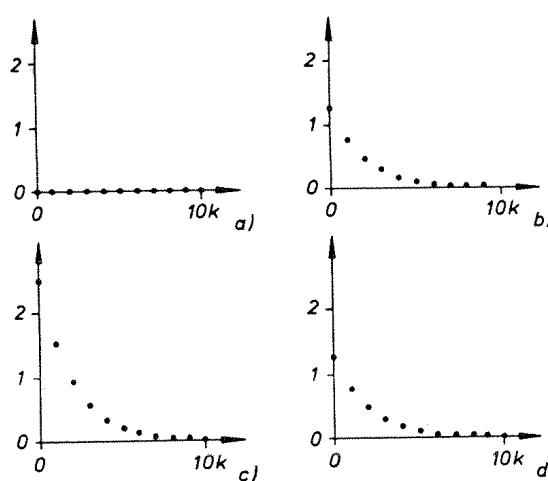


Figura 5.7 Salida del sistema del Ejemplo 5.7 para los estados iniciales
a) $[0.5, 1]$, b) $[1.5, 0.5]$, c) $[2.5, 0]$ y
d) $[1, -0.5]$.

$$\begin{aligned}x(k+1) &= \begin{bmatrix}\Phi_{11} & \Phi_{12} & 0 & 0 \\ 0 & \Phi_{22} & 0 & 0 \\ \Phi_{31} & \Phi_{32} & \Phi_{33} & \Phi_{34} \\ 0 & \Phi_{42} & 0 & \Phi_{44}\end{bmatrix}x(k) + \begin{bmatrix}\Gamma_1 \\ 0 \\ \Gamma_3 \\ 0\end{bmatrix}u(k) \\y(k) &= [C_1 \ C_2 \ 0 \ 0]x(k)\end{aligned}$$

donde Φ_{ij} , Γ_i y C_i son matrices de los órdenes adecuados. El espacio de estados se descompone en cuatro partes, que corresponden a los estados que son alcanzables y observables, no alcanzables pero observables, alcanzables pero no observables y ni alcanzables ni observables.

Mediante sencillas manipulaciones algebraicas, el operador de transferencia discreta resulta ser

$$H(q) = C_1(qI - \Phi_{11})^{-1}\Gamma_1$$

El operador de transferencia discreta queda entonces determinado por la parte alcanzable y observable del sistema. El teorema siguiente resume estos resultados.

Teorema 5.9 – Descomposición de Kalman. Un sistema lineal puede descomponerse en cuatro subsistemas con las siguientes propiedades:

- S_{or} Subsistema observable y alcanzable.
- S_{of} Subsistema observable pero no alcanzable.
- $S_{\bar{o}r}$ Subsistema no observable pero alcanzable.
- $S_{\bar{o}f}$ Subsistema no observable y no alcanzable.

Además, la función de transferencia discreta del sistema queda determinada de forma única por el subsistema que es a la vez observable y alcanzable. \square

Un diagrama del bloques de esta descomposición puede verse en la Figura 5.8, que muestra cómo están interconectados los subsistemas. La figura también muestra que la relación entrada-salida viene dada solamente por el subsistema S_{or} .

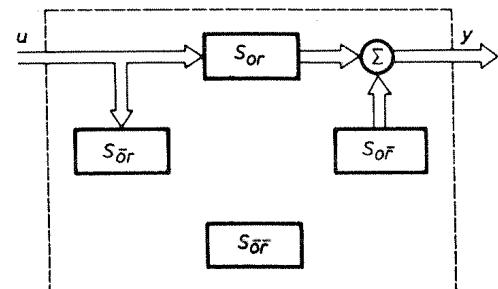


Figura 5.8 Diagrama de bloques de la descomposición de Kalman.

Pérdida de la alcanzabilidad y la observabilidad por el muestreo

El muestreo de un sistema continuo produce un sistema discreto con matrices del sistema que dependen del período de muestreo. ¿Cómo influirá éste sobre la observabilidad y alcanzabilidad del sistema muestreado? Para tener un sistema discreto alcanzable, es necesario que el sistema continuo sea también alcanzable, ya que las señales de control permitidas para el sistema muestreado —señales constantes a tramos— son un subconjunto de las señales de control permitidas para el sistema continuo. Sin embargo, puede ocurrir que se pierda la alcanzabilidad para algunos períodos de muestreo.

Las condiciones para la inobservabilidad están más restringidas en los sistemas discretos debido a que la salida debe ser cero durante un intervalo de tiempo, mientras que la salida discreta sólo tiene que ser cero en los instantes de muestreo. Esto significa que la salida continua puede oscilar entre los instantes de muestreo y ser nula en ellos. A esta situación se la llama a veces *oscilación oculta*. El sistema muestreado puede entonces ser inobservable incluso aunque el correspondiente sistema continuo sea observable.

Se puede emplear el oscilador armónico para ilustrar la exposición anterior.

Ejemplo 5.8

El modelo discreto del oscilador armónico viene dado por (véase el Ejemplo A.3)

$$\begin{aligned}x(kh+h) &= \begin{bmatrix}\cos \omega h & \operatorname{sen} \omega h \\ -\operatorname{sen} \omega h & \cos \omega h\end{bmatrix}x(kh) + \begin{bmatrix}1 - \cos \omega h \\ \operatorname{sen} \omega h\end{bmatrix}u(kh) \\y(kh) &= [1 \ 0]x(kh)\end{aligned}$$



Los determinantes de las matrices de controlabilidad y observabilidad son

$$\det W_c = -2 \sin \omega h (1 - \cos \omega h)$$

y

$$\det W_o = \sin \omega h$$

Tanto la alcanzabilidad como la observabilidad se pierden para $\omega h = n\pi$, aunque el sistema continuo correspondiente, dado por (A.7), es controlable y observable.

5.4. Análisis de lazos de realimentación sencillos

En esta sección se va a estudiar el efecto de la realimentación sobre la estabilidad, el transitorio y el comportamiento en estado estacionario, y se analizarán sistemas de realimentación sencillos, como el de la Figura 5.3. El empleo de la realimentación tiene varias ventajas, tanto en los sistemas continuos como discretos. La realimentación, por ejemplo, realiza lo siguiente:

Mejora el comportamiento transitorio del sistema.

Hace disminuir la sensibilidad del sistema en lazo abierto a las variaciones en los parámetros.

Elimina los errores estacionarios si hay suficientes integradores en el sistema en lazo abierto.

Se puede estudiar la estabilidad de los sistemas en lazo cerrado empleando las herramientas indicadas en la Sección 5.2. El método del lugar de las raíces es una herramienta muy adecuada para analizar lazos de realimentación sencillos. Debido a que la realimentación cambia los polos del sistema, es importante comprender la conexión existente entre los polos discretos y el comportamiento transitorio del sistema. Esto se ha tratado en la Sección 3.6.

Valores en estado estacionario

Cuando se analizan sistemas de control es importante calcular los valores en estado estacionario de la salida y del error del sistema. Supongamos un sistema de realimentación sencillo, como el presentado en la Figura 5.3. Para generalizar, puede suponerse que la realimentación negativa unidad que tiene, -1 , se sustituye por $-H(q)$. Entonces el error $e(k)$ viene dado por

$$e(k) = \frac{1}{[1 + H_o(q)H_c(q)]} u_c(k) \quad (5.20)$$

Se puede emplear el teorema del valor final (Sección 3.5, Tabla 3.2) para calcular el valor estacionario de $e(k)$, suponiendo que existe. Si la señal de entrada es un salto

escalón, el error estacionario puede calcularse de forma sencilla haciendo $q=1$ en (5.20).

El número de integradores en el sistema en lazo abierto determina la clase de los valores de referencia que pueden ser seguidos sin errores estacionarios. Si el sistema en lazo abierto tiene p integradores, entonces el error será cero en estado estacionario (suponiendo que el sistema en lazo cerrado es asintóticamente estable) para señales de referencia que sean polinomios en k de orden menor o igual a $p-1$.

Ejemplo 5.9

Consideremos el sistema

$$y(k) = H_o(q)u(k) = \frac{q - 0.5}{(q - 0.8)(q - 1)} u(k)$$

Al cerrar el sistema, como en la Figura 5.3, se tiene

$$e(k) = \frac{(q - 0.8)(q - 1)}{(q - 0.8)(q - 1) + q - 0.5} u_c(k)$$

Supongamos una entrada u_c que es un salto escalón unidad. Como el sistema en lazo cerrado es estable, se puede usar el teorema del valor final para comprobar que el error estacionario es cero. Esto se puede ver simplemente con solo hacer $q=1$. Otra manera es observar que el sistema en lazo abierto contiene un integrador; es decir, un polo en $+1$.

Si u_c es una rampa unidad, se emplea la Tabla 3.3 para encontrar la transformada z de la rampa. Entonces el error estacionario viene dado por

$$\lim_{k \rightarrow \infty} e(k) = \lim_{z \rightarrow 1} \frac{(z - 0.8)(z - 1)}{(z - 0.8)(z - 1) + z - 0.5} \frac{z(1 - z^{-1})}{(z - 1)^2} = 0.4$$

Para homogeneizar la notación suele ser conveniente considerar las señales de entrada, o las perturbaciones, como si fueran generadas por un sistema dinámico (véase la Fig. 5.9). Se supone que la entrada al sistema dinámico es un pulso unidad δ_k ; es decir:

$$u_c(k) = H_c(q)\delta_k$$

Para generar un salto escalón se emplea $H_c(q) = q/(q-1)$; para generar una rampa, se emplea $H_c(q) = q/(q-1)^2$. El valor final se obtiene así fácilmente mediante el teorema del valor final.

Simulación

La simulación es una buena forma de estudiar el comportamiento de los sistemas dinámicos; por ejemplo, el comportamiento intermuestreos de los sistemas controlados.

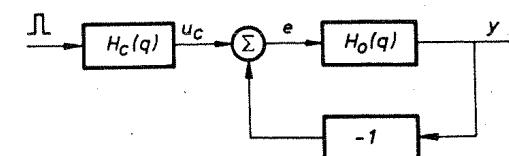


Figura 5.9 Generación del valor de referencia mediante un sistema dinámico con una entrada pulso.



dos por computador. La simulación con computador es una herramienta muy buena, pero debe recordarse que la simulación y el análisis deben emplearse juntas. Cuando se llevan a cabo simulaciones, no siempre es posible probar todas las combinaciones que son desfavorables, por ejemplo, desde el punto de vista de la estabilidad, observabilidad o alcanzabilidad. Estos casos pueden hallarse mediante el análisis.

Es importante que el programa de simulación sea tan sencillo de usar que la persona principalmente interesada pueda ser involucrada en la simulación y en la evaluación de los resultados de la simulación.

Al comienzo de los años sesenta se desarrollaron varios paquetes de simulación digital. Estos eran básicamente una realización digital de la simulación analógica. La programación se hacía mediante diagramas de bloques y módulos de operaciones prefijadas. Posteriormente se desarrollaron programas en los que se daban los modelos directamente como ecuaciones.

Para la simulación es importante tener una buena comunicación hombre-máquina; el usuario debe ser capaz de cambiar los parámetros y modificar el modelo fácilmente. El empleo de sistemas operativos de tiempo compartido hace posible la realización de programas de *simulación interactiva*, lo que significa que el usuario interacciona en línea con el computador y decide el paso siguiente en base a los resultados obtenidos hasta el momento. Una forma de realizar la interacción es que el computador haga preguntas y el usuario elija entre un conjunto de respuestas preseleccionadas. Esto se conoce como *interacción por menú*. Otra posibilidad es la *interacción por órdenes*, que es como un lenguaje de alto nivel de resolución de problemas en el que el usuario puede escoger libremente entre todas las órdenes disponibles en el sistema. Se trata de una forma más flexible de comunicarse con el computador, y es muy eficaz para el usuario experimentado, mientras que el programa tipo pregunta-respuesta es más fácil de usar para el que carece de ella.

En un paquete de simulación también es importante disponer de una forma flexible de presentar los resultados, que a menudo son gráficas. Finalmente, para ser capaz de resolver el tipo de problemas que trata este libro, es importante poder mezclar sistemas continuos y discretos.

Simnon es un paquete de simulación interactivo por órdenes desarrollado en el Departamento de Control Automático del Instituto de Tecnología de Lund. La mayor parte de los ejemplos de este libro se han generado mediante Simnon. En el Apéndice C puede encontrarse un resumen de las características de Simnon. (Véase también la lista de Referencias.)

Control del integrador doble

Como principal ejemplo para mostrar cómo varía el comportamiento con diferentes controladores emplearemos el integrador doble. La función de transferencia discreta del integrador doble para un período de muestreo $h=1$ es

$$H_0(q) = \frac{0.5(q+1)}{(q-1)^2} \quad (5.21)$$

Supongamos que el objetivo del control es hacer que la salida siga los cambios en el valor de referencia. Supongamos también que el proceso está controlado por un computador con realimentación proporcional; es decir,

$$u(k) = K[u_c(k) - y(k)] = Ke(k), \quad K > 0 \quad (5.22)$$

donde u_c es el valor de referencia. La ecuación característica del sistema en lazo cerrado es

$$(q-1)^2 + 0.5K(q+1) = 0 \quad (5.23)$$

Un esbozo del lugar de las raíces (véase la Fig. 5.10) muestra que el sistema en lazo cerrado es inestable para todos los valores de la ganancia K . Lo mismo puede verse por el criterio de estabilidad de Jury. (Compárese con el Ejemplo 5.2.)

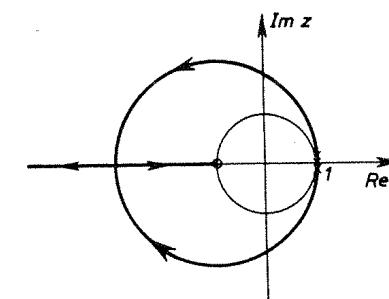


Figura 5.10 Lugar de las raíces de (5.23) cuando $K > 0$.

Para lograr un sistema estable debe modificarse el controlador. Se sabe por la síntesis de sistemas continuos que la acción derivadora mejora la estabilidad, con lo que se puede probar para los sistemas discretos la realimentación proporcional y derivadora; es decir,

$$u(k) = K[e(k) - T_D\dot{y}(k)] \quad (5.24)$$

(véase la Fig. 5.11). Entonces se muestrea también la velocidad \dot{y} y se realimenta. Para encontrar el modelo entrada-salida del sistema en lazo cerrado con el controlador (5.24), nótese que

$$\frac{dy}{dt} = u$$

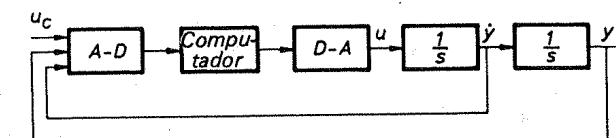


Figura 5.11 Controlador discreto del integrador doble con realimentación de la posición y de la velocidad.

Debido a que u es constante durante los intervalos de muestreo,

$$\dot{y}(k+1) - \dot{y}(k) = u(k)$$

o

$$\dot{y}(k) = \frac{1}{q-1}u(k) \quad (5.25)$$

Las ecuaciones (5.21), (5.24) y (5.25) dan el sistema en lazo cerrado

$$y(k) = \frac{0.5K(q+1)}{(q-1)(q-1+T_DK)+0.5K(q+1)}u_c(k) \quad (5.26)$$

El sistema es de segundo orden y tiene dos parámetros libres, K y T_D , que se pueden utilizar para seleccionar los polos en lazo cerrado. El sistema en lazo cerrado es estable si $K > 0$, $T_D > 0,5$ y $T_DK < 2$. El lugar de las raíces respecto de K de la ecuación característica de (5.26) puede verse en la Figura 5.12 cuando $T_D = 1,5$.

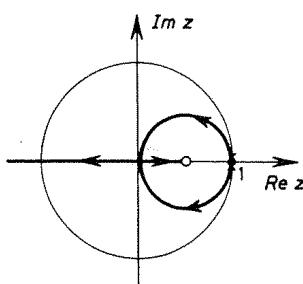


Figura 5.12 Lugar de las raíces de la ecuación característica del sistema de (5.26) cuando $T_D = 1,5$ con respecto al parámetro K .

Tomemos como señal de referencia el salto escalón. La Figura 5.13 muestra la salida continua para cuatro valores diferentes de K . El comportamiento del sistema en lazo cerrado varía de una respuesta oscilatoria a otra correctamente amortiguada.

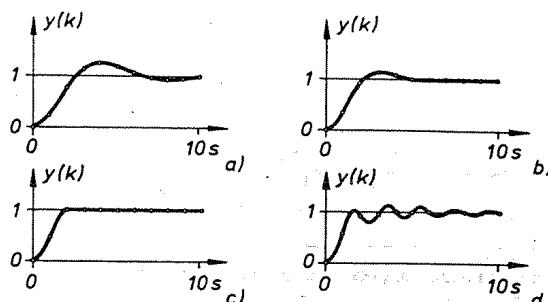


Figura 5.13 Salida continua del sistema de la Figura 5.11 cuando $T_D = 1,5$ y $K = 0,5, 0,75, 1$ y $1,25$, respectivamente.

Cuando $K = 1$, la salida es igual al valor de referencia después de dos períodos de muestreo. Esto es lo que se conoce como *control de tiempo finito* que se tratará más adelante en los Capítulos 9 y 10. Cuando $K > 1$ la salida y la señal de control oscilan debido al polo discreto situado sobre el eje real negativo. Los polos están dentro del círculo unidad si $K < \frac{4}{3}$.

Para determinar la respuesta en lazo cerrado es importante comprender la conexión entre los polos discretos y la respuesta del sistema. Esto se expuso en la Sección 3.6. En la Figura 3.7 puede verse que $K = 0,75$ corresponde a un amortiguamiento de $\zeta = 0,4$. La distancia al origen es una medida de la velocidad del sistema.

Se ha obtenido entonces un buen control utilizando el modelo discreto. Sin embargo, es preciso tener cuidado. Para ver por qué, comenzemos con el modelo de (5.21), que puede escribirse como la ecuación en diferencias

$$y(k) = 2y(k-1) - y(k-2) + 0.5u(k-1) + 0.5u(k-2)$$

El propósito del control es seguir la trayectoria de referencia $u_c(k)$. Si se escoge la señal de control de forma que la parte derecha sea igual al valor de referencia en el instante $k-1$, se obtiene el siguiente controlador causal:

$$u(k) = \frac{2q}{q+1}u_c(k) - \frac{2(2q-1)}{q+1}y(k) \quad (5.27)$$

El sistema en lazo cerrado viene dado por

$$y(k) = u_c(k-1)$$

La salida es igual al valor de referencia después de un período. Si se emplea el controlador de (5.24) con $K = 1$ y $T_D = 1,5$, lo será después de dos períodos. En la Figura 5.14 se pueden ver la respuesta escalón y la señal de control cuando se emplea la ley de control (5.27). En los instantes de muestreo, el sistema tiene el comportamiento que se había previsto, pero en la salida continua hay una oscilación. Para conocer

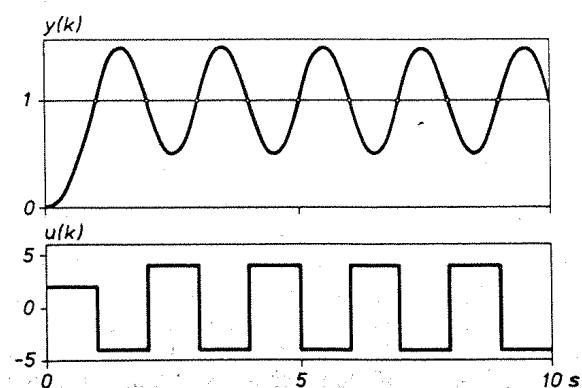


Figura 5.14 Respuesta salto y señal de control del integrador doble cuando se utiliza el controlador de (5.27).



el comportamiento entre instantes de muestreo es importante simular el sistema. Sin embargo, es posible tener una indicación del problema mediante el análisis del sistema discreto. En la señal de control puede verse que algo va mal.

La ley de control de (5.27) y (5.21) da el sistema en lazo cerrado

$$\begin{aligned} y(k) &= \frac{q(q+1)}{(q+1)[q^2 - 2q + 1 - (-2q+1)]} u_c(k) \\ &= \frac{q(q+1)}{q^2(q+1)} u_c(k) = u_c(k-1). \end{aligned}$$

El sistema en lazo cerrado es de tercer orden, el proceso tiene dos modos y el controlador uno. El cero en el límite de estabilidad se cancela con un polo. Este modo no es observable en el sistema discreto en lazo cerrado. Esto significa que la observabilidad del sistema en lazo cerrado se ha perdido por una elección incorrecta del controlador.

Hemos expuesto el comportamiento del integrador doble con algunos controladores sencillos; los resultados se pueden generalizar a sistemas más complejos. También se ha ilustrado la importancia del análisis y la simulación.

Oscilaciones ocultas

Las Figuras 5.13 y 5.14 muestran que una salida continua del proceso puede tener oscilaciones que no se ven en los instantes de muestreo. Reciben el nombre de *oscilaciones ocultas* o *rizado intermuestreo*. La simulación es una herramienta eficaz para encontrar las oscilaciones ocultas. También se puede emplear la transformada z modificada o (3.2) para calcular la salida continua entre los instantes de muestreo; sin embargo, también es ilustrativo llevar a cabo algún tipo de análisis.

El rizado intermuestreo se determina esencialmente por la dinámica en lazo abierto, ya que el sistema opera en lazo abierto entre los instantes de muestreo. Pueden distinguirse dos casos.

Sistemas en lazo abierto o cerrado, en los que hay una oscilación en la salida continua que no puede verse en la señal de control.

Oscilaciones entre los puntos de muestreo causadas por una oscilación en la señal de control.

El primer caso de rizado intermuestreo puede ocurrir si se pierde la observabilidad del sistema en lazo abierto por causa del muestreo. Entonces la función de transferencia discreta tiene polos y ceros que se cancelan. Por tanto, el efecto de los modos cancelados no se ve en los instantes de muestreo. Puede haber oscilaciones ocultas si el sistema continuo en lazo abierto tiene modos oscilatorios y si el período de muestreo coincide con la frecuencia de estos modos. Este tipo de oscilaciones ocultas sólo ocurre para unos ciertos valores del período de muestreo. Un cambio en éste hace observable al sistema y se puede ver la oscilación en la salida muestreada.

Para detectar este tipo de rizado intermuestreo es necesario comprobar la observabilidad del sistema muestreado. (Compárese con el Ejemplo 5.8.)

El segundo tipo de oscilación oculta tiene lugar si existen en el sistema en lazo abierto ceros poco amortiguados que cancela el controlador. En este caso se puede ver la oscilación en la señal de control. Esta es la causa del rizado intermuestreo que aparece en las Figuras 5.13 y 5.14. Este tipo de oscilación oculta no se detectará aunque se cambie el período de muestreo, ya que el diseño mismo hace que se cancelen los ceros del proceso.

Para resumir, no hay oscilaciones ocultas si los modos no observables en lazo abierto no son oscilatorios y si el controlador no cancela los ceros inestables o mal amortiguados del proceso.

Ejemplo 5.10

Consideremos el sistema continuo con función de transferencia

$$G(s) = \frac{1}{s+1} + \frac{\pi}{(s+0.02)^2 + \pi^2}$$

Al muestrear este sistema con $h=2$ tenemos la función de transferencia discreta

$$H(z) = \frac{1-a}{z-a}$$

donde $a=e^{-2}$.

El sistema discreto es de primer orden mientras que el sistema continuo es de tercer orden. La cancelación de polos y ceros que dan lugar a oscilaciones es una indicación de que puede haber una oscilación oculta. La Figura 5.15 muestra la respuesta escalón del sistema continuo. Los puntos de muestreo se indican por pequeños círculos. El sistema se comporta como un sistema de primer orden en los instantes de muestreo.

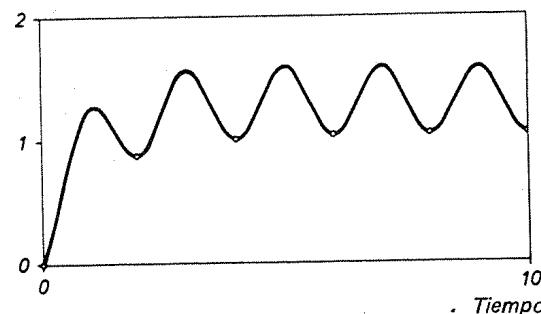


Figura 5.15 Respuesta continua a un salto escalón del sistema del Ejemplo 5.10. Los círculos indican los instantes de muestreo.

5.5. Problemas

5.1. Determinar si las siguientes ecuaciones tienen todas sus raíces dentro del círculo unidad:

- a) $z^2 - 1.5z + 0.9 = 0$
- b) $z^3 - 3z^2 + 2z - 0.5 = 0$

c) $z^3 - 2z^2 + 2z - 0.5 = 0$
d) $z^3 + 5z^2 - 0.25z - 1.25 = 0$
e) $z^3 - 1.7z^2 + 1.7z - 0.7 = 0$

- 5.2. Considérese el sistema de la Figura 5.3 y sea

$$H_0(z) = \frac{K}{z(z - 0.2)(z - 0.4)} \quad K > 0$$

Determinar los valores de K para los que el sistema en lazo cerrado es estable.

- 5.3. Considérese el sistema de la Figura 4.20a). Supóngase que se muestrea periódicamente con período h y que el convertidor D-A mantiene constante la señal de control durante el intervalo de muestreo. Se supone que el algoritmo de control es

$$u(kh) = K[u_c(kh - \tau) - y(kh - \tau)]$$

donde $k > 0$ y τ es el tiempo de cálculo. La función de transferencia del proceso es

$$G(s) = \frac{1}{s}$$

- a) ¿Para qué valores de la ganancia del regulador, K , el sistema en lazo cerrado será estable si $\tau = 0$ y $\tau = h$?
b) Compárese este sistema con el sistema continuo correspondiente; es decir, cuando hay un controlador proporcional continuo y un retardo en el proceso.

- 5.4. Determinar la curva de Nyquist para el sistema

$$H_0(z) = \frac{1}{z - 0.5}$$

- 5.5. Determinar una función de Lyapunov para la ecuación no lineal del Ejemplo 1.5 y demostrar que la solución es asintóticamente estable.

- 5.6. Del sistema

$$x(k + 1) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}x(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u(k)$$

$$y(k) = [0 \ 1]x(k)$$

se obtienen los siguientes valores

$$\begin{aligned} y(1) &= 0 & u(1) &= 1 \\ y(2) &= 1 & u(2) &= -1 \end{aligned}$$

Determinar el valor del estado en $k = 3$.

- 5.7. El siguiente sistema, ¿es a) observable, b) alcanzable?

$$x(k + 1) = \begin{bmatrix} 0.5 & -0.5 \\ 0 & 0.25 \end{bmatrix}x(k) + \begin{bmatrix} 6 \\ 4 \end{bmatrix}u(k)$$

$$y(k) = [2 \ -4]x(k)$$

- 5.8. ¿Es alcanzable el siguiente sistema?

$$x(k + 1) = \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix}x(k) + \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}u(k)$$

Supongamos que se introduce una entrada escalar $u'(k)$ tal que

$$u(k) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}u'(k)$$

El sistema, ¿es sólo alcanzable desde $u'(k)$?

- 5.9. Dado el sistema

$$x(k + 1) = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}x(k) + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}u(k)$$

- a) Determinar una secuencia de control tal que lleve al sistema desde el estado inicial $x^T(0) = [1, 1, 1]$ hasta el origen.
b) ¿Cuál es el mínimo número de pasos para el que se resuelve el problema en a)?
c) Explicar por qué no es posible encontrar una secuencia tal que alcance el estado $[1, 1, 1]^T$ desde el origen.

- 5.10. Comprobar la fórmula para W_c^{-1} dada en el Ejemplo 5.6 para un sistema de orden n .

- 5.11. El sistema

$$x(k + 1) = \Phi x(k) + \Gamma u(k)$$

se ha obtenido a partir del sistema

$$z(k + 1) = Fz(k) + G u(k)$$

por la transformación lineal

$$z = Tx$$

- a) Emplear el resultado de la Sección 5.3 para deducir una fórmula para T cuando $\dim(u) = 1$; $\dim(u) = r$.
b) Emplear el resultado para resolver el Problema 3.7.

- 5.12. Determinar la estabilidad y el valor estacionario de la salida del sistema descrito en la Figura 5.16 con

$$H_0(q) = \frac{1}{q(q - 0.5)}$$

cuando u_c es una función salto escalón y a) $H_r(q) = K$ (controlador proporcional) $K > 0$ y
b) $H_r(q) = Kq/(q - 1)$ (controlador integral) $K > 0$.

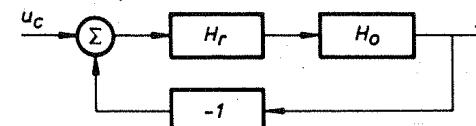


Figura 5.16

- 5.13. Considérese el sistema del Problema 5.12. Determinar el error estacionario entre la señal de mando, u_c , y la salida cuando u_c es una rampa unidad; es decir, $u_c(k) = k$. Supóngase que H_r es a) un controlador proporcional y b) un controlador integral.

- 5.14. Muestrear el sistema

$$G(s) = \frac{s + 1}{s^2 + 0.2s + 1}$$

y determinar los intervalos de muestreo para los que la respuesta del sistema tendrá oscilaciones ocultas. Verificarlo mediante simulaciones.

- 5.15. Considérese el sistema de tanques con el operador de transferencia discreta dado en el Problema 3.10b).

- a) Introdúzcase un controlador como en la Figura 5.15. Utilícese una entrada de control tipo salto y determínese el error estacionario cuando se emplea un controlador proporcional K ; un controlador integral $K/(1 - q^{-1})$.
 b) Simúlese el sistema con los controladores de a). Escoger K de forma tal que los polos del sistema en lazo cerrado correspondan a un amortiguamiento $\zeta = 0.7$.

- 5.16. Considérese el sistema de la Figura 5.3. Deducir una fórmula para el *coeficiente de error* de velocidad. Esta es la expresión del error estacionario cuando u_c es la rampa unidad.

5.6. Referencias

Los trabajos originales sobre las pruebas para comprobar si un polinomio tiene todos sus polos dentro del círculo unidad son

- SCHUR, J. (1918): «Über Potenzreihen, die im Inneren des Einheitskreises beschränkt sind, II,» *Zeitschrift für die reine und angewandte Matematik*, 148, 122-45.
 COHN, A. (1922): «Über die Anzahl der Wurzeln einer algebraischen Gleichung in einem Kreise,» *Mathematische Zeitschrift*, 14, 110-48.

La prueba de Jury es una simplificación de la prueba de Schur-Cohn y se encuentra en JURY, E. I. y J. BLANCHARD (1961): «A Stability Test for Linear Discrete Time Systems in Table Form,» *Proc. IRE*, 49, 1947-48.

Una demostración sencilla de la prueba de Jury está en
 ÅSTRÖM, K. J. (1970): *Introduction to Stochastic Control Theory*. New York: Academic Press.

El empleo de la teoría de Lyapunov para los sistemas de control discretos se introduce en KALMAN, R. E. y J. E. BERTRAM (1960): «Control System Analysis and Design via the Second Method of Lyapunov: II Discrete-Time Systems,» *Trans. ASME Journal Basic Eng.*, Series D, no. 3, 394-400.

La controlabilidad y la observabilidad son conceptos introducidos por Kalman en conexión con el análisis de los sistemas de control óptimo. Véase

KALMAN, R. E. (1961): «On the General Theory of Control Systems,» Proc. First IFAC Congress, Moscow: Butterworths, 1, 481-92.

KALMAN, R. E., Y. C. HO, y K. S. NARENDRA (1963): Controllability of Linear Dynamical Systems. In *Contributions to Differential Equations*, 1, 189-213, New York: John Wiley.

El concepto de oscilaciones ocultas y su causa se exponen en

JURY, E. I. (1975): «Hidden Oscillations in Sampled-Data Control Systems,» *AIEE Trans.*, 75, pt. II, 391-95.

Los aspectos generales de la simulación digital se tratan en

GORDON, G. (1969): *System Simulation*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.

El paquete de simulación Simon se describe en

ELMQVIST, H. (1977): «Simon: An Interactive Simulation Program for Nonlinear Systems,» *Proc. International Symposium Simulation '77*, Montreux.

ÅSTRÖM, K. J. (1982): A Simon Tutorial. Department of Automatic Control, Lund Inst. of Tech., CODEN: LUTFD2/(TFRT-3168)/1-52/(1982).

ÅSTRÖM, K. J. (1983): «Computer-Aided Modeling, Analysis and Design of Control Systems: A perspective,» *IEEE Control Systems Magazine*, 3 no. 2, 4-16.

Modelos de perturbaciones

Objetivo:

Discutir diferentes modos de eliminar las perturbaciones y sus efectos. Introducir modelos determinísticos y estocásticos para las perturbaciones. Analizar la respuesta de sistemas lineales frente a perturbaciones.

6.1. Introducción

La presencia de perturbaciones es una de las principales razones para utilizar la teoría de control. Sin perturbaciones no es necesaria la realimentación. El carácter de las perturbaciones impone limitaciones fundamentales a los sistemas de control. La cantidad de ruido en un servosistema limita el ancho de banda final del sistema en lazo cerrado. La naturaleza de las perturbaciones determina la calidad de la regulación en un sistema de control de procesos. Las perturbaciones también proporcionan importante información sobre las propiedades del sistema. Por tanto, investigando las características de las perturbaciones es posible detectar el estado del proceso, incluyendo indicios de malos funcionamientos.

En este capítulo estudiaremos diferentes métodos para tratar las perturbaciones en un sistema y analizar sus efectos. En primer lugar se da una revisión de los diferentes métodos para eliminar perturbaciones. Esto incluye el uso de realimentación, prealimentación y predicción. La exposición justifica las diferentes formas de descripción de perturbaciones.

Los modelos clásicos de perturbaciones: impulso, escalón, rampa y sinusoide, se describen en la Sección 6.3. Todas estas perturbaciones se pueden considerar generadas por sistemas lineales con determinadas condiciones iniciales. Por tanto, el problema de analizar el efecto de las perturbaciones en un sistema lineal puede reducirse a un problema de condiciones iniciales. Desde el punto de vista de entrada-salida, una perturbación también puede modelarse como la respuesta a un impulso de un filtro lineal. En este caso el análisis de perturbaciones puede reducirse al cálculo de

una respuesta. Esto es particularmente útil para perturbaciones de tipo escalón o senoidal. En cualquier caso, el análisis de perturbaciones puede realizarse con las herramientas desarrolladas en el Capítulo 5. Cuando se necesita conocer la respuesta de un sistema frente a una perturbación específica es con frecuencia necesario recurrir a la simulación. Esto es fácil de realizar con un programa de simulación, ya que el análisis de perturbaciones se reduce de nuevo a un problema de valores iniciales.

Cuando la fuente de la perturbación no puede eliminarse ni medirse, es necesario recurrir a la predicción. Para hacer esto es necesario disponer de modelos de perturbaciones que proporcionen una formulación razonable del problema de predicción. Para este propósito, en la Sección 6.4 se introduce el concepto de perturbaciones deterministas a trozos.

Otro modo de abordar el problema de la predicción consiste en describir las perturbaciones como un proceso aleatorio. Esta formulación se presenta en la Sección 6.5. También se presenta una versión simple de la famosa teoría de predicción de Wiener-Kolmogorov-Kalman. Como veremos en los Capítulos 11 y 12, el error de la predicción supone una limitación fundamental para la calidad de la regulación. En la Sección 6.6 se discuten brevemente los procesos estocásticos continuos. Estos modelos se requieren para formular modelos y especificaciones continuas en el tiempo. En la Sección 6.7 se trata el muestreo de modelos continuos.

6.2. Reducción de los efectos de las perturbaciones

Antes de entrar en los detalles de los modelos de perturbaciones, resulta útil discutir cómo pueden reducirse los efectos de las perturbaciones en un sistema. Las perturbaciones se pueden reducir en la fuente que las produce. Los efectos de las perturbaciones también se pueden reducir mediante realimentación local o por prealimentación desde la fuente de perturbaciones. También puede utilizarse la predicción para estimar perturbaciones que no pueden medirse. En este caso, la parte predecible de la perturbación se puede reducir por prealimentación. Los diferentes enfoques se discutirán más adelante.

Reducción en la fuente

El método más obvio de reducir los efectos de las perturbaciones consiste en intentar reducir las perturbaciones en la misma fuente. Este método está estrechamente relacionado con el procesado de señales. Ejemplos típicos son:

Reducir las variaciones de la composición introduciendo tanques con un mezclador eficaz.

Reducir las fuerzas de fricción en un servo mejorando los cojinetes.

Mover los sensores a posiciones en que haya menos perturbaciones.

Modificar los sensores electrónicos para obtener el mínimo ruido.

Sustituir los sensores por otros que tengan mejor respuesta.

Cambiar el procedimiento de muestreo espaciando las muestras en el tiempo del modo más adecuado o espaciándolas para obtener la mejor representación de las características del proceso.

Son unos pocos ejemplos pero es muy importante tener en cuenta estas posibilidades.

Reducción por realimentación local

Si la perturbación no puede reducirse en la fuente se debe intentar reducirla mediante realimentación local. El principio genérico de este método se ilustra en la Figura 6.1. Para este método se necesita que la perturbación se introduzca localmente en el sistema por un camino bien conocido. También es necesario poder medir la variable perturbada. Y acceder a una variable de control que se introduzca al sistema en las proximidades de la perturbación. En este caso, el efecto de la perturbación se puede reducir utilizando realimentación local. La dinámica que relaciona la variable medida con la variable de control debería permitir el uso de un lazo de control con alta ganancia.

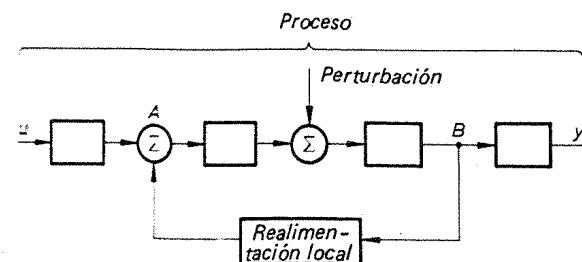


Figura 6.1 Reducción de perturbaciones por realimentación local. La perturbación debe entrar en el sistema entre los puntos A y B. La dinámica entre A y B debe ser tal que se pueda utilizar en el lazo una ganancia grande.

Frecuentemente, este uso de la realimentación es muy simple y efectivo, puesto que no es necesario disponer de información detallada sobre las características del proceso, con la condición de que se utilice una ganancia alta en el lazo. Sin embargo, se requiere un lazo extra de realimentación. A continuación, se presentan ejemplos típicos de realimentación local:

Reducir las variaciones de la presión de alimentación a las válvulas, instrumentos y reguladores, introduciendo un regulador de presión.

Reducir las variaciones en el control de temperatura estabilizando la fuente de alimentación de tensión.

Reducción por prealimentación

Las perturbaciones medibles también pueden reducirse por prealimentación. El principio genérico se ilustra en la Figura 6.2. Se mide la perturbación y se genera una señal de control que aplicada al proceso intenta contrarrestar el efecto de la perturbación. Si las funciones de transferencia que relacionan la salida y con la perturbación w y con el control u son H_w y H_p , respectivamente, es fácil ver que la función de transferencia H_{ff} del compensador de prealimentación, idealmente sería

$$H_{ff} = -H_p^{-1}H_w$$

Si la función de transferencia es inestable o irrealizable, debe elegirse, en su lugar, una aproximación conveniente. El diseño del compensador de prealimentación a menudo se basa en un modelo estático simple. En este caso, la función de transferencia H_{ff} es simplemente una ganancia estática. La prealimentación es particularmente útil para perturbaciones generadas por cambios en la consigna o señal de referencia, o bien, para procesos en cascada en los que las perturbaciones en un proceso están generadas por variaciones en los procesos precedentes.

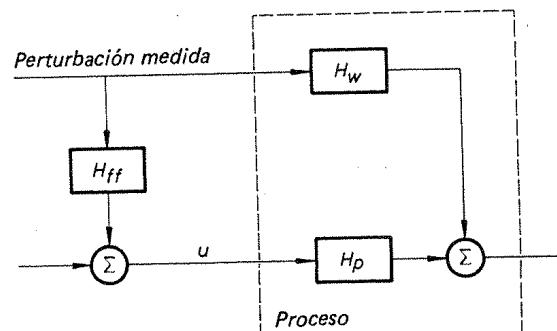


Figura 6.2 Reducción de perturbaciones por prealimentación.

Reducción por predicción

La reducción por predicción es una extensión del principio de prealimentación que puede utilizarse cuando la perturbación no puede medirse. El principio es muy simple: la perturbación se predice midiendo señales, y la prealimentación se genera a partir de la predicción.

Es importante observar que no es necesario predecir la propia perturbación; es suficiente modelar una señal que represente el efecto de la perturbación en las variables del proceso que interesen.

Objetivos del modelado

Puesto que se han especificado diferentes métodos de reducción de los efectos de las perturbaciones, es fácil comprender ahora, cómo se utilizarán los modelos de perturbaciones. Para evaluar las necesidades de reducción de perturbaciones es necesario poder estimar la influencia de las perturbaciones en las variables de interés del sistema, lo cual es básicamente un problema de análisis de respuesta de un sistema frente a una entrada determinada. Los modelos utilizados para las perturbaciones pueden ser bastante simples, siempre que representen las características reales de perturbaciones auténticas. Análogamente, los modelos simples también se pueden utilizar para estimar posibles mejoras obtenidas con realimentación local y prealimentación.

Si se aplica predicción se necesitan modelos de perturbaciones más precisos. En este caso, las mejoras obtenidas dependen críticamente del carácter de las perturbaciones. También existen algunas dificultades fundamentales en la formulación de los modelos de perturbaciones que proporcionan un problema apreciable de predicción.

6.3. Los modelos clásicos de perturbaciones

A continuación se presenta una discusión de los modelos comunes de perturbaciones. En primer lugar se discute el carácter de las perturbaciones. Las perturbaciones se clasifican como perturbaciones en la carga, errores de medida y variación de parámetros. Se discuten algunos modelos de perturbaciones simples: impulso, escalón, rampa y sinusoidal. Se demuestra que estas perturbaciones se pueden generar como salidas de sistemas lineales simples. Esta observación sugiere una clase general de perturbaciones que se generan a partir de sistemas lineales arbitrarios. La observación, también hace posible la simplificación del análisis de los efectos de las perturbaciones en un sistema.

Carácter de las perturbaciones

Se acostumbra a distinguir entre diferentes tipos de perturbaciones, tales como perturbaciones en la carga, errores de medida y variación de parámetros.

Perturbaciones en la carga. Las perturbaciones en la carga influyen en las variables del proceso. Pueden representar fuerzas de perturbación en un sistema mecánico; por ejemplo, ráfagas de viento en una antena estabilizada, olas en un barco, carga en un motor. En el control de procesos, las perturbaciones en la carga pueden ser condiciones de variación del flujo de alimentación o variaciones en el flujo que se demanda. En un sistema térmico las perturbaciones en la carga pueden ser variaciones en la temperatura ambiente. Las perturbaciones en la carga varían, generalmente, muy despacio. También pueden ser periódicas; por ejemplo, las olas en un sistema de control de un barco.

Errores en la medida. Los errores en la medida se introducen en los sensores. Pueden existir errores en estado estacionario en algunos sensores debidos a errores en la calibración. Sin embargo, los errores en la medida generalmente tienen componentes de alta frecuencia. También puede haber errores dinámicos debidos a la dinámica del sensor. Un caso típico es un termopar, el cual puede tener una constante de tiempo de 10-50 s dependiendo de su encapsulado. También pueden producirse complicadas interacciones dinámicas entre los sensores y el proceso. Ejemplos típicos son las medidas giroscópicas y las medidas de nivel de líquido en los reactores nucleares. El carácter del error de la medida depende con frecuencia del filtrado en los instrumentos. Es a menudo una buena idea, observar el instrumento y modificar el filtro para corregir un problema particular.

En algunos casos no es posible medir directamente la variable controlada; en este caso el valor de la variable se infiere por medida indirecta de otras variables. La relación entre las variables controladas y medidas puede ser bastante compleja. Esta relación puede caracterizarse como un sistema dinámico no lineal y variable con el tiempo. Una situación común es que un instrumento puede dar una indicación rápidamente con grandes errores y otro instrumento puede dar una medida precisa después de un tiempo largo. Por ejemplo, las medidas precisas pueden obtenerse en los laboratorios de medida.

Variación de parámetros. En este libro se utiliza la teoría de sistemas lineales. Las perturbaciones en la carga y la medida de ruido aparecen adicionalmente. Sin embargo, los sistemas reales son con frecuencia no lineales. Esto significa que las perturbaciones se pueden introducir de un modo más complejo. Ya que los modelos lineales se obtiene linealizando los modelos no lineales, algunas perturbaciones pueden aparecer como variación en los parámetros del modelo lineal.

Modelos simples de perturbaciones

Hay cuatro tipos diferentes de perturbaciones—impulso, escalón, rampa y sinusoidal—que se utilizan comúnmente en el análisis de sistemas de control. Estas perturbaciones se ilustran en la Figura 6.3 y a continuación se presenta un estudio de sus propiedades.

El impulso y el pulso. El *impulso* y el *pulso* son idealizaciones simples de perturbaciones repentinas de corta duración. Pueden representar tanto perturbaciones

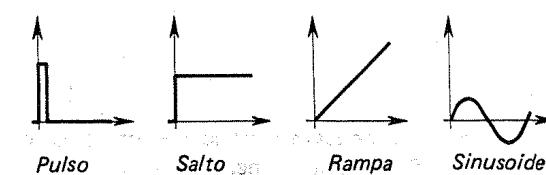


Figura 6.3 Modelos idealizados de perturbaciones sencillas.

en la carga como errores en las medidas. Para sistemas continuos la perturbación es un impulso (una función delta); para sistemas muestreados la perturbación se modela como un pulso de amplitud unidad y de un período de muestreo de duración.

El pulso e impulso también son importantes por razones teóricas, ya que la respuesta de un sistema lineal continuo está completamente especificada por su respuesta ante un impulso, y en un sistema lineal y discreto por su respuesta a un pulso.

El escalón. La señal escalón es otro prototipo de perturbación (véase la Fig. 6.3). Se utiliza típicamente para representar una perturbación en la carga o como una desviación en una medida.

La rampa. La rampa es una señal nula para instantes de tiempo negativos y crece linealmente para instantes positivos (véase la Fig. 6.3). Se utiliza para representar errores bruscos de medidas y perturbaciones que aparecen repentinamente para desaparecer posteriormente. En la práctica, las perturbaciones suelen estar acotadas; sin embargo, la rampa es una idealización útil.

La sinusode. La onda sinusoidal es el prototipo de perturbación periódica. La elección de la frecuencia hace posible la representación de perturbaciones de baja frecuencia en la carga, así como medidas de ruido de alta frecuencia.

Es conveniente ver a las perturbaciones como elementos generados por sistemas dinámicos. Desde el punto de vista de entrada-salida, las perturbaciones pueden describirse como respuestas a impulsos. Las perturbaciones también se pueden considerar como la respuesta de sistemas dinámicos a una entrada nula pero con condiciones iniciales no nulas. En ambos casos las características más importantes de las perturbaciones están descritas por la dinámica del sistema que las genera. El método puede, por supuesto, aplicarse tanto a sistemas continuos como a sistemas discretos.

El escalón puede generarse con un integrador, la rampa con un doble integrador y la sinusode con un oscilador armónico (comparar los Ejemplos A.1 y A.3 del Apéndice A).

Análisis de los efectos de las perturbaciones

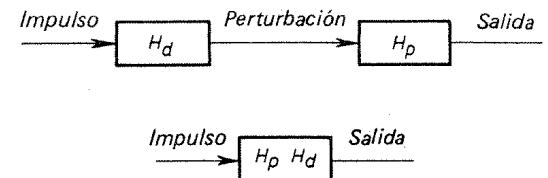
Puesto que las perturbaciones se pueden generar como salidas de sistemas dinámicos lineales, los efectos de las perturbaciones en un sistema se pueden analizar directamente.

Ánalisis en el espacio de estados. Ya que las perturbaciones se pueden considerar como soluciones de ecuaciones diferenciales lineales con condiciones iniciales, es fácil estudiar el efecto de las perturbaciones en un sistema lineal. Las variables de estado del sistema simplemente se incrementan con las variables de estado empleadas para

describir las perturbaciones. Entonces el conjunto de ecuaciones se resuelve para las condiciones iniciales, las cuales corresponden a las perturbaciones. De este modo, el análisis de perturbaciones puede tratarse con las herramientas para análisis y simulación presentadas en el Capítulo 5.

Análisis de entrada-salida. Ya que también puede considerarse que las perturbaciones se generan aplicando un impulso a un sistema lineal, es posible analizar el efecto de las perturbaciones en un sistema lineal situándolo en cascada con otro sistema que genere las perturbaciones (véase la Fig. 6.4). Este punto de vista es particularmente útil cuando se emplean los teoremas del valor inicial y valor final o los diagramas de Bode para realizar estimaciones groseras del efecto de las perturbaciones. También es útil en simulación.

Figura 6.4 Cómo se puede reducir el análisis de los efectos de las perturbaciones al cálculo de una respuesta a un impulso (compárese con la Figura 5.9).



6.4. Perturbaciones determinísticas a trozos

Los modelos de perturbaciones discutidos en la Sección 6.3 son útiles para analizar los efectos de las perturbaciones en un sistema. Con estos modelos, también pueden estudiarse posibles mejoras utilizando realimentación local o prealimentación. Sin embargo, los modelos de perturbación discutidos no son adecuados para investigar la reducción de perturbaciones por predicción. Para predecir una señal es necesario disponer de modelos realistas de perturbaciones. Es un problema fundamental construir modelos que permitan una formulación razonable del problema de predicción. Este problema fundamental es el primer asunto de esta sección. Un estudio del problema conduce a la introducción de las perturbaciones determinísticas a trozos. Modelos alternativos, que también permiten la formulación del problema de predicción, se discuten en las Secciones 6.5 y 6.6.

Un problema fundamental

No es trivial construir modelos de perturbaciones que permitan una formulación razonable de un problema de predicción.

Ejemplo 6.1 – Predictor para una señal escalón

Para predecir el valor futuro de una señal escalón parece natural utilizar los valores actuales de la señal. Para señales discretas el predictor es

$$\hat{y}[(k+m)h | kh] = y(kh)$$

La notación $\hat{y}(t|s)$ indica la predicción de $y(t)$ basada en los datos disponibles en el instante s .

Este predictor tiene un error de predicción en los instantes $t=0, h, 2h, \dots, (m-1)h$; es decir, m pasos después del cambio en escalón en y . A partir de este punto, predice la señal sin error.

Ejemplo 6.2 – Predictor para una señal en rampa

Se puede construir un predictor para rampas calculando la pendiente entre las observaciones pasada y presente y efectuando una extrapolación lineal, que puede expresarse por la fórmula

$$\begin{aligned}\hat{y}[(k+m)h|kh] &= y(kh) + m[y(kh) - y(kh-h)] \\ &= (1+m)y(kh) - my(kh-h)\end{aligned}$$

Este predictor tiene un error inicial para $t=h, 2h, \dots, mh$. Despues, predice la señal sin error.

La idea básica

Estos ejemplos indican que el error de predicción será cero excepto en unos pocos puntos. Esta observación no está muy de acuerdo con la experiencia práctica de que las perturbaciones son muy difíciles de predecir. La explicación está en que el escalón y la rampa no son buenos modelos para el problema de la predicción. De hecho, no es una tarea trivial encontrar señales que conduzcan a un problema de predicción razonable. Las señales analíticas son inútiles porque una señal analítica está definida de forma única por sus valores en un intervalo arbitrariamente corto. El escalón y la rampa son funciones analíticas en todos los puntos salvo en el origen.

Una posibilidad de construcción de señales que son poco regulares, consiste en introducir más puntos de irregularidad. De este modo se pueden introducir señales que se generan con sistemas dinámicos lineales con puntos irregulares. En lugar de tener un pulso en el origen, se introducen entradas que son diferentes de cero en varios puntos. Se obtiene una clase interesante de señales si se supone que los pulsos están aislados y extendidos al menos por n muestras, donde n es el orden del sistema. Se supone que no se conoce *a priori* cuándo ocurren los pulsos. La amplitud de los pulsos también es desconocida. Estas señales se denominan *señales determinísticas a trozos*. El nombre proviene del hecho de que las señales son determinísticas excepto en puntos aislados, donde cambian de un modo impredecible. Un ejemplo de señales determinísticas a trozos se muestra en la Figura 6.5.

Se pueden formular problemas de predicción razonables para señales determinísticas a trozos.

Modelos del espacio de estados

Sea una señal generada por el sistema dinámico

$$\begin{aligned}x(k+1) &= \Phi x(k) + v(k) \\ y(k) &= Cx(k)\end{aligned}\quad (6.1)$$

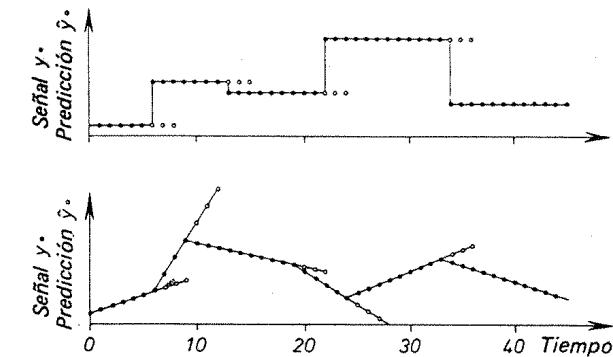


Figura 6.5 Señales constantes a trozos y lineal a trozos y sus predicciones a m pasos cuando $m=3$.

Se supone que la salida y es un escalar y que el sistema es completamente observable. Se supone que la entrada v es nula salvo en puntos aislados. Si se conoce el estado del sistema se puede predecir directamente el estado sobre un intervalo donde la entrada es cero. Sin embargo, cuando hay un pulso, el estado puede cambiar de forma arbitraria, pero después de un pulso siempre habrá un intervalo donde la entrada es cero. Como el sistema es observable, puede calcularse el estado del proceso. Entonces pueden darse predicciones exactas hasta que aparezca un nuevo pulso. Este argumento puede formularse matemáticamente del siguiente modo. De la deducción de la condición de observabilidad en la Sección 5.3 se encuentra que el estado está definido por

$$x(k-n+1) = W_o^{-1}[y(k-n+1) \cdots y(k)]^T \quad (6.2)$$

donde W_o es la matriz de observabilidad dada por la ecuación (5.19). El siguiente predictor da el estado de m pasos anteriores

$$\hat{x}(k+m|k) = \Phi^{m+n-1} W_o^{-1}[y(k-n+1) \cdots y(k)]^T \quad (6.3)$$

El predictor para la señal se obtiene de una combinación lineal de n valores de la señal medida. El predictor se puede expresar como

$$\hat{x}(k+m|k) = P^*(q^{-1})y(k)$$

donde P es un polinomio de grado $n-1$.

El predictor también se puede representar por la ecuación recursiva

$$\begin{aligned}\hat{x}(k|k) &= \Phi\hat{x}(k-1|k-1) + K[y(k) - C\Phi\hat{x}(k-1|k-1)] \\ \hat{x}(k+m|k) &= \Phi^m\hat{x}(k|k)\end{aligned}\quad (6.4)$$

donde la matriz K se elige de modo que todos los valores propios de la matriz $(I - KC)\Phi$ sean nulos.

Cálculos simples para un integrador y un integrador doble dan los mismos predictores que en los Ejemplos 6.1 y 6.2. Esto es una consecuencia de que las características importantes de las perturbaciones se captan por la dinámica del sistema que genera las perturbaciones. Esta dinámica determina los predictores de modo único; no importa si al sistema se le ha aplicado un pulso simple o varios pulsos. Las propiedades del predictor se ilustran en la Figura 6.5.

Modelos de entrada-salida

Ya que el predictor para una señal determinista a trozos es un polinomio, parece normal obtenerlo directamente mediante cálculo de polinomios. Para este propósito se supone que la señal está generada con el sistema dinámico

$$y(k) = \frac{C(q)}{A(q)} w(k) \quad (6.5)$$

donde la entrada w es una señal nula excepto en una serie de puntos aislados, los cuales están más espaciados que $\text{grad } A + m$ y $\text{grad } C = \text{gra } A$. Definamos $F(z)$ y $G(z)$ a través de la identidad

$$z^{m-1}C(z) = A(z)F(z) + G(z)$$

Puede demostrarse que el predictor de m pasos para y está definido por la ecuación en diferencias

$$C(q)\hat{y}(k+m|k) = qG(q)y(k) \quad (6.6)$$

En la Sección 6.10 se da una referencia en la que se demuestra esta expresión.

Nota. Obsérvese que las señales discutidas en esta sección son similares a las de la Sección 6.3 en el sentido de que están caracterizadas por sistemas dinámicos. La única diferencia entre las señales es que las entradas al sistema son diferentes. Esta idea se amplía en la siguiente sección.

6.5. Modelos estocásticos de perturbaciones

Es normal utilizar conceptos *estocásticos* o aleatorios para describir las perturbaciones. Por este método se puede describir una clase amplia de perturbaciones, las cuales permiten una buena formulación del problema de predicción. De hecho, la teoría de procesos aleatorios y la teoría de predicción se desarrollaron en estrecha interacción.

La teoría general de procesos estocásticos es bastante compleja. Afortunadamente, para trabajar con computadores es generalmente suficiente utilizar un caso especial de la teoría general, que requiere menos complejidad. Esta teoría se desarrolla en esta sección. En primer lugar, se muestran algunos elementos de la teoría de procesos aleatorios, para discutir posteriormente el ruido blanco discreto. De este modo,

perturbaciones se modelan como salidas de sistemas dinámicos cuyas entradas tienen ruido blanco. Por tanto, los modelos de perturbaciones son similares a los modelos discutidos en las secciones previas; la única diferencia es el carácter de las entradas al sistema. También se proporcionan herramientas para analizar las propiedades del sistema.

Procesos estocásticos

El concepto de proceso estocástico es difícil. Se han necesitado cientos de años de trabajo de grandes investigadores para encontrar ideas correctas. El concepto maduró en el trabajo realizado por el matemático Kolmogorov alrededor de 1930. Mostraremos una representación simple de las ideas. Los lectores interesados en el tema deben consultar las referencias.

Los procesos estocásticos (procesos aleatorios, funciones aleatorias) pueden considerarse como una familia de variables estocásticas $\{x(t), t \in T\}$. Las variables estocásticas son funciones del parámetro t , el cual pertenece a T , denominado *conjunto de índices*. En la teoría de control estocástico, la variable t se interpreta como el tiempo. Entonces T es un conjunto de variables reales. Cuando se consideran sistemas muestreados, como se hace en este libro, T es el conjunto de los intervalos de muestreo; es decir, $T = \{\dots, -h, 0, h, \dots\}$ o $T = \{\dots, -1, 0, 1, \dots\}$ cuando el período de muestreo se elige como unidad de tiempo.

Un proceso aleatorio puede considerarse como una función $x(t, \omega)$ de dos variables. Para $\omega = \omega_0$ la función $x(\cdot, \omega_0)$ es una función temporal denominada *realización*. Para $t = t_0$, la función $x(t_0, \cdot)$ es una variable aleatoria. De este modo, un proceso aleatorio se puede considerar producido por un generador de señales aleatorias. A menudo se suprime el argumento ω en la notación.

Procesos estocásticos completamente determinísticos. Una posibilidad de obtener un proceso aleatorio consiste en escoger las condiciones iniciales de una ecuación diferencial como variable aleatoria y generar las funciones temporales para resolver la ecuación diferencial. Sin embargo, estos tipos de procesos aleatorios no son muy interesantes porque no son lo suficientemente aleatorios. Esto se ve claramente si se considera el proceso estocástico generado por un integrador, con condiciones iniciales aleatorias. Puesto que la salida del integrador es constante se deduce que

$$x(t, \omega) - x(t-h, \omega) = 0$$

para todo t, h y ω . Un proceso estocástico con esta propiedad se denomina *proceso estocástico completamente determinístico*, porque sus valores futuros se pueden predecir exactamente a partir de sus valores pasados.

En general se dirá que un proceso aleatorio $x(t, \omega)$ se denomina *completamente determinístico* si

$$\ell x(t, \omega) = 0, \quad \text{para casi todos los } \omega$$

donde ℓ es un operador lineal arbitrario no nulo. Esto significa que los procesos aleatorios completamente determinísticos pueden predecirse exactamente con predic-



tores lineales para casi todos los ω . (Casi todos los ω significa todos los ω excepto un posible conjunto de puntos de medida cero.)

El proceso aleatorio completamente determinístico está ligado estrechamente a las señales consideradas en la Sección 6.3. Estas señales serán un proceso aleatorio completamente determinístico si las condiciones iniciales del sistema dinámico se eligen como un proceso aleatorio. El proceso completamente determinístico, generalmente no se considera porque es demasiado regular para ser interesante.

Conceptos. A continuación se muestran algunos conceptos importantes para los procesos aleatorios.

Los valores de un proceso aleatorio en n intervalos diferentes son variables aleatorias n -dimensionales. La función

$$F(\xi_1, \dots, \xi_n; t_1, \dots, t_n) = P\{x(t_1) \leq \xi_1, \dots, x(t_n) \leq \xi_n\} \quad (6.7)$$

donde P indica la probabilidad, se denomina *función de distribución de dimensión finita* del proceso aleatorio. Se ilustra en la Figura 6.6. Un proceso aleatorio se denomina *Gaussiano* o *normal*, si todas las distribuciones de dimensión finita son normales.

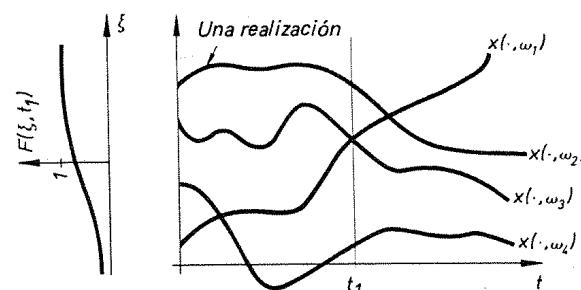


Figura 6.6 Proceso estocástico y función de distribución de dimensión finita.

La función *valor medio* de un proceso aleatorio x se define

$$m(t) = Ex(t) = \int_{-\infty}^{\infty} \xi dF(\xi, t) \quad (6.8)$$

La función valor medio es una función normal del tiempo. La definición de momentos superiores es similar.

La función *covarianza* de un proceso se define por

$$\begin{aligned} r_{xx}(s, t) &= \text{cov}[x(s), x(t)] \\ &= E[x(s) - m(s)][x(t) - m(t)]^T \\ &= \iint [\xi_1 - m(s)][\xi_2 - m(t)]^T dF(\xi_1, \xi_2; s, t) \end{aligned} \quad (6.9)$$

Un proceso aleatorio Gausiano está caracterizado completamente por su función valor medio y su función de covarianza.

La función de covarianza cruzada

$$r_{xy}(s, t) = \text{cov}[x(s), y(t)]$$

de dos procesos estocásticos se define de modo análogo.

Un proceso estocástico se denomina *estacionario* si la distribución de dimensión finita de $x(t_1), x(t_2), \dots, x(t_n)$ es idéntica a la distribución de $x(t_1 + \tau), x(t_2 + \tau), \dots, x(t_n + \tau)$ para todo τ, n, t_1, \dots, t_n . El proceso se denomina *débilmente estacionario* si para todo τ los dos primeros momentos de la distribución coinciden.

La función valor medio de un proceso (débilmente) estacionario es constante. La función de covarianza cruzada de un proceso débilmente estacionario es función únicamente de la diferencia de los argumentos, $s - t$. Con abuso en la notación de funciones, escribimos

$$r_{xy}(s, t) = r_{xy}(s - t) \quad (6.10)$$

La función de covarianza cruzada de procesos (débilmente) estacionarios es una función de un único argumento. Por tanto,

$$r_{xy}(\tau) = \text{cov}[x(t + \tau), y(t)] \quad (6.11)$$

Cuando x es una escalar, la función

$$r_x(\tau) = r_{xx}(\tau) = \text{cov}[x(t + \tau), x(t)] \quad (6.12)$$

se denomina función de *autocovarianza*.

La *densidad espectral cruzada* de procesos (débilmente) estacionarios es la transformada de Fourier de su función de covarianza. Por tanto,

$$\phi_{xy}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r_{xy}(k) e^{-ik\omega} \quad (6.13)$$

y

$$r_{xy}(k) = \int_{-\pi}^{\pi} e^{ik\omega} \phi_{xy}(\omega) d\omega \quad (6.14)$$

Es costumbre referirse a ϕ_{xx} y ϕ_{xy} como la *densidad autoespectral* y la *densidad espectral cruzada*. La densidad autoespectral también se denomina, por simplicidad, *densidad espectral*.

Interpretación de la covarianza y la densidad espectral. Los procesos gaussianos estacionarios están caracterizados completamente por sus funciones de valor medio y sus funciones de covarianza. En las aplicaciones es útil tener una buena intuición para comprender cómo las propiedades de los procesos estocásticos están reflejadas en estas funciones.

La función valor medio es casi siempre autoexplicativa. El valor $r_x(0)$ de la función de covarianza en el origen es la varianza del proceso. Indica la magnitud de las fluctuaciones del proceso. La desviación estándar de las variaciones es igual a la

raíz cuadrada de $r_x(0)$. Si la función de covarianza está normalizada por $r_x(0)$, se obtiene la *función de correlación*, que se define por

$$\rho_x(\tau) = \frac{r_x(\tau)}{r_x(0)}$$

De la desigualdad de Schwartz se deduce que

$$|r_x(\tau)| \leq r_x(0)$$

Por tanto, el valor de la función de correlación es inferior a la unidad.

El valor $\rho_x(\tau)$ da la correlación entre los valores del proceso espaciados τ . Valores próximos a la unidad significan que hay una fuerte correlación, valores nulos indican que no existe correlación y valores negativos indican que la correlación es negativa. El estudio de las formas de la función de correlación indica la interdependencia temporal del proceso.

Es muy útil estudiar realizaciones de procesos estocásticos y sus funciones de covarianza para discernir entre sus relaciones. En la Figura 6.7 se muestran algunos ejemplos. Todos los procesos tienen varianza unidad.

La función densidad espectral también tiene una fácil interpretación física. La integral

$$2 \int_{\omega_1}^{\omega_2} \phi(\omega) d\omega \quad (6.15)$$

representa la potencia de la señal en la banda de frecuencia (ω_1, ω_2) . El área bajo la curva de densidad espectral representa la potencia de la señal en una cierta banda de frecuencia. La presencia de picos en el espectro indica que tiene componentes casi periódicos. El área total bajo la curva es proporcional a la varianza total de la señal. En los trabajos prácticos es interesante comprender cómo se relacionan las propiedades de las señales con el espectro (compárese con la Fig. 6.7).

Obsérvese que la función valor medio, la función covarianza y la densidad espectral se caracterizan únicamente por los dos primeros momentos de la distribución. Señales cuyas realizaciones son muy diferentes pueden tener iguales los primeros momentos. Una señal telegráfica aleatoria que comuta entre los valores 0 y 1 tiene el mismo espectro que el ruido de un simple circuito *RC*.

Ruido blanco discreto

A continuación se introduce un proceso aleatorio simple y útil. Consideremos como instantes de tiempo el conjunto de los números enteros. Supongamos un proceso x estocástico discreto y estacionario tal que $x(t)$ y $x(s)$ son independientes si $t \neq s$. Entonces, el proceso estocástico puede considerarse como una secuencia $\{x(t, \omega), t = \dots, -1, 0, 1, \dots\}$ de variables aleatorias independientes e igualmente distribuidas.

Función de covarianza $r_y(\tau)$ Espectro $|\Phi_y(\omega)|$ Realización $y(k)$

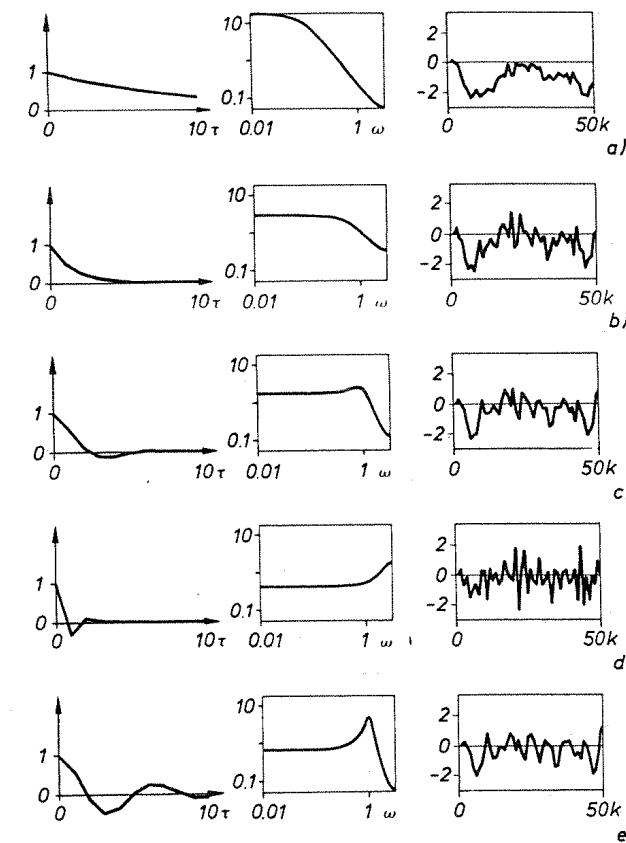


Figura 6.7 Funciones de covarianza, densidades espiracrales y funciones de muestra para algunos procesos aleatorios estacionarios. Todos los procesos tienen varianza unidad.

La función de covarianza es

$$r(\tau) = \begin{cases} \sigma^2 & \tau = 0 \\ 0 & \tau = \pm 1, \pm 2, \dots \end{cases}$$

Un proceso con esta función de covarianza se denomina *ruido blanco discreto*. De la ecuación (6.13) se deduce que su densidad espectral es

$$\phi(\omega) = \frac{\sigma^2}{2\pi}$$

La densidad espectral es, por tanto, constante para todas las frecuencias. La analogía con las propiedades de la luz blanca explica el nombre dado al proceso.

El ruido blanco juega un papel muy importante en la teoría de control estocástico. Todos los procesos estocásticos que se necesiten se generarán simplemente filtrando un ruido blanco. De este modo, el ruido blanco es equivalente a los pulsos de los sistemas determinísticos.

Procesos ARMA

Se pueden generar muchas clases de procesos estocásticos alimentando a los sistemas lineales con ruido blanco. Sea $\{e(k), k = \dots, -1, 0, 1, \dots\}$ un ruido blanco discreto. El proceso generado por

$$y(k) = e(k) + b_1e(k-1) + \dots + b_ne(k-n)$$

se denomina de *media móvil*, o proceso MA (del inglés «Moving Average»). El proceso generado por

$$y(k) + a_1y(k-1) + \dots + a_ny(k-n) = e(k)$$

se denomina *autoregresivo*, o proceso AR. El proceso

$$y(k) + a_1y(k-1) + \dots + a_ny(k-n) = e(k) + b_1e(k-1) + \dots + b_ne(k-n)$$

se denomina, por tanto, proceso ARMA.

Modelos en el espacio de estados

El concepto de estado tiene sus raíces en la relación causa-efecto de la mecánica clásica. El movimiento de un sistemas de partículas está determinado de modo único para todos los instantes futuros por los valores actuales de las posiciones y momento de las partículas y las fuerzas futuras. No importa cómo se alcanzaron las posiciones y los momentos actuales. El estado es una abstracción de esta propiedad; es mínima la información que se requiere de la historia de un sistema para predecir su futuro.

Para sistemas estocásticos no puede exigirse que los movimientos próximos estén determinados con precisión. Una extensión natural de la noción de estado para sistemas estocásticos consiste en requerir que la distribución de probabilidad de los estados futuros se determine de modo único por los estados actuales. Los procesos estocásticos que disponen de esta propiedad se denominan procesos de Markov. Los *procesos de Markov* son los procesos estocásticos equivalentes al modelo de espacio de estados. Formalmente se definen del siguiente modo.

Definición 6.1 – Procesos de Markov. Sean t_i y t elementos del conjunto de índices T tales que $t_1 < t_2 < \dots < t_n < t$. Un proceso estocástico $\{x(t), t \in T\}$ se denomina proceso de Markov si

$$P\{x(t) \leq \xi | x(t_1), \dots, x(t_n)\} = P\{x(t) \leq \xi | x(t_n)\}$$

donde $P\{\cdot | x(t_1), \dots, x(t_n)\}$ denota la probabilidad condicional dadas $x(t_1), \dots, x(t_n)$.

Un proceso de Markov está determinado completamente por la *distribución de probabilidad inicial*

$$F(\xi; t_0) = P\{x(t_0) \leq \xi\}$$

y la *distribución de probabilidad de transición*

$$F(\xi_s, t | \xi_s, s) = P\{x(t) \leq \xi_s | x(s) = \xi_s\}$$

Todas las distribuciones de dimensión finita se pueden generar a partir de esta distribución utilizando las reglas de multiplicación de probabilidades condicionales.

El proceso de Markov es el concepto natural que se utiliza cuando se extiende la noción del modelo de estados al caso estocástico.

Ecuaciones en diferencias lineales estocásticas. Consideremos un sistema discreto donde el período de muestreo se supone que es la unidad de tiempo. Supongamos que el estado en el instante k es $x(k)$. Entonces, la distribución de probabilidad del estado en el instante $k+1$ es una función de $x(k)$. Si el valor medio es lineal en $x(k)$ y la distribución alrededor del valor medio es independiente de $x(k)$, entonces $x(k+1)$ puede representarse como

$$x(k+1) = \Phi x(k) + v(k) \quad (6.16)$$

donde $v(k)$ es una variable aleatoria con valor medio nulo que es independiente de $x(k)$ e independiente de todos los valores pasados de x . Esto implica que $v(k)$ también es independiente de todos los valores pasados de v . La secuencia $\{v(k), k = \dots, -1, 0, 1, \dots\}$ es una secuencia de variables aleatorias independientes igualmente distribuidas. Por tanto, el proceso estocástico $\{v(k)\}$ es ruido blanco discreto.

La ecuación (6.16) se denomina *ecuación en diferencias estocástica lineal*. Para definir completamente el proceso aleatorio $\{x(k)\}$ es necesario especificar las condiciones iniciales. Se supone que el estado inicial tiene una media m_0 y una matriz de covarianza R_0 . La covarianza de la variable aleatoria v se denota por R_1 .

Propiedades de las ecuaciones en diferencias estocásticas lineales. Vamos a estudiar el carácter de los procesos aleatorios definidos por la ecuación en diferencias estocástica lineal (6.16) y a calcular el primer y segundo momentos del proceso.

Para obtener la función valor medio

$$m(k) = E x(k)$$

simplemente tomaremos el valor medio de ambos términos de (6.16). Ya que v tiene valor medio cero, se obtiene la siguiente ecuación en diferencias

$$m(k+1) = \Phi m(k) \quad (6.17)$$

La condición inicial es

$$m(0) = m_0$$

Así, el valor medio se propagará del mismo modo que en el sistema no perturbado.

Para calcular la función de covarianza, se introduce

$$P(k) = \text{cov}[x(k), x(k)] = E\tilde{x}(k)\tilde{x}^T(k)$$

donde

$$\tilde{x} = x - m$$

Se demuestra con las ecuaciones (6.16) y (6.17) que x satisface la ecuación (6.16) con unas condiciones iniciales que tengan un valor medio nulo. Para calcular la covarianza, se forma la expresión

$$\begin{aligned}\tilde{x}(k+1)\tilde{x}^T(k+1) &= \{\Phi\tilde{x}(k) + v(k)\}\{\Phi\tilde{x}(k) + v(k)\}^T \\ &= \Phi\tilde{x}(k)\tilde{x}^T(k)\Phi^T + \Phi\tilde{x}(k)v^T(k) + v(k)\tilde{x}^T(k)\Phi^T + v(k)v^T(k)\end{aligned}$$

Tomando el valor medio se obtiene

$$P(k+1) = \Phi P(k)\Phi^T + R_1$$

puesto que $v(k)$ y $\tilde{x}(k)$ son independientes. Las condiciones iniciales son

$$P(0) = R_0$$

La ecuación recursiva para P muestra cómo se propaga la covarianza.

Para calcular la función de covarianza del estado, debe observarse que

$$\tilde{x}(k+1)\tilde{x}^T(k) = [\Phi\tilde{x}(k) + v(k)]\tilde{x}^T(k)$$

Como $v(k)$ y $\tilde{x}(k)$ son independientes y $v(k)$ tiene valor medio cero,

$$r_{xx}(k+1, k) = \text{cov}[x(k+1), x(k)] = \Phi P(k)$$

Repetiendo este análisis,

$$r_{xx}(k+\tau, k) = \Phi P(k)$$

El resultado obtenido es tan importante que merece ser resumido.

Teorema 6.1. Considérese un proceso aleatorio definido por la ecuación en diferencias lineal estocástica (6.16), donde $\{v(k)\}$ es un proceso de ruido blanco con valor medio cero y covarianza R_1 . Supongamos que el estado inicial tiene valor medio m_0 y covarianza R_0 . La función valor medio del proceso está definida por

$$m(k+1) = \Phi m(k), \quad m(0) = m_0 \quad (6.18)$$

y la función de covarianza por

$$r(k+\tau, k) = \Phi P(k), \quad \tau \geq 0 \quad (6.19)$$

donde $P(k) = \text{cov}[x(k), x(k)]$ está dada por

$$P(k+1) = \Phi P(k)\Phi^T + R_1, \quad P(0) = R_0 \quad (6.20)$$

□

Nota 1. Si la variable aleatoria es gaussiana, el proceso estocástico está determinado de modo único por su función valor medio m y su función de covarianza r .

Nota 2. Si el sistema tiene una salida $y = Cx$, la función valor medio de y está definida por

$$m_y = Cm$$

y su varianza

$$r_{yy} = Cr_{xx}C^T$$

La covarianza cruzada entre y y x es

$$r_{yx} = Cr_{xx}$$

Nota 3. Obsérvese que la ecuación en diferencias (6.20) para la matriz P es similar a la ecuación (5.14), la cual se utilizó para calcular las funciones de Lyapunov en el Capítulo 5.

Nota 4. Los diferentes términos de (6.20) tienen una interpretación física perfecta. La covarianza P puede representar la incertidumbre del estado; el término $\Phi P(k)\Phi^T$ muestra cómo se propaga la incertidumbre en los instantes k debido a la dinámica del sistema, y el término R_1 describe el incremento de la incertidumbre debido a la perturbación v .

Ejemplo 6.3

Considérese el sistema de primer orden

$$x(k+1) = ax(k) + v(k)$$

donde v es una secuencia de variables aleatorias no correlacionadas con valor medio cero y covarianza r_1 . Supongamos que el estado en el instante k_0 tiene media m_0 y covarianza r_0 . Se demuestra de (6.18) que el valor medio

$$m(k) = Ex(k)$$

viene dado por

$$m(k+1) = am(k), \quad m(k_0) = m_0$$

Por tanto,

$$m(k) = a^{k-k_0}m_0$$

La ecuación (6.20) proporciona

$$P(k+1) = a^2 P(k) + r_1, \quad P(k_0) = r_0$$

Por tanto,

$$P(k) = a^{2(k-k_0)}r_0 + \frac{1 - a^{2(k-k_0)}}{1 - a^2}r_1$$

Además

$$r_x(l, k) = a^{l-k}P(k), \quad l \geq k$$

y

$$r_x(l, k) = a^{k-l} P(l), \quad l < k$$

Si $|a| < 1$ y $k_0 \rightarrow -\infty$, se deduce que

$$\begin{aligned} m(k) &\rightarrow 0 \\ P(k) &\rightarrow \frac{r_1}{1-a^2} \\ r_x(k+\tau, k) &\rightarrow \frac{r_1 a^{|\tau|}}{1-a^2} \end{aligned}$$

El proceso llega a ser estacionario porque m es constante y la función de covarianza sólo es función de τ .

Si se incluye una salida

$$y(k) = x(k) + e(k)$$

donde e es una secuencia de variables aleatorias no correlacionadas con valor medio cero y covarianza r_e , se demuestra que la función de covarianza de y es

$$r_y(\tau) = \begin{cases} r_2 + \frac{r_1}{1-a^2} & \tau = 0 \\ \frac{r_1 a^{|\tau|}}{1-a^2} & \tau \neq 0 \end{cases}$$

La densidad espectral se obtiene de (6.13). Por tanto,

$$\begin{aligned} \phi_y(\omega) &= \frac{1}{2\pi} \left[r_2 + \frac{r_1}{(e^{i\omega}-a)(e^{-i\omega}-a)} \right] \\ &= \frac{1}{2\pi} \left[r_2 + \frac{r_1}{1+a^2-2a \cos \omega} \right] \end{aligned}$$

Modelos de entrada-salida

Como aclaración adicional se da una descripción de entrada-salida de señales generadas por ecuaciones en diferencias lineales. Obsérvese que la señal x dada por (6.16) puede describirse como la salida de un sistema dinámico lineal alimentado con ruido blanco. Desde este punto de vista es normal estudiar cómo cambian las propiedades de los procesos estocásticos cuando se filtran mediante sistemas dinámicos.

Análisis. Considérese un sistema como el mostrado en la Figura 6.8. Por simplicidad se supone que el período de muestreo es la unidad. Supongamos que la entrada u es un proceso estocástico con una función valor medio m_u y una función de covarianza r_u . Consideremos como respuesta impulsional del sistema a $\{h(k), k=0, 1, \dots\}$. Obsérvese que h también se utiliza para designar el período de muestreo. Sin embargo, del contexto se infiere claramente cuál debe ser h . La relación entrada-salida es

$$y(k) = \sum_{n=-\infty}^k h(k-n)u(n) = \sum_{n=0}^{\infty} h(n)u(k-n) \quad (6.21)$$

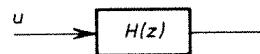


Figura 6.8 Generación de perturbaciones por la introducción de un ruido blanco en un sistema dinámico.

Tomando valores medios

$$\begin{aligned} m_y(k) &= E[y(k)] = E \sum_{n=0}^{\infty} h(n)u(k-n) \\ &= \sum_{n=0}^{\infty} h(n)E[u(k-n)] = \sum_{n=0}^{\infty} h(n)m_u(k-n) \end{aligned} \quad (6.22)$$

El valor medio de la salida se obtiene enviando el valor medio de la entrada a través del sistema.

Para determinar la covarianza, debe observarse en primer lugar que restando (6.22) de (6.21) se obtiene

$$y(k) - m_y(k) = \sum_{n=0}^{\infty} h(n)[u(k-n) - m_u(k-n)]$$

De este modo, la diferencia entre la señal de entrada y su valor medio se propaga a través del sistema del mismo modo que la propia señal de entrada. Cuando se calcula la covarianza, puede suponerse que el valor medio es cero. Esto simplifica la escritura.

La definición de la función de covarianza da

$$\begin{aligned} r_y(\tau) &= E[y(k+\tau)y^T(k)] \\ &= E \sum_{n=0}^{\infty} h(n)u(k+\tau-n) \left[\sum_{l=0}^{\infty} h(l)u(k-l) \right]^T \\ &= \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} h(n)E[u(k+\tau-n)u^T(k-l)]h^T(l) \\ &= \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} h(n)r_u(\tau+l-n)h^T(l) \end{aligned} \quad (6.23)$$

Un cálculo similar proporciona la siguiente fórmula para la covarianza cruzada entre la entrada y la salida:

$$\begin{aligned} r_{yu}(\tau) &= E[y(k+\tau)u^T(k)] \\ &= E \sum_{n=0}^{\infty} h(n)u(k+\tau-n)u^T(k) \\ &= \sum_{n=0}^{\infty} h(n)E[u(k+\tau-n)u^T(k)] \\ &= \sum_{n=0}^{\infty} h(n)r_u(\tau-n) \end{aligned} \quad (6.24)$$

Obsérvese que se ha supuesto que existen todas las sumas infinitas y que la suma infinita y las operaciones matemáticas se han intercambiado libremente en estos cálculos. Por supuesto, esto deberá justificarse; lo cual es fácil de realizar en el sentido

de que converja la media de los cuadrados, si se supone que el cuarto momento de la señal de entrada es finito.

Las relaciones expresadas en las ecuaciones (6.23) y (6.24) pueden expresarse de una forma muy simple si se introduce la densidad espectral.

La definición de densidad espectral en (6.13) da

$$\phi_y(\omega) = \phi_{yy}(\omega) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{-in\omega} r_y(n)$$

Introduciendo r_y se obtiene de (6.23) que

$$\begin{aligned}\phi_y(\omega) &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{-in\omega} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} h(k) r_u(n+l-k) h^T(l) \\ &= \frac{1}{2\pi} \sum_{k=0}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{l=0}^{\infty} e^{-ik\omega} h(k) e^{-i(n+l-k)\omega} r_u(n+l-k) e^{il\omega} h^T(l) \\ &= \frac{1}{2\pi} \sum_{k=0}^{\infty} e^{-ik\omega} h(k) \sum_{n=-\infty}^{\infty} e^{-in\omega} r_u(n) \sum_{l=0}^{\infty} e^{il\omega} h^T(l)\end{aligned}$$

Introduzcamos la función de transferencia discreta H del sistema. Está relacionada con la respuesta impulsional h por

$$H(z) = \sum_{n=0}^{\infty} z^{-n} h(n) \quad (6.25)$$

La ecuación para la densidad espectral puede escribirse como

$$\phi_y(\omega) = H(e^{i\omega}) \phi_u(\omega) H^T(e^{-i\omega})$$

Analógamente,

$$\begin{aligned}\phi_{yu}(\omega) &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{-in\omega} r_{yu}(n) \\ &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{-in\omega} \sum_{k=0}^{\infty} h(k) r_u(n-k) \\ &= \frac{1}{2\pi} \sum_{k=0}^{\infty} e^{-ik\omega} h(k) \sum_{n=-\infty}^{\infty} e^{-in\omega} r_u(n) \\ &= H(e^{i\omega}) \phi_u(\omega)\end{aligned}$$

Resultado principal. Para obtener el resultado general, debe estudiarse la propagación del valor medio a través del sistema.

Teorema 6.2 – Filtrado de procesos estacionarios. Consideremos un sistema dinámico discreto estacionario con período de muestreo 1 y función de transferencia discreta H . Consideremos la señal de entrada como un proceso estocástico estacionario de valor medio m_u y densidad espectral Φ_u . Si el sistema es estable, entonces la salida es también un proceso estacionario de valor medio

$$m_y = H(1)m_u \quad (6.26)$$

y densidad espectral

$$\phi_y(\omega) = H(e^{i\omega}) \phi_u(\omega) H^T(e^{-i\omega}) \quad (6.27)$$

La densidad espectral cruzada entre la entrada y la salida viene dada por

$$\phi_{yu}(\omega) = H(e^{i\omega}) \phi_u(\omega) \quad (6.28) \quad \square$$

Nota 1. El resultado tiene una sencilla interpretación física. El valor $|H(e^{i\omega})|$ es la amplitud en estado estacionario de la respuesta del sistema a una señal seno de frecuencia ω . Entonces, el valor de la densidad espectral de la salida es el producto del cuadrado de la ganancia $|H(e^{i\omega})|^2$ y la densidad espectral de la entrada $\phi_u(\omega)$.

Nota 2. De la ecuación (6.28) se deduce que la densidad espectral cruzada es igual a la función de transferencia del sistema si la entrada es un ruido blanco con densidad espectral unidad. Esta propiedad puede utilizarse para determinar la función de transferencia discreta de un sistema.

El resultado se ilustra con un ejemplo.

Ejemplo 6.4

Considérese el proceso $\{x(k)\}$ del Ejemplo 6.3. Desde el punto de vista entrada-salida, puede considerarse que el proceso está generado por una señal de ruido blanco que alimenta a un filtro con función de transferencia discreta

$$H(z) = \frac{1}{z - a}$$

Puesto que la densidad espectral de $\{v(k)\}$ es

$$\phi_v(\omega) = \frac{r_1}{2\pi}$$

se demuestra de (6.27) que la densidad espectral de $\{x(k)\}$ es

$$\begin{aligned}\phi_x(\omega) &= H(e^{i\omega}) H(e^{-i\omega}) \frac{r_1}{2\pi} \\ &= \frac{r_1}{2\pi} \cdot \frac{1}{(e^{i\omega} - a)(e^{-i\omega} - a)} = \frac{r_1}{2\pi(1 + a^2 - 2a \cos \omega)}\end{aligned}$$

Entonces, el proceso

$$y(k) = x(k) + v(k)$$

tiene la densidad espectral

$$\phi_y(\omega) = \frac{1}{2\pi} \left[r_2 + \frac{r_1}{1 + a^2 - 2a \cos \omega} \right]$$

(Compárese con el cálculo del Ejemplo 6.3.)



Factorización espectral

El Teorema 6.2 da la densidad espectral de un proceso estocástico que se obtiene filtrando otro proceso estocástico. La densidad espectral de una señal obtenida filtrando ruido blanco se obtiene como un caso especial. El problema inverso se discute a continuación. Se determinará un sistema lineal que proporciona una salida con una densidad espectral determinada cuando se le alimenta con ruido blanco. Este problema es importante porque muestra cómo puede generarse una señal con densidad espectral determinada filtrando ruido blanco. La solución al problema también mostrará la generalidad del modelo (6.16). Con el Teorema 6.2 se demuestra que el proceso aleatorio generado por un sistema lineal para una entrada con ruido blanco tiene la densidad espectral dada por (6.27). Si el sistema tiene dimensión finita, entonces H es una función racional en $\exp(i\omega)$ y la densidad espectral ϕ también será racional en $\exp(i\omega)$ o lo que es equivalente en $\cos \omega$. Por abuso del lenguaje a esta densidad espectral se le denomina *racional*. Introduciendo

$$z = e^{i\omega}$$

el término de la derecha de (6.27) se puede escribir como

$$F(z) = H(z)H^T(z^{-1})$$

Si z_i es un cero de $H(z)$, entonces z_i^{-1} es un cero de $H(z^{-1})$. Los ceros de la función F son simétricos respecto al eje real y al círculo unidad. Si los coeficientes de la función racional H son reales, los ceros de la función F también serán simétricos respecto al eje real. El mismo argumento sirve para los polos de H . De este modo los polos y ceros de F tendrán la estructura mostrada en la Figura 6.9.

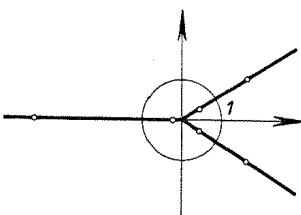


Figura 6.9 Simetría de los polos y ceros de la función de densidad espectral.

Ahora resulta inmediato encontrar una función H que corresponda a una densidad espectral racional dada: en primer lugar, se determinan los polos p_i y los ceros z_i de la función F asociada a la densidad espectral. Se demuestra de la simetría de los polos y ceros, lo cual ya se ha establecido, que los polos y ceros siempre aparecen en pares de modo que

$$z_i z_j = 1$$

$$p_i p_j = 1$$

En cada par se elige el polo o el cero de magnitud menor o igual a la unidad; entonces se forma la función de transferencia deseada a partir de los polos y ceros elegidos en la forma

$$H(z) = K \frac{\prod (z - z_i)}{\prod (z - p_i)} = \frac{B(z)}{A(z)}$$

Puesto que el proceso estocástico es estacionario, los polos p_i elegidos serán estrictamente menores que la unidad. Sin embargo, pueden existir ceros con magnitud unidad. El resultado se resume en el siguiente teorema.

Teorema 6.3 – Teorema de factorización espectral. Dada una densidad espectral $\phi(\omega)$, que es racional en $\cos \omega$, existe un sistema lineal con función de transferencia discreta

$$H(z) = \frac{B(z)}{A(z)} \quad (6.29)$$

tal que la salida obtenida cuando el sistema se alimenta con ruido blanco es un proceso aleatorio estacionario con densidad espectral ϕ . El polinomio $A(z)$ tiene todos sus ceros dentro del círculo unidad. El polinomio $B(z)$ tiene todos sus ceros dentro o sobre la frontera del círculo unidad. \square

Nota 1. El teorema de factorización espectral es muy importante. Implica que todos los procesos aleatorios estacionarios pueden considerarse generados con un sistema lineal estable alimentado con ruido blanco; es decir, un tipo especial de proceso ARMA. Esto supone una simplificación considerable tanto en la teoría como en la práctica. Es suficiente comprender cómo se comporta un sistema cuando se excita con ruido blanco, para lo cual es necesario poder simular ruido blanco. El resto de procesos estacionarios con densidad espectral racional se pueden formar mediante filtrado.

Nota 2. Puesto que una función continua se puede aproximar uniformemente de forma arbitraria, en un intervalo compacto, por una función racional, se demuestra que los modelos (6.16) y (6.21) pueden dar señales cuyos espectros se aproximan a funciones continuas. Obsérvese, sin embargo, que hay modelos con densidades espectrales no racionales. En la teoría de turbulencias, por ejemplo, hay densidades espectrales que decaen según potencias fraccionarias de ω para ω elevados.

Una consecuencia importante del teorema de factorización espectral es que para sistemas con una salida, siempre es posible representar los efectos de todas las perturbaciones de la red mediante *una* perturbación equivalente. Esta perturbación se obtiene calculando la densidad espectral total de la señal de salida y aplicando el teorema de factorización espectral.

Nota 3. Con frecuencia se supone que el polinomio $B(z)$ tiene todos sus ceros dentro del círculo unidad. Esto significa que la inversa del sistema H es estable.

Los resultados se ilustran con un ejemplo.

Ejemplo 6.5

Considérese el proceso $\{y(k)\}$ de los Ejemplos 6.3 y 6.4. Este proceso tiene la densidad espectral

$$\begin{aligned}\phi_y(\omega) &= \frac{1}{2\pi} \left[r_2 + \frac{r_1}{(z-a)(z^{-1}-a)} \right]_{z=e^{i\omega}} \\ &= \frac{1}{2\pi} \left[\frac{r_1 + r_2(1+a^2) - r_2a(z+z^{-1})}{(z-a)(z^{-1}-a)} \right]_{z=e^{i\omega}}\end{aligned}$$

El denominador está expresado en forma factorial. Para factorizar el numerador, observemos que puede escribirse como

$$\lambda^2(z-b)(z^{-1}-b) \equiv r_1 + r_2(1+a^2) - r_2a(z+z^{-1})$$

Identificando los coeficientes de las potencias iguales en z , tenemos

$$z^0: \quad \lambda^2(1+b^2) = r_1 + r_2(1+a^2)$$

$$z^1: \quad \lambda^2 b = r_2 a$$

Eliminando λ se obtiene una ecuación algebraica de segundo orden para b . Esta ecuación tiene la solución

$$b = \frac{r_1 + r_2(1+a^2) - \sqrt{[r_1 + r_2(1+a^2)][r_1 + r_2(1-a^2)]}}{2ar_2}$$

Las demás raíces se descartan por estar fuera del círculo unidad. Por tanto, la variable λ está definida por

$$\lambda^2 = \frac{1}{2} [r_1 + r_2(1+a^2) + \sqrt{[r_1 + r_2(1+a^2)][r_1 + r_2(1-a^2)]}]$$

Representaciones de las innovaciones

El Teorema 6.3 tiene algunas consecuencias conceptuales importantes. Se deduce del teorema que un proceso con densidad espectral racional puede representarse como

$$y(k) = \sum_{n=-\infty}^k h(k-n)e(n) \quad (6.30)$$

donde e es un ruido blanco discreto y h es la respuesta impulsional correspondiente a la función de transferencia discreta (6.29). El sistema tiene una inversa estable si el polinomio $B(z)$ tiene todos los ceros dentro del círculo unidad. Esto significa que

$$e(k) = \sum_{n=-\infty}^k g(k-n)y(n)$$

donde g es la respuesta impulsional, correspondiente a la función de transferencia estable $A(z)/B(z)$. De esto se deduce que las secuencias $y(k), y(k-1), \dots$ y $e(k), e(k-1), \dots$ son equivalentes en el sentido de que una secuencia se puede calcular a partir de la otra.

Consideremos ahora

$$\begin{aligned}y(k+1) &= \sum_{n=-\infty}^{k+1} h(k+1-n)e(n) \\ &= \sum_{n=-\infty}^k h(k+1-n)e(n) + h(0)e(k+1) \\ &= \sum_{n=-\infty}^k h(k+1-n) \sum_{l=-\infty}^k g(n-l)y(l) + h(0)e(k+1)\end{aligned}$$

La variable $y(k+1)$ puede escribirse como suma de dos términos: un término es una función lineal de $y(k), y(k-1), \dots$, y el otro término es $h(0)e(k+1)$. De este modo, $e(k+1)$ se puede interpretar como la parte de $y(k+1)$ que contiene nueva información que no estaba presente en los valores pasados $y(k), y(k-1), \dots$. Por tanto, al proceso estocástico $\{e(k)\}$ se le denomina las *innovaciones* de los procesos $\{y(k), k \in T\}$ y la representación (6.30) se denomina *representación de las innovaciones* del proceso. Esta representación es muy importante en relación con los problemas de filtrado y predicción.

El término

$$\sum_{n=-\infty}^k h(k+1-n) \sum_{l=-\infty}^n g(n-l)y(l)$$

es de hecho la mejor predicción de la media de los cuadrados de $y(k+1)$ basada en $y(k), y(k-1), \dots$. Esto se expondrá con detalle en los Capítulos 11 y 12.

Ejemplo 6.6

Consideremos el proceso $\{y(k)\}$ del Ejemplo 6.3. El proceso tiene una densidad espectral

$$\phi_y(\omega) = \frac{1}{2\pi} \left[r_2 + \frac{r_1}{1+a^2 - 2a \cos \omega} \right]$$

Se demuestra del Ejemplo 6.5 que la densidad espectral se puede factorizar como

$$\phi_y(\omega) = \frac{\lambda^2 (z-b)(z^{-1}-b)}{2\pi (z-a)(z^{-1}-a)}$$

De este modo, el proceso y se puede generar enviando un ruido blanco a través de un sistema con función de transferencia discreta

$$H(z) = \frac{z-b}{z-a}$$

La relación entrada-salida de este sistema se puede escribir como

$$y(k+1) = ay(k) + e(k+1) - be(k)$$

donde $\{e(k)\}$ es un ruido blanco con varianza λ^2 .

Cálculo de varianzas

La varianza de una señal obtenida filtrando un ruido blanco se puede calcular con la ecuación recursiva (6.20), si el modelo está dado en la representación del

espacio de estados. Para un sistema descrito por su función de transferencia es posible utilizar la misma ecuación, si el modelo se transforma previamente a su representación en el espacio de estados. Naturalmente, es conveniente tener expresiones similares cuando el sistema está dado en su representación de entrada-salida. A continuación se presentan estas expresiones.

Considérese una señal generada por

$$y(k) = \frac{B(q)}{A(q)} e(k) \quad (6.31)$$

donde e es un ruido blanco con varianza unidad. Se demuestra del Teorema 6.2 que la densidad espectral de la señal y está dada por

$$\phi(\omega) = \frac{1}{2\pi} \frac{B(z)B(z^{-1})}{A(z)A(z^{-1})}$$

donde $z = \exp(i\omega)$. También se obtiene del Teorema 6.2 que la varianza de la señal y está definida por la integral compleja

$$\begin{aligned} E[y^2] &= \int_{-\pi}^{\pi} \phi(\omega) d\omega \\ &= \frac{1}{i} \int_{-\pi}^{\pi} \phi(\omega) e^{-i\omega} d(e^{i\omega}) \\ &= \frac{1}{2\pi i} \oint \frac{B(z)B(z^{-1})}{A(z)A(z^{-1})} \frac{dz}{z} \end{aligned} \quad (6.32)$$

La evaluación de integrales de esta forma está relacionada con el criterio de estabilidad de Jury (compárese con Sec. 5.2). Para evaluar la integral, se forma la siguiente tabla.

a_0	a_1	\dots	a_{n-1}	a_n	α_n	b_0	b_1	\dots	b_{n-1}	b_n	β_n
a_n	a_{n-1}	\dots	a_1	a_0	α_n	a_n	a_{n-1}	\dots	a_1	a_0	β_n
a_0^{n-1}	$-a_1^{n-1}$	\dots	a_{n-1}^{n-1}			b_0^{n-1}	b_1^{n-1}	\dots	b_{n-1}^{n-1}		
a_{n-1}^{n-1}	a_{n-2}^{n-1}	\dots	a_0^{n-1}		α_{n-1}	a_{n-1}^{n-1}	a_{n-2}^{n-1}	\dots	a_0^{n-1}		β_{n-1}
\vdots											
a_0	a_1			b_0	b_1						β_1
a_1	a_0			α_1	a_1	a_0					
a_0				1		b_0					β_0

donde

$$\begin{aligned} \alpha_n &= a_n/a_0 & \beta_n &= b_n/a_0 \\ \alpha_k &= a_k/a_0 & \beta_k &= b_k/a_0 \end{aligned}$$

y

$$\begin{aligned} a_i^{k-1} &= a_i^k - \alpha_k a_{k-i}^k \\ b_i^{k-1} &= b_i^k - \beta_k a_{k-i}^k \end{aligned}$$

Como resultado se tiene el siguiente teorema.

Teorema 6.4. La integral (6.32) está definida por

$$I_n = \frac{1}{a_0} \sum_{i=0}^n b_i^i \beta_i \quad (6.33) \quad \square$$

La aplicación del teorema da los siguientes valores de la integral para $n=1$ y 2 .

$$\begin{aligned} I_1 &= \frac{(b_0^2 + b_1^2)a_0 - 2b_0 b_1 a_1}{a_0(a_0^2 - a_1^2)} \\ I_2 &= \frac{B_0 a_0 e_1 - B_1 a_0 a_1 + B_2 (a_1^2 - a_2 e_1)}{a_0[(a_0^2 - a_2^2)e_1 - (a_0 a_1 - a_1 a_2)a_1]} \end{aligned}$$

donde

$$\begin{aligned} B_0 &= b_0^2 + b_1^2 + b_2^2 \\ B_1 &= 2(b_0 b_1 + b_1 b_2) \\ B_2 &= 2b_0 b_2 \\ e_1 &= a_0 + a_2 \end{aligned}$$

6.6. Procesos estocásticos continuos

Puede resultar útil formular especificaciones y modelos continuos incluso cuando se utiliza un computador para realizar la ley de control. Por tanto, se presenta un breve resumen de los procesos estocásticos continuos.

Definiciones

Los procesos estocásticos continuos se pueden definir del mismo modo que los procesos estocásticos discretos. La única diferencia está en que el conjunto T de índices es el conjunto de los números reales en lugar de un conjunto discreto. Las funciones de covarianza y los procesos estacionarios se definen para un proceso discreto utilizando las funciones de distribución de dimensión finita. También se puede introducir una densidad espectral como la transformada de Fourier de la función de covarianza. Entonces, la ecuación (6.13) se reemplaza por

$$\phi_{xy}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} r_{xy}(t) dt \quad (6.34)$$

La transformación inversa es

$$r_{xx}(t) = \int_{-\infty}^{\infty} e^{i\omega t} \phi_{xx}(\omega) d\omega \quad (6.35)$$

la cual reemplaza a (6.14). La densidad espectral tiene la misma interpretación que en los sistemas discretos.

Ruido blanco

El ruido blanco se define como un proceso estacionario con densidad espectral constante. Si

$$\phi(\omega) = \frac{r_0}{2\pi} \quad (6.36)$$

se demuestra formalmente de (6.35) que la covarianza correspondiente es una función delta; es decir,

$$r(t) = r_0 \delta(t) \quad (6.37)$$

El ruido blanco continuo tiene la propiedad de que los valores de la señal en diferentes instantes no están correlacionados como en el ruido blanco discreto. Sin embargo, el ruido blanco continuo tiene varianza infinita. Esto causa algunas dificultades matemáticas. Intuitivamente, el ruido blanco continuo es análogo a las funciones delta en la teoría de sistemas lineales.

Se pueden evitar algunas de las dificultades del ruido blanco continuo introduciendo un proceso estocástico que formalmente es la integral en el tiempo

$$w(t) = \int_0^t e(s) ds \quad (6.38)$$

del ruido blanco e . El proceso estocástico w tiene valor medio cero. Sus incrementos sobre intervalos disjuntos no están correlacionados. Si la función de covarianza de e es

$$\text{cov}[e(t), e(s)] = r_0 \delta(t - s)$$

entonces las varianzas de los incrementos de w es

$$E[w(t) - w(s)]^2 = |t - s| \cdot r_0$$

El proceso estocástico $\{w(t), t \in T\}$ se denomina *proceso de Wiener* si es también gaussiano. El proceso de Wiener es un modelo para comportamientos aleatorios. El incremento infinitesimal

$$dw = w(t + dt) - w(t)$$

tiene de varianza

$$E(dw)^2 = r_0 dt$$

El incremento dw tiene el valor $\sqrt{r_0 dt}$ en el sentido de la media de los cuadrados. Al número $r_0 dt$ se le denomina *covarianza incremental* del proceso de Wiener.

Modelos de espacio de estados

Los modelos de estados para procesos continuos pueden obtenerse mediante la generalización formal de (6.16) a

$$\frac{dx}{dt} = Ax + \dot{v}$$

donde \dot{v} es un vector cuyos elementos son procesos estocásticos de ruido blanco. Como \dot{v} tiene varianza infinita, se acostumbra a escribir la ecuación en términos diferenciales como

$$dx = Ax dt + dv \quad (6.39)$$

donde v es la integral de \dot{v} .

Se supone que la señal v tiene valor medio cero, incrementos no correlacionados y la varianza

$$\text{cov}[v(t), v(t)] = R_1 t \quad (6.40)$$

También se supone que dv no está correlacionada con x .

Se puede dar un significado preciso a (6.39) sin hacer referencia al ruido blanco. Este método es común en los textos con orientación matemática. También es útil para recordar que dv es un valor proporcional a \sqrt{dt} .

La ecuación (6.39) se denomina *ecuación diferencial estocástica*. Para especificarla completamente, es necesario conocer la distribución de la probabilidad de x en el instante inicial. A continuación se obtiene un teorema para procesos continuos análogo al Teorema 6.1.

Teorema 6.5. Considérese un proceso estocástico definido por la ecuación diferencial lineal estocástica (6.39) donde el proceso v tiene media nula y covarianza incremental $R_1 dt$. El estado inicial tiene media m_0 y covarianza R_0 . Entonces la función valor medio del proceso x es

$$\frac{dm(t)}{dt} = Am(t), \quad m(0) = m_0 \quad (6.41)$$

y la función de covarianza es

$$\text{cov}[x(s), x(t)] = e^{A(t-s)} P(t), \quad s \geq t \quad (6.42)$$

donde $P(t) = \text{cov}[x(t), x(t)]$ es

$$\frac{dP(t)}{dt} = AP(t) + P(t)A^T + R_1, \quad P(0) = R_0 \quad (6.43)$$

Demostración. La expresión (6.41) para el valor medio se obtiene simplemente tomando el valor medio de (6.39). Obsérvese que dv tiene valor medio cero.



Para obtener la ecuación diferencial (6.43), obsérvese que

$$\begin{aligned} d(xx^T) &= (x + dx)(x + dx)^T - xx^T \\ &= x dx^T + dx x^T + dx dx^T \end{aligned}$$

Entonces, la ecuación (6.39) da

$$d(xx^T) = x[Ax dt + dv]^T + [Ax dt + dv]x^T + [Ax dt + dv][Ax dt + dv]^T$$

Tomando valores medios

$$\begin{aligned} d(Exx^T) &= (Exx^T)A^T dt + A(Exx^T) dt + E dv dv^T \\ &\quad + A(Exx^T)A^T(dt)^2 \end{aligned}$$

puesto que dv no está correlacionada con x . Además, se deduce de (6.40) que

$$E dv dv^T = R_1 dt$$

Por tanto,

$$dP = PA^T dt + AP dt + R_1 dt + APA^T(dt)^2$$

Dividiendo por dt y calculando el límite cuando dt tiende a cero se obtiene la ecuación diferencial (6.43). Para obtener la ecuación (6.42), consideremos $s \geq t$ e integremos (6.39). Por tanto,

$$x(s) = e^{A(s-t)}x(t) + \int_t^s e^{A(s-s')} dv(s')$$

Multiplicando por la derecha por $x^T(t)$ y operando se obtiene (6.42). Obsérvese que $dv(s')$ no está correlacionada con $x(t)$ si $s' \geq t$. \square

Ejemplo 6.7

Considérese la ecuación diferencial estocástica escalar

$$dx = -ax dt + dv$$

$$x(t_0) = m_0, \quad \text{var}[x(t_0)] = r_0$$

donde el proceso $\{v(t), t \in T\}$ tiene la covarianza incremental $r_1 dt$. Se deduce de (6.41) que la función valor medio es

$$\frac{dm}{dt} = -am, \quad m(t_0) = m_0$$

Esta ecuación tiene la solución

$$m(t) = m_0 e^{-a(t-t_0)}$$

La función de covarianza es

$$r(s, t) = \text{cov}[x(s), x(t)] = e^{-a(t-s)}P(t), \quad s \geq t$$

y

$$r(s, t) = e^{-a(t-s)}P(s), \quad s \leq t$$

La ecuación (6.43) proporciona la siguiente ecuación diferencial para P

$$\frac{dP}{dt} = -2aP + r_1, \quad P(t_0) = r_0$$

Esta ecuación diferencial tiene la solución

$$\begin{aligned} P(t) &= e^{-2a(t-t_0)}r_0 + \int_{t_0}^t e^{-2a(t-s)}r_1 ds \\ &= e^{-2a(t-t_0)}r_0 + \frac{r_1}{2a}[1 - e^{-2a(t-t_0)}] \end{aligned}$$

Cuando $t_0 \rightarrow -\infty$, la función valor medio se anula y la función de covarianza tiende a

$$r(s, t) = \frac{r_1}{2a}e^{-a|t-s|}$$

Ya que la función límite de covarianza depende únicamente de la diferencia de argumentos $s-t$, el proceso límite es (débilmente) estacionario y su función de covarianza puede escribirse como

$$r(\tau) = \frac{r_1}{2a}e^{-a|\tau|}$$

La ecuación (6.34) proporciona la correspondiente densidad espectral

$$\phi(\omega) = \frac{r_1}{2\pi} \frac{1}{\omega^2 + a^2}$$

Filtrado de procesos continuos

El análisis de sistemas lineales cuyas entradas son procesos estocásticos continuos es análogo al correspondiente análisis de los sistemas discretos. Consideremos un sistema estable invariante en el tiempo con respuesta impulsional h . La relación entre entrada y salida es

$$y(t) = \int_{-\infty}^t h(t-s)u(s) ds = \int_0^\infty h(s)u(t-s) ds \quad (6.44)$$

[compárese con la ecuación (6.21)]. Sea la señal de entrada u un proceso estocástico con una función valor medio m_u y una función de covarianza r_u .

El siguiente resultado es análogo al Teorema 6.2 para sistemas discretos.

Teorema 6.6 – Filtrado de procesos estacionarios. Considérese un sistema lineal estacionario con función de transferencia G . Sea la señal de entrada un proceso estocástico continuo y estacionario con valor medio m_u y densidad espectral ϕ_u . Si el sistema es estable, entonces la salida es también un proceso estacionario de valor medio

$$m_y = G(0)m_u \quad (6.45)$$

y densidad espectral

$$\phi_y(\omega) = G(i\omega)\phi_u(\omega)G^*(-i\omega) \quad (6.46)$$



La densidad espectral cruzada entre la entrada y la salida es

$$\phi_{yu}(\omega) = G(i\omega)\phi_u(\omega) \quad (6.47)$$

□

El resultado se puede interpretar del mismo modo que el resultado correspondiente para sistemas discretos. Compárese con las notas 1 y 2 del Teorema 6.2.

Ejemplo 6.8

Supongamos el sistema del Ejemplo 6.7. El proceso x se puede considerar como el resultado obtenido al filtrar ruido blanco con varianza $r_1/2\pi$ a través de un sistema de función de transferencia

$$G(s) = \frac{1}{s+a}$$

Se deduce de (6.46) que la densidad espectral es

$$\phi(\omega) = \frac{r_1}{2\pi} \cdot \frac{1}{i\omega + a} \cdot \frac{1}{-i\omega + a} = \frac{r_1}{2\pi} \cdot \frac{1}{\omega^2 + a^2}$$

Factorización espectral

Se deduce de (6.46) que si la entrada es ruido blanco con $\phi_u = 1$, entonces la densidad espectral de la salida es

$$\phi_y(\omega) = G(i\omega)G^T(-i\omega) \quad (6.48)$$

Esto significa que cualquier perturbación cuya densidad espectral se pueda escribir en esta forma, se puede generar enviando ruido blanco a través de un filtro de función de transferencia G .

Puesto que los sistemas lineales de dimensión finita tienen funciones de transferencia racionales, se deduce que las señales con densidades espectrales racionales arbitrarias pueden generarse con sistemas lineales de dimensión finita. La función de covarianza es no negativa y simétrica. Entonces se deduce de (6.34) que ϕ también es simétrica. Si ϕ es racional se deduce que sus polos y ceros son simétricos respecto a los ejes real e imaginario. La función de transferencia G de (6.48) puede elegirse de modo que todos los polos están en el semiplano izquierdo y todos sus ceros en el semiplano izquierdo o sobre el eje imaginario. De este modo se obtiene el teorema análogo al Teorema 6.3.

Teorema 6.7 – Factorización espectral. Dada una densidad espectral racional $\phi(\omega)$, existe un sistema lineal de dimensión finita con función de transferencia racional

$$G(s) = \frac{B(s)}{A(s)}$$

tal que la salida que se obtiene cuando el sistema se alimenta con ruido blanco es un proceso estocástico estacionario con la densidad espectral dada. El polinomio A tiene todos los ceros en el semiplano izquierdo. El polinomio B no tiene ceros en el semiplano derecho. □

6.7. Muestreo de una ecuación diferencial estocástica

Si los modelos de los procesos continuos se presentan como ecuaciones diferenciales estocásticas, resulta útil muestrear estas ecuaciones para obtener modelos discretos.

Consideremos el proceso descrito por

$$dx = Ax dt + dv \quad (6.49)$$

donde el proceso v tiene valor medio cero e incrementos no correlacionados. La covarianza incremental de v es $R_1 dt$. Sea $\{t_k; k=0, 1, \dots\}$ el conjunto de los instantes de muestreo. La integración de (6.49) sobre un período de muestreo da

$$x(t_{k+1}) = e^{A(t_{k+1}-t_k)} x(t_k) + \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-s)} dv(s)$$

Consideremos la variable aleatoria

$$e(t_k) = \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-s)} dv(s)$$

Esta variable tiene media cero porque v tiene media cero. Las variables aleatorias $e(t_k)$ y $e(t_1)$ tampoco están correlacionadas para $k \neq 1$ porque los incrementos de v sobre intervalos disjuntos no están correlacionados. La covarianza de $e(t_k)$ está definida por

$$\begin{aligned} E[e(t_k), e^T(t_k)] &= E \left[\int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-s)} dv(s) \right] dv^T(t) e^{AT(t_{k+1}-t)} \\ &= \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-s)} R_1 e^{AT(t_{k+1}-s)} ds \end{aligned} \quad (6.50)$$

De este modo se encuentra que la secuencia aleatoria $\{x(t_k), k=0, 1, \dots\}$ obtenida muestreando el proceso $\{x(t)\}$ se describe por la ecuación en diferencias

$$x(t_{k+1}) = e^{A(t_{k+1}-t_k)} x(t_k) + e(t_k)$$

donde $\{e(t_k)\}$ es una secuencia de variables aleatorias no correlacionadas con media cero y covarianza (6.50).

6.8. Conclusiones

El propósito principal de este capítulo es desarrollar modelos matemáticos para perturbaciones. El resultado es un enfoque uniforme a los modelos para una amplia variedad de señales. Las señales se consideran generadas por sistemas dinámicos alimentados con un pulso, una secuencia de pulsos o un ruido blanco. De modo



equivalente, las señales se pueden considerar generadas por sistemas dinámicos con condiciones iniciales.

En primer lugar se demuestra en la Sección 6.3 que las perturbaciones simples como el escalón, rampa y sinusode se pueden generar como salidas de sistemas lineales alimentados con un pulso. Las perturbaciones más complicadas se pueden considerar generadas como respuestas a pulsos de sistemas más complejos.

En la Sección 6.4 se muestra que la clase de perturbaciones podrían ampliarse alimentando a los sistemas con señales compuestas de varios pulsos. Esto conduce a las señales determinísticas a trozos. Una nueva expresión se da en la Sección 6.5, donde se elige un ruido blanco como señal de entrada al sistema generador de la perturbación.

Se ha obtenido un método para modelar diferentes tipos de perturbaciones. Las perturbaciones se caracterizan por un sistema dinámico

$$y(k) = \frac{B(q)}{A(q)} \epsilon(k) \quad (6.51)$$

donde la entrada ϵ puede ser un pulso, varios pulsos o un ruido blanco. El sistema se denomina *generador de perturbaciones*. Por supuesto, el sistema dinámico puede representarse en el espacio de estados.

El problema de la predicción es importante cuando se controlan sistemas con perturbaciones que no pueden medirse. El problema de predecir una señal definida por (6.51) consiste esencialmente en calcular ϵ a partir de y . Esto es lo mismo que invertir el sistema dinámico (6.51). Para obtener un sistema estable, el polinomio $B(q)$ debe tener todos sus ceros dentro del círculo unidad. En general, éste no será el caso para perturbaciones determinísticas. Una consecuencia es que el funcionamiento de la predicción se deteriora. Se emplea mucho tiempo para obtener la predicción. Del teorema de factorización espectral se deduce para un sistema estocástico que el polinomio $B(q)$ tiene todos sus ceros dentro o sobre la frontera del círculo unidad.

El punto esencial de la discusión es que los predictores para las señales están definidos de modo único por la función de transferencia discreta $H = B/A$. Los predictores son, por tanto, los mismos para entradas formadas por pulsos, trenes de pulsos o un ruido blanco. Esto significa que los predictores diseñados para perturbaciones determinísticas pueden trabajar muy bien para perturbaciones estocásticas si el generador de perturbaciones para las señales es el mismo.

El intento de unificar el modelado de perturbaciones conduce a una simplificación sustancial de la teoría porque es suficiente trabajar únicamente con unos pocos prototipos de perturbaciones.

6.9. Problemas

- 6.1. Enumerar situaciones en las que es posible reducir la influencia de las perturbaciones por a) reducción en la fuente, b) realimentación local y c) predicción.
- 6.2. Determinar H_d en la Figura 6.4 de modo que la salida muestreada sea a) una sinusode de frecuencia ω rad/s y b) lo mismo para $t \exp(-t)$ en los puntos de muestreo cuando la entrada es un pulso.

- 6.3. Demostrar que el predictor (6.4) es equivalente al predictor (6.3).
- 6.4. Determíñese el predictor de m pasos para el modelo de perturbación

$$y(k) = \frac{C(q)}{A(q)} w(k)$$

donde $w(k)$ es cero excepto en puntos aislados que están más espaciados que gra A . Utilizar el resultado para determinar la señal y el predictor cuando $A(q) = q - 0.5$, $C(q) = q$ y $m = 3$ y cuando $w(k)$ es cero excepto para $k = 0$ y 5. Se suponen condiciones iniciales nulas.

- 6.5. Utilizar el Teorema 6.1 para calcular la función de covarianza estacionaria del proceso

$$x(k+1) = \begin{bmatrix} 0.4 & 0 \\ -0.6 & 0.2 \end{bmatrix} x(k) + v(k)$$

donde v es un proceso de ruido blanco con media cero y varianza

$$R_1 = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

- 6.6. Considérese un proceso estocástico estacionario generado por

$$\begin{aligned} x(k+1) &= \Phi x(k) + v(k) \\ y(k) &= C x(k) \end{aligned}$$

donde $v(k)$ es una secuencia de variables estocásticas igualmente distribuidas de media cero. Supongamos que ϕ tiene la ecuación característica

$$z^n + a_1 z^{n-1} + \cdots + a_n = 0$$

Demostrar que las funciones de autocovarianza de la salida $r_y(\tau)$ satisfacen

$$r_y(\tau) + a_1 r_y(\tau - 1) + \cdots + a_n r_y(\tau - n) = 0$$

para $\tau \geq n+1$. (Esta ecuación se denomina *ecuación de Yule-Walker*).

- 6.7. Considérese el proceso

$$\begin{aligned} x(k+1) &= \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix} x(k) + v(k) \\ y(k) &= [1 \ 1] x(k) \end{aligned}$$

donde $v(k)$ es un ruido blanco de media cero y matriz de covarianza

$$R_1 = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

Demostrar que $y(k)$ puede representarse en la forma

$$y(k) = \lambda \frac{q+c}{(q+a)(q+b)} e(k)$$

donde $e(k)$ es un ruido blanco de media cero y varianza unidad. Encontrar la relación para determinar λ y c .



- 6.8. Un proceso estocástico $y(k)$ se describe por

$$\begin{aligned}x(k+1) &= ax(k) + v(k) \\y(k) &= x(k) + e(k)\end{aligned}$$

donde v y e son procesos de ruido blanco normalmente distribuidos con las propiedades

$$Ev = Ee = 0$$

$$\text{Var } v = 1$$

$$\text{Var } e = r_2$$

$$Ev(k)e(j) = r_{12} \quad \text{cuando } k=j \text{ y } 0 \text{ en los demás casos}$$

Demostrar que $y(k)$ puede representarse como la salida de un filtro lineal

$$y(k) = \lambda \frac{q - c}{q - a} e(k), \quad |c| < 1$$

- donde $e(k)$ es un ruido blanco de media cero y varianza unidad. Determinar λ y c .
- 6.9. Determíñese la función de covarianza, $r_y(\tau)$, y el espectro, $\phi_y(\omega)$, del proceso $y(k)$ cuando

$$y(k) - 0.7y(k-1) = e(k) - 0.5e(k-1)$$

donde $e(k)$ es ruido blanco de varianza unidad.

- 6.10. Determinar la varianza del proceso estocástico $y(k)$ definida por

$$y(k) - 1.5y(k-1) + 0.7y(k-2) = e(k) + 0.2e(k-1)$$

donde e es ruido blanco con varianza unidad.

6.10. Referencias

Los principios para reducir las perturbaciones por realimentación y prealimentación y los modelos clásicos de perturbaciones son los elementos claves de la teoría de realimentación clásica. Véanse

- BROWN, G. S., y D. P. CAMPBELL (1948): *Principles of Servomechanisms*. New York: John Wiley.
CHESTNUT, H., y R. W. MAYER (1959): *Servomechanisms and Regulating System Design*. (vol. 1), New York: John Wiley.
GILLE, J. C., M. J. PELEGREN, y P. DECAULNE (1959): *Feedback Control Systems*. New York: McGraw-Hill.

La noción de señales determinísticas a trozos se introdujo en

- ÅSTRÖM, K. J. (1980): «Piece-wise Deterministic Signals», en *Time Series*, ed. O. D. Anderson. Amsterdam: North Holland.

donde se demuestran las fórmulas de predicción y se dan algunos detalles al respecto. Tablas de integrales para valores pequeños de n se dan en

- JURY, E. I. (1964): *Theory and Application of the z-Transform Method*. New York: John Wiley. Second printing: Krieger, 1973.

Las ideas de representar las perturbaciones como procesos estocásticos también fue parte de la teoría clásica de realimentación. Véanse

- JAMES, H. M., N. B. NICHOLS y R. S. PHILIPS (1974): *Theory of Servomechanisms*. New York: McGraw-Hill.
TSIEN, H. S. (1955): *Engineering Cybernetics*. New York: McGraw-Hill.
LANING, J. H., y R. H. BATTIN (1956): *Random Processes in Automatic Control*. New York: McGraw-Hill.
NEWTON, G. C., JR., L. A. GOULD y J. F. KAISER (1957): *Analytical Design of Linear Feedback Controls*. New York: John Wiley.

Un tratamiento razonable completo de procesos estocásticos requiere un libro entero. Los libros

- KARLIN, S. (1966): *A First Course in Stochastic Processes*. New York: Academic Press.
PARZEN, E. (1962): *Stochastic Processes*. San Francisco: Holden-Day.
PAPOULIS, A. (1965): *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill.

proporcionan una buena base. También hay resúmenes breves en los libros sobre control estocástico que se mencionan a continuación.

La teoría de predicción se originó en los artículos

- KOLMOGOROV, A. N. (1941): «Interpolation and Extrapolation of Stationary Random Sequences», *Bull. Moscow Univ.*, USSR. Ser. Math. 5.
WIENER, N. (1949): *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. New York: John Wiley.
KALMAN, R. E. (1960): «A New Approach to Linear Filtering and Prediction Problems», *Trans. ASME, Ser. D. Journal Basic Eng.*, 82, 34-45.
KALMAN, R. E., y R. S. BUCY (1961): «New Results in Linear Filtering and Prediction Theory», *Trans. ASME, Ser. D., Journal Basic Eng.*, 83, 95-107.

Los artículos de Wiener y Kolmogorov no son de lectura fácil. El trabajo de Kolmogorov se puede comprender en

- WHITTLE, P. (1963): *Prediction and Regulation by Linear Least-Squares Methods*. London: English Universities Press.

Los resultados de Wiener se publicaron originalmente en un informe del MIT en 1942. Se conocieron como el peligro amarillo por el color amarillo de la cubierta y por el estilo de escritura. El artículo de Kalman escrito en (1960), que trata de los procesos discretos, es fácil de comprender. También existen libros completos dedicados a la predicción, teoría de filtrado y control estocástico.

- ÅSTRÖM, K. J. (1970): *Introduction to Stochastic Control Theory*, New York: Academic Press.
ANDERSON, B. D. O., y J. B. MOORE (1979): *Optimal Filtering*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.
BOX, G. E. P., y G. M. JENKINS (1970): *Time Series Analysis, Forecasting, and Control*. San Francisco: Holden-Day.
MCGARTY, T. P. (1974): *Stochastic Systems and State Estimation*. New York: John Wiley.

Diseño: visión de conjunto

Objetivo:

Exponer el diseño de sistemas de control y mostrar cómo encaja dentro del diseño y operación del proceso. Exponer la estructuración y selección de los principios del control. Revisar los métodos para especificar y diseñar lazos sencillos.

7.1. Introducción

Este capítulo enfoca el problema del control con una perspectiva más amplia. En la práctica, a menudo se dedica más tiempo a formular los problemas de control que a resolverlos. Por tanto, será útil estar al tanto de estos problemas más generales, aunque raramente se tratan en los libros de texto corrientes.

La mayor parte de los problemas de control surgen al diseñar sistemas en ingeniería. Normalmente estos problemas son de una gran dimensión y borrosos. Tareas típicas son el diseño de plantas de energía, robots industriales, aviones, vehículos espaciales y sistemas biomédicos. La teoría de control, por otra parte, trata problemas bien definidos de pequeña escala. Un problema típico es diseñar una ley de control para un sistema dado, que se representa por medio de ecuaciones diferenciales lineales con coeficientes constantes, por lo que el sistema en lazo cerrado tiene unos polos dados.

Una dificultad importante en el diseño de sistemas de control es cómo compaginar los problemas reales, borrosos y de gran dimensión con los problemas sencillos y bien definidos que puede manejar la teoría de control. Sin embargo, precisamente es en esta zona intermedia donde el ingeniero de control puede emplear eficazmente su creatividad e ingenio. Esta situación no es sólo exclusiva del ingeniero de control. Situaciones similares pueden encontrarse en casi todos los campos del diseño en ingeniería. Sin embargo, el control es un campo de la ingeniería en el que hace falta una teoría comparativamente sofisticada para comprender los problemas.

Es útil adoptar una cierta perspectiva respecto al proceso de diseño y un cierto

sentido del papel de la teoría en el proceso de diseño. En primer lugar, un buen diseño de ingeniería debe satisfacer un gran número de especificaciones, y a menudo hay muchas soluciones igualmente buenas para un problema concreto de diseño. Un buen diseño suele ser un compromiso basado en concesiones razonables entre coste y prestaciones. Desgraciadamente con demasiada frecuencia es cierto que lo mejor es el peor enemigo de lo bueno. En consecuencia cuando en este contexto se empleen palabras como *óptimo*, no deben tomarse al pie de la letra.

Otro aspecto es que a menudo se llega al diseño por interacción entre el cliente y el vendedor. En esta interacción intervienen muchos factores subjetivos; tales como orgullo, tradición y ambición. La situación es particularmente confusa con respecto a las preferencias del cliente cuando la tecnología está cambiando. Ejemplos típicos son las discusiones respecto a los reguladores neumáticos o electrónicos o al control analógico frente al digital, que proliferan en las revistas comerciales.

La contribución que la teoría *puede* hacer al proceso de diseño es proporcionar una penetración y comprensión del problema. En particular, la teoría puede señalar con precisión las limitaciones fundamentales respecto a las prestaciones del control. Hay también algunos problemas ideales de diseño que pueden resolverse teóricamente. Estas soluciones pueden proporcionar una buena penetración en las estructuras y algoritmos adecuados.

También es útil recordar que los problemas de control pueden ser de naturaleza muy diferente. Pueden variar desde el diseño de un lazo de control sencillo hasta el de un sistema integrado de control para un proceso completo. El enfoque del diseño también será muy diferente para sistemas producidos en masa y para sistemas únicos. Para sistemas producidos en masa hay que hacer un esfuerzo sustancial para obtener un sistema barato que tenga un buen comportamiento. Para sistemas únicos, a menudo es mucho mejor instalar un sistema comercial flexible y ajustarlo *in situ*.

También es importante la relación entre el diseño del proceso y el del control. Tradicionalmente los sistemas de control se han introducido en procesos dados para simplificar o mejorar su funcionamiento. Sin embargo, cada vez resulta más claro que puede ganarse mucho si se consideran conjuntamente el diseño del proceso y el del control, ambos en un mismo contexto. La posibilidad de disponer de un sistema de control siempre proporciona al diseñador un grado de libertad adicional, que con frecuencia se puede emplear en mejorar las prestaciones o la economía. De forma semejante, hay muchas situaciones en las que surgen problemas de control de difícil solución por causa de un diseño del proceso inadecuado. Compárese con la exposición sobre eliminación de perturbaciones de la Sección 6.2.

En la Sección 7.2 se discutirán algunos aspectos operacionales de los sistemas de control. En ella se incluyen las interfaces con el proceso, el operador y el computador. También se exponen diversos aspectos del diseño, puesta en servicio y operación del proceso. En la Sección 7.3 se presentan los problemas de estructuración. El problema básico es la descomposición de un problema grande y complicado en un conjunto de problemas más pequeños y sencillos. Esto incluye la elección de los principios de control, la selección y emparejamiento de las variables de control y de medida. También se tratan los principios normales de estructuración: aproximación descendente, ascendente, del medio hacia los extremos y de dentro hacia fuera. En la Sección



7.4 se trata la aproximación descendente. Incluye la elección de los principios de control y la selección y agrupamiento de las señales de control y de medida. En la Sección 7.5 se ve la aproximación ascendente, incluyendo la exposición de las estructuras de control elementales, realimentación, prealimentación, predicción, estimación, optimización y adaptación. También se tratan las combinaciones de estos tipos. En la Sección 7.6 se analiza el diseño de lazos sencillos. Se habla de las especificaciones del comportamiento como servo y en regulación. También se revisan los métodos de diseño de lazos sencillos. En los siguientes capítulos se lleva a cabo un tratamiento detallado de estos métodos.

7.2. Aspectos operacionales

Es muy útil comprender cómo interacciona el sistema de control con su entorno. Esta sección trata las interfases entre el diseño del proceso y del regulador. También se tratan la puesta en servicio, la operación y la modificación del sistema.

Diseño del proceso y del regulador

En las etapas iniciales de la automatización siempre se diseñaba el sistema de control cuando estaba completo el diseño del proceso. En muchos casos todavía es así. Como el diseño del proceso está ampliamente basado en consideraciones estáticas, este diseño puede conducir a un proceso difícil de controlar. Por esto es de gran utilidad considerar el diseño del control justamente con el del proceso. El hecho de que un proceso se vaya a controlar automáticamente también da a sus diseñadores un grado de libertad adicional, que puede usarse para llegar a mejores situaciones de compromiso. Por tanto, el regulador y el proceso deben diseñarse conjuntamente. Un ejemplo ilustrará esta idea.

Ejemplo 7.1 – Eliminación de perturbaciones mediante el mezclado

La eliminación de inhomogeneidades en un flujo de producto es uno de los mayores problemas en el control de procesos. Una posibilidad para reducir las variaciones es introducir grandes tanques de almacenamiento con lo que se incrementa el material almacenado en el proceso. Además un sistema con grandes tanques de mezclado tiene una respuesta lenta. Y en un sistema de este tipo llevará mucho tiempo el cambiar la calidad del producto. Una consecuencia de esto es que el producto puede estar fuera de especificaciones durante mucho tiempo durante un cambio de la calidad. Otra posibilidad para eliminar las inhomogeneidades es medir la calidad del producto y reducir las variaciones por un control de realimentación. En este caso es posible emplear tanques mucho más pequeños y lograr sistemas con respuesta mucho más rápida. El sistema de control resulta más complicado. Como el sistema global siempre tendrá una anchura de banda finita, siempre habrá que emplear *pequeños* tanques de mezclado para eliminar las variaciones rápidas.

Estabilidad frente a controlabilidad (maneobrabilidad)

Sucede frecuentemente que la estabilidad y la controlabilidad tienen requisitos contrapuestos. Por ejemplo, esto ha sido evidente en el diseño de vehículos. Los hermanos Wright tuvieron éxito en el diseño de su avión porque decidieron hacerle manejable aunque fuera inestable, mientras sus competidores, por el contrario, estaban diseñando aviones estables. En el diseño de barcos, un barco estable normalmente es difícil hacerlo girar mientras que un barco que gira fácilmente tiende a ser inestable. Tradicionalmente la tendencia ha sido dar importancia a la estabilidad. Sin embargo, es interesante comprobar que si se emplea un sistema de control, el sistema básico puede por el contrario diseñarse para que sea controlable. La estabilidad requerida puede proporcionarse mediante el sistema de control. Para demostrar que se pueden conseguir ahorros considerables con este enfoque se suele utilizar un ejemplo sacado del diseño de aviones.

Ejemplo 7.2 – Diseño de un avión supersónico

Para un avión de elevadas prestaciones, que ha de operar en un amplio rango de velocidades, el centro de presión se mueve hacia atrás con velocidad creciente. Para un moderno avión de combate supersónico, el desplazamiento del centro de presiones puede llegar a ser de 1 m. Si se diseña el avión para que sea estáticamente estable a velocidades subsónicas, el centro de masas estará unos pocos decímetros por delante del centro de presiones a bajas velocidades. A velocidades supersónicas, la distancia entre el centro de masas y el de presiones aumentará hasta más de 1 m. Habrá entonces un elevado par estabilizador, que tenderá a mantener el avión en línea recta. El par será proporcional al producto del impulso por la distancia entre los centros de masa y de presión. Para maniobrar el avión a altas velocidades será necesario un gran timón de dirección. Sin embargo, un timón muy grande producirá una resistencia considerable.

Hay una gran ventaja en cambiar el diseño para que el centro de masas esté en la mitad del rango de variación del centro de presión. Entonces puede utilizarse un timón mucho menor, y la resistencia producida por el timón se reduce. La reducción de la resistencia puede ser del orden del 10 %. Un avión de este tipo, sin embargo, será estáticamente inestable a bajas velocidades, esto es, ¡en el despegue y aterrizaje! Se puede conseguir la estabilidad necesaria usando un sistema de control. Por supuesto que un sistema de control de este tipo ha de ser absolutamente fiable.

Las ideas actuales en diseño de aviones se están desplazando hacia el diseño de aviones que son estáticamente inestables a bajas velocidades y lograr una estabilidad suficiente utilizando un sistema de control. Ejemplos similares son corrientes en el diseño de otros vehículos.

También se dan casos análogos en el control de procesos químicos. El siguiente es un caso típico.

Ejemplo 7.3 – Reactor químico exotérmico

Para obtener un rendimiento elevado en un reactor químico exotérmico puede interesar hacer funcionar el reactor en condiciones de operación en las que éste es inestable en lazo abierto. Obviamente la seguridad de la operación depende de forma crítica del sistema de control que estabilice al reactor.

Controlabilidad, observabilidad y dinámica

Cuando se diseña un proceso es muy importante estar seguro de que todas las variables importantes del proceso pueden cambiarse de la forma que convenga. La palabra *controlabilidad* se emplea a menudo en este contexto, aunque se interpreta en un sentido mucho más amplio que en el de los conceptos formales de controlabilidad introducidos en la Sección 5.3.

Para obtener plantas que sean controlables en un sentido amplio lo primero que hace falta es tener un número suficiente de actuadores. Si hay cuatro variables importantes en el proceso que pueden manipularse separadamente, debe haber al menos cuatro actuadores. Es más, el sistema debe ser tal que la relación estática entre las variables del proceso y los actuadores sea uno a uno. Para lograr un buen control, la relación dinámica entre los actuadores y las variables del proceso debe ser idealmente tal que sea posible un control muy preciso. Esto significa que deben evitarse los retardos y las relaciones de fase no mínima. Idealmente las relaciones dinámicas deben ser como un integrador o un retardo de fase de primer orden. Sin embargo, normalmente es difícil obtener procesos de este tipo. En la dinámica de los procesos industriales son corrientes los lazos de fase no mínima.

A menudo son muy útiles los modelos de dinámica simple para establecer la dinámica de los sistemas en la etapa de diseño del proceso. Los actuadores deben diseñarse de forma que se puedan variar las variables del proceso dentro de un rango suficiente con una buena resolución. También las relaciones deben ser tales que la ganancia no cambie demasiado en todo el rango de operación. Una equivocación frecuente en los sistemas de flujo es escoger una válvula de control demasiado grande. Esto conduce a una relación fuertemente no lineal entre la apertura de la válvula y el flujo. El flujo cambia muy poco cuando se reduce la apertura de la válvula hasta que está casi cerrada. Entonces se produce un cambio brusco en el flujo en un rango muy pequeño de la posición de la válvula.

El proceso también debe tener los sensores apropiados, cuyas señales estén íntimamente relacionadas con las variables importantes del proceso. Los sensores deben colocarse para suministrar las señales que son representativas para las variables importantes del proceso. Por ejemplo, debe tenerse cuidado de no colocar los sensores en bolsas donde las propiedades del fluido del proceso es posible que no sean típicas. Deben evitarse los retardos de tiempo. Estos pueden ocurrir debido a factores tales como transportes o encapsulación de los sensores de temperatura.

Los modelos dinámicos simples junto con el análisis de observabilidad, son muy útiles para lograr los emparejamientos adecuados de sensores y actuadores. También es muy útil para esto estimar las constantes de tiempo a partir de modelos dinámicos sencillos.

Diseño de regulador o sintonizado en línea

Otro hecho que influye de forma drástica en el diseño del regulador es el esfuerzo que puede emplearse en el diseño. Para sistemas que se van a producir en grandes cantidades, puede ser posible emplear mucho esfuerzo de ingeniería para diseñar un

regulador. Se puede diseñar un regulador de parámetros fijos que no requieren ningún ajuste. En muchos casos, sin embargo, no es económicamente factible emplear mucho esfuerzo en el diseño del regulador. En estas aplicaciones suele usarse un regulador normal de propósito general con parámetros ajustables. Se instala el regulador y se ponen los parámetros adecuados mediante un proceso de ajuste.

Con el control con computador han aumentado drásticamente las posibilidades de diseñar reguladores flexibles de propósito general. Cuando se realiza un regulador sobre un computador es posible proporcionar al sistema herramientas de ayuda con computador que simplifican el diseño y el sintonizado.

En el control de procesos, la mayoría de los lazos para el control del nivel de líquido, temperatura, flujo y presión se diseñan mediante reglas empíricas y se ajustan sobre el proceso. Sin embargo, para el control de composición y pH se aplican técnicas sistemáticas de diseño, lo mismo que para el control de sistemas no lineales multivariados y distribuidos como las columnas de destilación.

Interacción entre proceso, regulador y operador

El regulador y el proceso deben trabajar conjuntamente de forma correcta. Normalmente un regulador se diseña para operar en estado estacionario, que es un estado de operación. Sin embargo, es necesario estar seguro de que el sistema también trabajará correctamente durante el arranque y la parada y bajo condiciones de emergencia, como fallos drásticos del proceso. Durante condiciones normales lo habitual es diseñar para conseguir el máximo rendimiento. En caso de fallo, puede ser mucho más importante el recuperarse y volver rápidamente a una condición de operación segura.

En control de procesos ha sido costumbre emplear la regulación automática para la operación en condiciones estacionarias. En otras condiciones de operación, el regulador se pasa a manual y un operador se encarga del control. Con un nivel de automatización mayor, se requerirá un buen control sobre mayor número de estados de operación.

7.3. Principios de estructuración

Como hemos mencionado anteriormente, los problemas de control reales son de gran tamaño y borrosos, y la teoría de control trata problemas pequeños y bien definidos. Según el diccionario, *estructurar* puede significar construir un sistema de referencia sistemático para algo. Sin embargo, en este contexto se emplea estructurar para describir el proceso de salvar la distancia entre los problemas reales y los que la teoría de control puede manejar.

Para el diseño de sistemas de control son muy importantes los problemas asociados con la estructuración. Desgraciadamente estos problemas todavía no pueden disponerse en una estructura sistemática completa. Por esto es por lo que frecuentemente se eluden estos problemas tanto en los libros de texto como en la investigación. Haciendo una analogía se puede decir que hay la misma relación entre

estructuración y diseño de sistemas de control que entre gramática y redacción. Está claro que es imposible escribir bien sin saber gramática. También lo es que un escrito que no tiene fallos gramaticales no es necesariamente un buen escrito. La estructuración de un sistema de control debe basarse en los principios científicos dados por la teoría de control. Sin embargo, la estructuración también contiene elementos de creatividad, ingenio y arte. Quizás la mejor forma de introducir la estructuración es enseñarla como una *destreza o habilidad*.

El problema de la estructuración se presenta en muchas disciplinas. También se han desarrollado planteamientos formales. La terminología empleada aquí se ha tomado de los campos de la informática y la resolución de problemas, donde la estructuración de grandes programas ha sido el tema de gran cantidad de trabajos. Hay dos enfoques principales, llamados descendente (*top-down*) y ascendente (*bottom-up*).

El planteamiento *descendente* comienza por la definición del problema. Entonces se divide sucesivamente el problema en partes más pequeñas, añadiendo cada vez más detalles.

El proceso termina cuando todas las partes corresponden a problemas perfectamente conocidos. Una característica de este enfoque es que al principio se dejan a un lado muchos detalles. Cuando se va subdividiendo el problema se van incorporando nuevos detalles. Por esto a esta forma de trabajar se le asocia frecuentemente la expresión: *refinamiento sucesivo*.

El planteamiento *ascendente* comienza, por el contrario, con las piezas pequeñas, que representan soluciones conocidas. Estas se combinan en partes cada vez mayores, hasta que se obtiene una solución al problema global.

A menudo se considera que el enfoque descendente es más sistemático y más lógico. Por supuesto, no es posible emplear este enfoque a menos que se conozcan muy bien los detalles del sistema. De manera similar, no es fácil utilizar el enfoque ascendente si no se conocen las características del problema completo. En la práctica, lo normal es utilizar diferentes combinaciones de ambas aproximaciones. A esto es a lo que a veces se le llama un enfoque *de dentro hacia fuera*.

La estructuración es un procedimiento iterativo. Pasará mucho tiempo antes de que se obtenga un procedimiento totalmente sistemático para realizar la estructuración. Es difícil percibir los problemas de la estructuración si no se abordan problemas de tamaño y complejidad considerables. Por tanto, la mayor parte del trabajo de estructuración se lleva a cabo en la industria. También es normal que muchas industrias tengan ingenieros capaces de realizar muy bien este trabajo de estructuración. Por eso se aconseja a los estudiantes que aprendan lo que los «expertos en estructuración» hacen, de la misma forma que los pintores han aprendido siempre de los grandes maestros.

7.4. Aproximación descendente

Esta sección describe una aproximación descendente al diseño de un sistema de control. Esto incluye la selección de los principios de control, la elección de las variables de control y de medida, y el emparejamiento de estas variables.

Principios de control

Un principio de control da una indicación amplia de cómo se debe controlar un proceso. Entonces, el principio de control dice cómo debe responder un proceso a las perturbaciones y a las señales de control. El establecimiento de un principio de control es el punto de arranque para el diseño descendente. A continuación se dan algunos ejemplos de principios de control.

Ejemplo 7.4 – Control de flujo

Cuando se controla una válvula es posible controlar la posición de la válvula, el flujo o ambos. Lo más sencillo y barato es controlar la posición. Como el flujo es, en general, una función no lineal de la apertura de la válvula, esto conduce a un sistema en que la relación entre la variable de control (posición de la válvula) y la variable física (flujo) es fuertemente no lineal. La relación cambiará con variables tales como las variaciones de presión y la apertura de la válvula. Estas dificultades se evitan si se controlan la posición de la válvula y el flujo simultáneamente. Sin embargo, un sistema para control de flujo es más complicado puesto que requiere un caudalímetro.

Ejemplo 7.5 – Control de composición

Cuando se controlan variables importantes referentes a la calidad del producto, lo que se quiere normalmente es mantenerlas lo más próximas posibles a los valores prefijados. Esto puede conseguirse minimizando la varianza de las variaciones de la calidad del producto. Si se introduce un flujo en un gran tanque de almacenamiento con mezclado, se minimizarán las variaciones de calidad. No es precisamente lo mismo que minimizar variaciones de calidad en el flujo dentro del tanque.

Ejemplo 7.6 – Control de una caldera de vapor

Consideremos una turbina y un generador, que están alimentados por una caldera de vapor. Para controlar el sistema bajo una demanda de potencia variable, la caldera puede controlarse de tal forma que la presión del colector de vapor sea constante cuando se producen los cambios de potencia. La caldera también puede controlarse de tal forma que cambie la presión con la demanda de potencia. Estos dos principios de control se conocen con los nombres de *presión constante* y *presión deslizante*, respectivamente. Si se mantiene constante la presión, sólo se producirán pequeñas tensiones térmicas sobre la caldera. Por otra parte, el control a presión constante no explota la energía almacenada en el colector de vapor para responder a las demandas repentinas de potencia.

Ejemplo 7.7 – Control de un barco

Cuando se diseña un piloto automático para un barco de gran maniobrabilidad, existen muchas alternativas en el diseño. Una posibilidad es diseñar el piloto automático de forma que el capitán pueda ordenar un giro a un nuevo rumbo con una velocidad de giro dada. Otra posibilidad es especificar el radio de giro en lugar de la velocidad de giro. La ventaja de especificar el radio de giro es que la trayectoria del barco será independiente de la velocidad del barco. El control del radio de giro conduce a un sistema de control más complicado, porque es necesario medir tanto la velocidad de giro como la del barco.

Ejemplo 7.8 – Control de un balance de material

Muchos procesos entrañan flujo y almacenamiento de materiales. Aunque los procesos son muy diferentes, todos incluyen el almacenamiento de materiales. La razón para

introducirlos es que suavizan las variaciones del flujo de material. Por tanto, no tiene sentido controlar estos sistemas de forma que los almacenes tengan un contenido constante. Por el contrario, los criterios consistirían en mantener lo siguiente:

Almacenamientos entre unos límites máximo y mínimo.

Un balance exacto de material a largo plazo entre la entrada y la salida.

Velocidades de flujo suaves.

Ejemplo 7.9 – Control de restricciones

Cuando se diseñan sistemas de control, con frecuencia es necesario considerar varias condiciones de operación. Esto significa que puede que haga falta tener en consideración las restricciones para mantener unas condiciones de tipo económico o de seguridad. También puede hacer falta considerar condiciones especiales durante el arranque y la parada. Durante estas situaciones el control se suele realizar mediante controladores lógicos. Hoy día el control lógico y el control analógico a menudo se realizan con el mismo equipo, a los que se suele llamar controladores lógicos programables (PCL, del inglés Programmable Logic Controller) o también autómatas programables. Esto significa que hay grandes posibilidades de integrar diferentes funciones del sistema de control.

La elección del principio de control es un aspecto importante. Un buen principio de control debe normalmente simplificar el problema del control. A menudo la selección incluye relaciones de compromiso técnicas y económicas. La selección del principio de control se suele basar en la investigación de los modelos del proceso. Los modelos utilizados para esto son modelos típicamente internos deducidos a partir de los principios físicos. Por tanto, es difícil definir reglas generales para encontrar los principios de control.

Elección de las variables de control

Después de escoger los principios de control, el siguiente paso lógico es escoger las variables de control. La elección de las variables de control está limitada normalmente por razones prácticas. Debido a que la selección del principio de control determina qué variables físicas deben controlarse, lo normal es escoger como variables de control las que tienen una relación más próxima con las variables dadas por el principio de control. Como para seleccionar los principios de control hacen falta los modelos matemáticos, también se les puede utilizar para los estudios de controlabilidad cuando se escogen las variables de control.

Elección de las variables de medida

Cuando se han escogido los principios de control, ya se ha hecho una primera elección de las variables de medida. Si las variables empleadas para expresar el principio de control no pueden medirse, lo natural es escoger unas variables de medida que estén fuertemente relacionadas con estas variables de control. Para realizar esta elección pueden ser muy útiles los modelos matemáticos y el análisis de

observabilidad. Pueden encontrarse ejemplos típicos en el control de procesos químicos, donde se usan las temperaturas —que son fáciles de medir— en lugar de las composiciones, que son difíciles y costosas de medir.

Emparejamiento de entrada y salidas

Un sistema de grandes dimensiones normalmente tendrá un gran número de entradas y salidas. Incluso aunque inicialmente se encuentre un principio de control que sólo involucre unas pocas variables, después de introducir las variables que pueden manejarse y medirse, hay que tener en consideración muchas más variables. Con una aproximación descendente, un sistema debe descomponerse en subsistemas más pequeños. Entonces interesa agrupar en bloques las diferentes entradas y salidas, para obtener un conjunto de sistemas más reducidos. Si es posible, este agrupamiento debe hacerse de forma que: 1) haya solamente acoplamiento débil entre los diferentes subsistemas, y 2) cada subsistema se comporte bien dinámicamente; es decir, que las constantes de tiempo sean de la misma magnitud, y se eviten los retardos y las variaciones bruscas y de fase no mínima en la dinámica del proceso.

No hay reglas generales para el agrupamiento. Tampoco hay ningún método bueno para determinar si es posible o no un agrupamiento con las propiedades deseadas. Una posibilidad es la aproximación por tanteos, combinado con el análisis de modelos.

Ejemplo 7.10 – Control del balance de material

En la Figura 7.1 puede verse un sistema en el que hay un flujo de material. El sistema consiste en una serie de tanques. Los flujos entre los tanques se controlan mediante bombas. La figura ilustra dos estructuras de control diferentes. En uno el flujo de salida de cada tanque se controla a partir del nivel del tanque. Es lo que se llama control *en la dirección del flujo*. Para mantener el equilibrio entre la producción y la demanda, hace falta controlar el flujo que entra en el primer tanque mediante una realimentación del nivel del último tanque. En el otro planteamiento se controla el flujo de entrada en cada tanque mediante el nivel del tanque. Esto se conoce como control *a contraflujo*. Esta forma de control es mejor, ya que todos los lazos de control son sistemas de primer orden sencillos y no hay problemas de estabilidad. Con el control en la dirección del flujo puede haber inestabilidades debido a la realimentación que rodea al conjunto de todos los tanques. También puede verse que el control a contraflujo puede llevarse a cabo con tanques de almacenamiento más pequeños.

7.5. Aproximación ascendente

En el enfoque ascendente, lo primero que hay que hacer es una elección de las variables de control y de medida. Después se introducen diferentes reguladores hasta que se consigue un sistema en lazo cerrado con las propiedades deseadas. Los reguladores que se emplean para construir el sistema son de tipo normal basados en las ideas de realimentación, prealimentación, predicción y estimación, optimización y

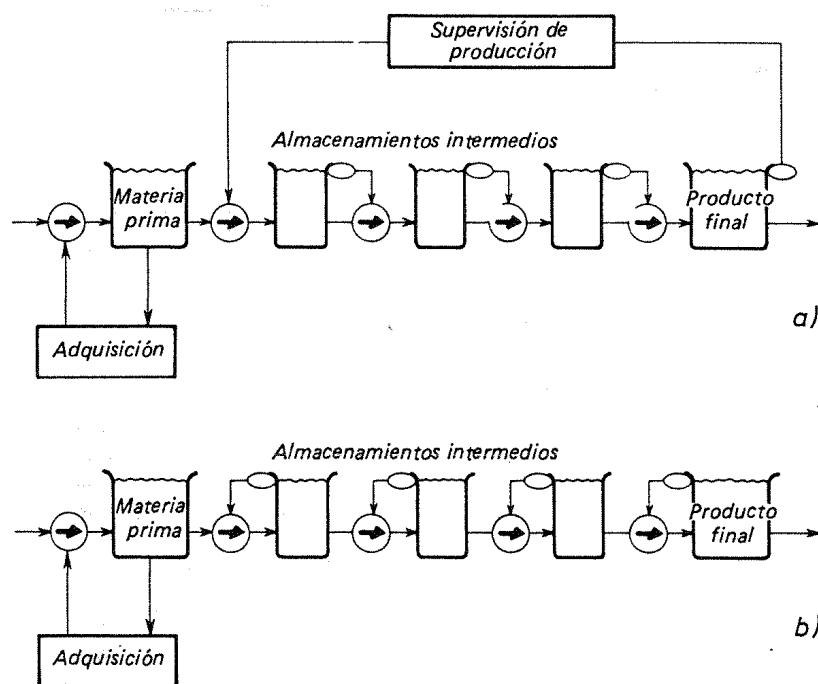


Figura 7.1 Control de balance de material en la dirección del flujo a) y a contraflujo b).

adaptación. Como estas técnicas son familiares por los cursos elementales, sólo se presentarán aquí de forma breve.

Realimentación

Los lazos de realimentación empleados incluyen, por ejemplo, reguladores PID sencillos y sus combinaciones en cascada. Cuando se emplean computadores digitales para realizar los reguladores, también resulta fácil el empleo de controles más sofisticados, como predictores de Smith para compensación de tiempos muertos, realimentación del estado y control con modelo de referencia.

La realimentación se usa en el contexto usual. Su ventaja es que puede reducir la sensibilidad a las perturbaciones y a las variaciones de los parámetros. La realimentación es más eficaz cuando la dinámica del proceso admite el empleo de una gran anchura de banda. En los siguientes capítulos se estudiará en profundidad el diseño de la realimentación.

Muchos sistemas que son difíciles de realizar empleando técnicas analógicas pueden realizarse fácilmente con tecnología de control con computador. El predictor de Smith es un ejemplo típico. En la Figura 7.2 puede verse un diagrama de bloques de este regulador, que muestra la necesidad de almacenar señales. Esto es difícil de llevar a cabo con dispositivos analógicos, pero muy fácil con un computador. En el Capítulo 8 se vuelve a tratar el predictor de Smith.

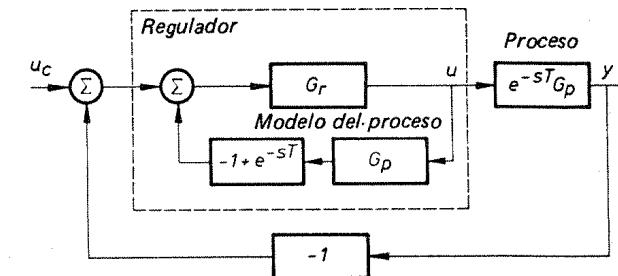


Figura 7.2 Diagrama de bloques de un predictor de Smith.

Prealimentación

La prealimentación es otro método de control. Se emplea, como se dijo en el Capítulo 6, para eliminar perturbaciones que pueden medirse. La idea básica es utilizar la perturbación medida para anticipar su influencia sobre las variables del proceso e introducir las acciones compensadoras de control apropiadas. La ventaja que tiene respecto a la realimentación es que se puede tomar la acción correctora antes de que la perturbación haya influido a las variables. Debido a que la prealimentación es una compensación en lazo abierto, requiere un buen modelo del proceso. Con el control digital es fácil incorporar este modelo. Por tanto, puede anticiparse que con el control digital aumentará el uso de la prealimentación.

El diseño de un compensador por prealimentación es en esencia el cálculo del inverso de un sistema dinámico.

Predicción y estimación

Las variables de estado y los parámetros no se pueden medir directamente. En este caso es conveniente procurar conocer las cantidades cuando se diseña la realimentación. Las variables desconocidas pueden sustituirse por estimaciones o predicciones. De hecho en algunos casos esta solución es la óptima. Por tanto, las nociones de predicción y estimación son importantes. Mediante técnicas analógicas pueden generarse fácilmente estimadores de variables de estado para sistemas lineales.

También se pueden realizar fácilmente usando un computador. Los estimadores de parámetros son más difíciles de realizar con métodos analógicos. Estos, sin embargo, sí pueden realizarse con un computador. Por tanto, la estimación y la predicción son más fáciles de usar en control con computador.

Optimización

Algunos problemas de control se pueden expresar adecuadamente como problemas de optimización. Con sistemas de control con computador es posible incluir algoritmos de optimización como elementos del sistema de control.

Combinaciones

Cuando se emplea un enfoque ascendente, las estructuras básicas de control se combinan para obtener una solución al problema de control. A menudo conviene hacer estas combinaciones de una forma jerarquizada. Muchas combinaciones, como el control en cascada, la realimentación de estado, y los observadores, se conocen de los cursos elementales. Los sistemas de control muy complejos pueden construirse combinando las estructuras sencillas. En la Figura 7.3 se puede ver un ejemplo. De hecho esta forma de diseñar el control empleando el enfoque ascendente es la más usada en el control de procesos. Su éxito depende en gran medida de la experiencia y la habilidad del diseñador.

En la Figura 7.4 puede verse un sistema adaptativo, que se ha obtenido combinando un estimador de parámetros con un regulador diseñado por un procedimiento normal.

7.6. Diseño de lazos sencillos

Si se emplea el procedimiento descendente, el método de diseño acabará en el diseño de lazos sencillos que contienen uno o varios controles, o medidas. Si se emplea la aproximación ascendente, el diseño comenzará con el diseño de lazos sencillos. Por tanto, el diseño de lazos sencillos es un paso importante en ambos enfoques. Además el diseño de lazos sencillos es un área en el que hay disponible una cantidad de estudios teóricos muy significativa, que se expondrá con detalle en los capítulos restantes. Para proporcionar una primera perspectiva, en esta sección daremos una visión de conjunto de los métodos de diseño existentes para lazos sencillos. Se analizarán los problemas modelo del diseño del regulador y del servo.

Criterios sencillos

Una forma sencilla de especificar las prestaciones de la regulación es fijar los errores permitidos para unas perturbaciones típicas. Por ejemplo, puede hacer falta

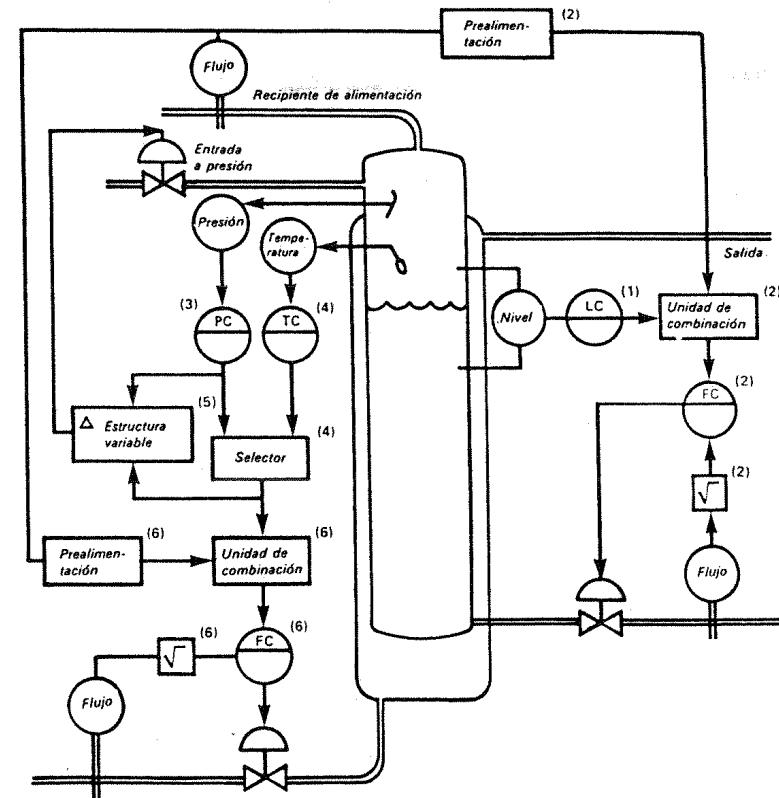


Figura 7.3 Ejemplo de un sistema de control complejo construido a partir de estructuras de control sencillas. (Adaptada por autorización de Foxboro Company.)

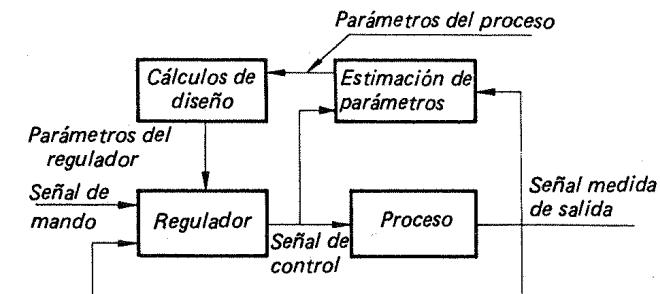


Figura 7.4 Diagrama de bloques de un regulador autosintonizado obtenido combinando un estimador de parámetros y cálculos de diseño.



que una perturbación tipo salto escalón no produzca error estacionario, y que el error debido a una perturbación tipo rampa sea una fracción de la velocidad de la rampa. Estas especificaciones se expresan normalmente en términos del comportamiento estacionario, como se expuso en la Sección 5.4. Los coeficientes de error dan los requisitos sólo para el comportamiento a bajas frecuencias. Por tanto, debe especificarse, además de los coeficientes de error, el ancho de banda del sistema.

Otra forma más completa de especificar las prestaciones de la regulación es dar las condiciones acerca de la función de transferencia entre las perturbaciones y la salida del proceso.

Especificaciones para el problema del regulador

El propósito de la regulación es mantener las variables del proceso lo más próximas posible a los valores especificados a pesar de las perturbaciones del proceso y de las variaciones en la dinámica del propio proceso.

Control de varianza mínima. Para la regulación de variables de calidad importantes, a menudo es posible establecer criterios objetivos para el comportamiento de la regulación. En la Figura 7.5 se presenta una situación corriente, que muestra la distribución de las variables de calidad. A menudo se especifica que un determinado porcentaje de la producción debe tener un nivel de calidad superior a un valor dado. Al reducir las variaciones de calidad, es posible entonces situar la referencia más cerca del objetivo. La mejora en el comportamiento puede expresarse en términos de una reducción en el consumo de energía o de materias primas o una mayor producción. En este caso se pueden expresar las reducciones en la variación de la calidad directamente en términos económicos.

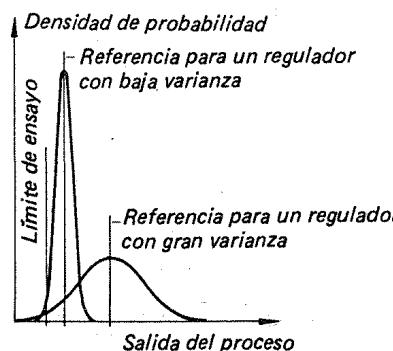


Figura 7.5 Expresa el comportamiento de la regulación en términos de la variación de unas variables de calidad.

Para procesos con un gran volumen de producción, las reducciones de una fracción de tanto por ciento pueden suponer una gran cantidad de dinero. Por ejemplo, una reducción del 1 % en la variación de la humedad en el control de una máquina de papel puede suponer un ahorro de 100.000 dólares anuales.

Si expresamos las variaciones de calidad por medio de distribuciones gaussianas, el criterio simplemente será minimizar la varianza de las variables de calidad. En estos problemas, las acciones de control necesarias son irrelevantes mientras no producen un desgaste excesivo ni señales demasiado grandes. La estrategia de control que minimiza la varianza de la salida del proceso recibe el nombre de *control de varianza mínima*.

Control óptimo. El control de varianza mínima es un ejemplo típico de cómo se puede especificar un problema de control como un problema de optimización. En un caso más general, no interesa minimizar la varianza de la salida. En su lugar hay un criterio del tipo

$$E \int_{t_0}^{t_1} g[x(s), u(s)] ds$$

donde x es la variable de estado, u es la variable de control y E representa el valor medio. A continuación se da un ejemplo de un criterio como éste.

Ejemplo 7.11 – Pilotaje de un barco

Puede verse que el incremento relativo de resistencia debido a las desviaciones de un movimiento en línea recta pueden expresarse de forma aproximada por

$$\frac{\Delta R}{R} = \frac{k}{T} \int_0^T [\psi^2(t) + \rho \delta^2(t)] dt$$

donde ψ es la desviación de la proa, δ es el ángulo del timón, R es la resistencia y ρ es un parámetro. Un buque tanque tiene como valores típicos de los parámetros $k=0,014$ y $\rho=0,1$.

Técnicas para el diseño de reguladores

Se pueden emplear muchos métodos para estimar el comportamiento de una regulación y para diseñar reguladores. A menudo los problemas de regulación se resuelven mediante una realimentación, pero las técnicas de prealimentación pueden ser muy útiles para eliminar perturbaciones que son medibles.

Si se dan las especificaciones en términos de la función de transferencia que relaciona la salida con la perturbación, lo normal es aplicar métodos que admitan el control de esta función de transferencia. Un método es la asignación de polos, que permite la especificación de la función de transferencia completa. Esta técnica directa de diseño se expone en detalle en los Capítulos 9 y 10. A menudo es demasiado restrictiva para especificar completamente la función de transferencia en lazo cerrado, lo que es un inconveniente.

Otra posibilidad es emplear un método de respuesta en frecuencia, que admite el control de la respuesta en frecuencia de la perturbación sobre la salida. La forma más apropiada de expresar los problemas de este tipo es mediante la teoría de los sistemas continuos. Los reguladores obtenidos se traducen a algoritmos de control digital mediante las técnicas descritas en el Capítulo 8.

Si los criterios se expresan como criterios de optimización, lo normal es emplear técnicas de diseño basadas en la optimización. En los Capítulos 11 y 12 se exponen técnicas basadas en la minimización de la varianza de la salida del proceso y de otros tipos de criterios cuadráticos.

Especificaciones para el problema del servo

En el problema del servo, la tarea es conseguir que las variables del proceso respondan a los cambios de la señal del mando de una determinada forma. Las técnicas de realimentación son útiles para el problema del servo. Como la señal de mando debe ser conocida, también es normal combinar la realimentación con la prealimentación.

Seguimiento del modelo. Una forma de expresar cómo debe responder el sistema a una señal de mando es dar un modelo de la respuesta deseada. Esto puede hacerse, por ejemplo, especificando la función de transferencia deseada entre la señal de control y las variables del proceso. Si no es posible alcanzar una coincidencia perfecta, puede minimizarse la desviación respecto de la respuesta del modelo.

Especificaciones en el dominio del tiempo y de la frecuencia. Como un sistema lineal queda perfectamente especificado por su respuesta a un salto escalón, es posible expresar la respuesta de un servo en términos de las especificaciones de la respuesta deseada a un salto en lazo cerrado. Estos tipos de especificaciones incluyen el tiempo de subida, T_r , rebasamiento, M , tiempo de asentamiento, T_s , errores estacionarios, e_0 , ancho de banda, ω_B , y pico de resonancia, M_p . El tiempo de establecimiento se define como el tiempo que transcurre hasta que la respuesta a un salto escalón está dentro de la banda $\pm p$ de su valor estacionario. Un valor empleado para p es 5 %. Compárense las Figuras 7.6 y 7.7.

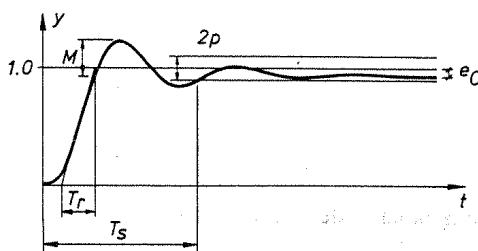
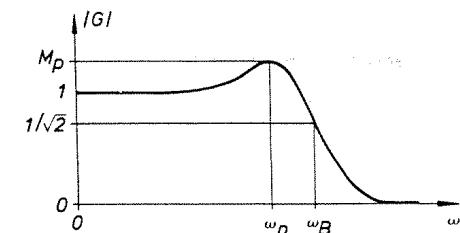


Figura 7.6 Expresión de las especificaciones del servo mediante los requisitos de la respuesta a un salto escalón.

Figura 7.7 Diagrama de Bode de la función de transferencia en lazo cerrado G_c ; ilustra cómo se pueden expresar las especificaciones del servo en términos de los requisitos de la respuesta en frecuencia. El sistema de la figura tiene ganancia estacionaria unidad.



Hay parámetros de diseño que pueden especificarse a partir del trazo de Bode en lazo abierto. Ejemplos de estos son el margen de ganancia, A_m , el margen de fase, φ_m , y la frecuencia de cruce ω_c . En esencia la frecuencia de cruce es proporcional al ancho de banda del sistema en lazo cerrado.

Técnicas para el diseño del servo

Hay muchas técnicas que pueden utilizarse para el diseño de servos. Si se pueden medir todas las variables de estado, lo normal es la asignación de polos. Si el sistema es controlable, se puede diseñar una realimentación de estado —que da los polos deseados en lazo cerrado. Si no se pueden medir todas las variables de estado, pueden reconstruirse mediante un observador o por un filtro de Kalman. En los Capítulos 9, 10 y 11 se presentan técnicas de diseño basadas en estas ideas, y también se muestra que este procedimiento de diseño puede interpretarse como un diseño de seguimiento de modelo que combina realimentación y prealimentación.

7.7. Conclusiones

Este capítulo presenta una visión de conjunto del problema de diseño. Hay un largo recorrido desde los grandes y borrosos problemas del mundo real hasta el pequeño y bien definido problema que la teoría de control puede manejar. Se han expuesto también los problemas de la estructuración.

Se ha introducido la noción de principio de control para aplicar la aproximación descendente. También se ha mostrado cómo se puede emplear la aproximación ascendente para construir sistemas complejos a partir de las sencillas estructuras de control de realimentación, prealimentación, estimación y optimización. Finalmente, se han expuesto las especificaciones y enfoques para el diseño de lazos de control sencillos.

Un proceso químico consiste en muchas operaciones unitarias, tales como reactores, mezcladores, y columnas de destilación. En un enfoque del diseño de sistemas de control ascendente se diseñan primero los lazos de control para las operaciones unitarias individuales. Entonces se añaden las interconexiones para obtener un sistema global. En el enfoque descendente, primero se postulan los



principios de control —como el control de composición y control de balance de material— para la totalidad de la planta. En la descomposición se aplican estos principios a las unidades y lazos individuales.

En el control de procesos la mayoría de los lazos para control de nivel de líquidos, flujo y presión, principalmente por conveniencia, se diseñan de forma intuitiva y se ajustan en línea. Sin embargo, el control de composición y de pH, así como el control de grandes sistemas no lineales distribuidos con fuertes interacciones, a menudo se diseñan con cuidado.

7.8. Problemas

- 7.1. Considérese el problema de balance de material presentado en la Figura 7.1. Supóngase que cada tanque (almacenamiento) es un integrador y que cada controlador es un controlador proporcional. Discútase la influencia sobre los dos sistemas de una perturbación tipo pulso en el almacenamiento de la materia prima.
- 7.2. Dedúzcase la función de transferencia del sistema en lazo cerrado de la Figura 7.2. Interprétese el comportamiento del sistema en lazo cerrado. ¿Cuál es la característica principal del predictor de Smith? (Sugerencia: Compárese con el caso en que no hay retardo y cuando el controlador es G_r .)
- 7.3. Identificar y explicar el empleo de a) el control en cascada, b) la prealimentación y c) los elementos no lineales en la Figura 7.3.

7.9. Referencias

El problema expuesto en este capítulo toca diferentes aspectos de la resolución de problemas. Un lector con interés en este aspecto puede encontrar interesante las lecturas

POLYA, G. (1945): *How to Solve It*. N. J.: Princeton University Press.

que toma muchos problemas del dominio de las matemáticas y

WIRTH, N. (1975): *Algorithms + Data Structures = Programs*. Englewood Cliffs, N. J.: Prentice Hall, Inc.

que lo aplica a la programación de computadores. En la literatura sobre el control de procesos hay algunos trabajos sobre el problema de la estructuración; véanse, por ejemplo,

BUCKLEY, P. S. (1964): *Techniques of Process Control*. New York: John Wiley.

SHINSKEY, F. G. (1979): *Process Control Systems* (2nd ed.). New York: McGraw-Hill.

BRISTOL, E. H. (1980): «Strategic Design: A Practical Chapter in a Textbook on Control,» Paper WA4-A, Preprints JACC, San Francisco.

No hay mucho escrito directamente sobre el problema de la estructuración. El trabajo

BUCKLEY, P. S. (1978): «Distillation Column Design Using Multivariable Control, Part 1: Process and Control Design; Part 2: Economics, Energy, and Equipment,» *Instrumentation Technology*, September, 115-22; October, 49-53.

contiene un material muy útil de interés general aunque trata un problema muy específico. El trabajo

Foss, A. S. (1973): «Critique of Chemical Process Control Theory,» *IEEE Trans. Autom. Control*, AC-18, 646-52.

tiene un enfoque más general.

Hay muy pocas áreas en las que se hayan considerado de forma conjunta el diseño del control y el del proceso. Un ejemplo notable es el diseño de aviones de elevadas prestaciones. Véanse

BURNS, B. R. A. (1975): «Fly-by-wire and Control Configured Vehicles—Rewards and Risks,» *Aeronautical Journal*, February.

BOUDREAU, J. A. (1976): «Integrated Flight Control System Design for CCV,» *Proc. AIAA Conf. on Flight Mechanics, Guidance and Control*.

Las especificaciones del comportamiento de un regulador en lazos sencillos se exponen en los textos clásicos de servomecanismos; véanse, por ejemplo,

DORF, R. C. (1980): *Modern Control Systems* (2nd ed.). Reading, Mass.: Addison-Wesley.

ELGERD, O. (1967): *Control System Theory*. New York: McGraw-Hill.

Conversión del diseño analógico

Objetivo:

Mostrar cómo se pueden obtener reguladores digitales mediante conversión y modificación del diseño analógico, incluyendo controladores digitales PID y rediseño de realimentación de estado.

8.1. Introducción

En algunas ocasiones, un sistema de control analógico se sustituye por un sistema controlado por computador simplemente porque es más fiable y barato. En este caso es normal buscar métodos de conversión de sistemas analógicos a digitales manteniendo las propiedades. Un método directo de solucionar este problema consiste en utilizar alguna aproximación discreta al controlador continuo utilizando un intervalo de muestreo corto. Este método se ilustra en el Ejemplo 1.3. La Figura 1.5 compara el controlador continuo y el controlador discreto aproximado. Obsérvese que la señal del controlador discreto se retrasa la mitad de un período de muestreo, por término medio, respecto de la señal de control continua. Este retraso origina un retardo en la fase, y deteriora las características del sistema en lazo cerrado. Por tanto, la elección del período de muestreo es importante. En general, la conversión de controladores continuos utilizando aproximaciones, conduce a períodos de muestreo bastante cortos. Sin embargo, es posible utilizar métodos digitales de diseño, que permiten períodos de muestreo mayores (compárese con la Fig. 1.6). Esto se estudia en los Capítulos 9-11.

En la Sección 8.2 se discuten algunos métodos de aproximación de controladores continuos. El controlador PID discreto se trata en la Sección 8.3. Se presentan diferentes modos de realización de controladores PID, junto a algunos aspectos operacionales tales como la saturación del término integral y la transferencia sin saltos. La Sección 8.4 muestra cómo pueden modificarse los controladores continuos

con realimentación de estado para permitir mayores períodos de muestreo, en comparación con la aproximación directa.

8.2. Métodos alternativos para aproximaciones

Esta sección supone que un controlador continuo está definido por una función de transferencia, $G(s)$. Se desea encontrar un algoritmo para computador de modo que el sistema digital se aproxime a la función $G(s)$ (véase la Fig. 8.1). Este problema es interesante para realizar controladores analógicos y filtros digitales. La aproximación puede hacerse de diferentes modos. La realización digital incluye una reconstrucción de datos, que también puede realizarse con diferentes métodos; por ejemplo, mantenedores de orden cero o primer orden. En esta sección se dan algunos métodos simples de aproximación.

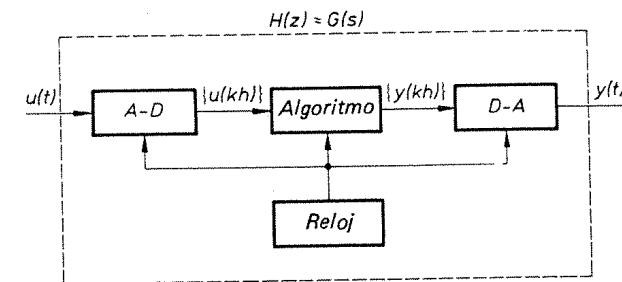


Figura 8.1 Aproximación de una función de transferencia continua, $G(s)$, mediante un computador.

Aproximación a la transformada z

Un método de aproximación de $G(s)$ consiste en utilizar los resultados del Capítulo 3 y calcular la función de transferencia discreta, $H(z)$, que corresponde a $G(s)$; es decir, muestrear el sistema. Esto puede hacerse a través de una tabla como la Tabla 3.1 o bien efectuando una realización de $G(s)$ en el espacio de estados y utilizando posteriormente (3.7) y (3.38). Entonces, la salida del sistema discreto se interpola entre los períodos de muestreo, por ejemplo, utilizando un mantenedor de orden cero. Esta solución proporciona los valores exactos de la salida en los instantes de muestreo si la entrada al filtro es constante durante los intervalos de muestreo. Este método puede funcionar bien si los cambios en la entrada durante un intervalo de muestreo son pequeños.

Aproximación en diferencias y aproximación de Tustin

Una función de transferencia representa a una ecuación diferencial. Es normal obtener una ecuación en diferencias approximando las derivadas por una diferencia adelantada (método de Euler)

$$px(t) = \frac{dx(t)}{dt} \approx \frac{x(t+h) - x(t)}{h} = \frac{q-1}{h}x(t)$$

o una diferencia retrasada

$$px(t) = \frac{dx(t)}{dt} \approx \frac{x(t) - x(t-h)}{h} = \frac{q-1}{qh}x(t)$$

En las variables transformadas esto equivale a sustituir s por $(z-1)/h$ o $(z-1)/(zh)$. En la Sección 3.6 se demuestra que las variables z y s están relacionadas en algunos aspectos por $z = \exp(sh)$. Las aproximaciones en diferencias corresponden a los desarrollos en serie

$$z = e^{sh} \approx 1 + sh \quad (\text{Método de Euler}) \quad (8.1)$$

$$z = e^{sh} \approx \frac{1}{1 - sh} \quad (\text{Diferencias en retroceso}) \quad (8.2)$$

Otra aproximación que corresponde al método trapezoidal de integración numérica es

$$z = e^{sh} \approx \frac{1 + sh/2}{1 - sh/2} \quad (\text{Método trapezoidal}) \quad (8.3)$$

En el contexto del control digital la aproximación (8.3) suele denominarse *aproximación de Tustin, o transformación bilineal*. Compárese con la transformación de Möbius de la Sección 5.2. Utilizando el método de aproximación anterior, la función de transferencia discreta $H(z)$ se obtiene simplemente sustituyendo el argumento s de $G(s)$ por s' , donde

$$s' = \frac{z-1}{h} \quad (\text{Método de Euler}) \quad (8.4)$$

$$s' = \frac{z-1}{zh} \quad (\text{Diferencias en retroceso}) \quad (8.5)$$

$$s' = \frac{2}{h} \frac{z-1}{z+1} \quad (\text{Aproximación de Tustin}) \quad (8.6)$$

Por tanto,

$$H(z) = G(s')$$

El método es muy simple de aplicar incluso manualmente.

La aproximación de Tustin tiene la ventaja de que el semiplano s izquierdo se transforma en el círculo unidad. Con el método de Euler, el semiplano s izquierdo se transforma en el semiplano $\text{Re } z < 1$. De este modo, un sistema continuo estable seguirá siendo estable después de la transformación si se aplica la aproximación de Tustin, mientras que puede approximarse por un sistema discreto inestable si se aplica el método de Euler.

Predesvío de frecuencias

Un problema de las aproximaciones anteriormente expuestas es que la escala de frecuencias se distorsiona. Por ejemplo, si se desea diseñar un filtro de banda pasante o banda eliminada, el filtro digital obtenido con las aproximaciones puede no dar las frecuencias correctas de cada banda, respectivamente. Este efecto se llama *desvío de frecuencias*. Considérese una aproximación obtenida con el método de Tustin. La transmisión de sinusoides por el filtro digital da

$$\frac{1}{i\omega h}(1 - e^{-i\omega h})H(e^{i\omega h}) = \frac{1}{i\omega h}(1 - e^{-i\omega h})G\left(\frac{2}{h} \frac{e^{i\omega h} - 1}{e^{i\omega h} + 1}\right)$$

Los dos primeros factores se deben a las operaciones del muestreador-mantenedor; compárese con (4.10). Considérese el argumento de G ; es decir,

$$\frac{2}{h} \frac{e^{i\omega h} - 1}{e^{i\omega h} + 1} = \frac{2}{h} \frac{e^{i\omega h/2} - e^{-i\omega h/2}}{e^{i\omega h/2} + e^{-i\omega h/2}} = \frac{2i}{h} \tan\left(\frac{\omega h}{2}\right)$$

Por ejemplo, si el sistema continuo no tiene transmisión de señal para una frecuencia ω —es decir, $G(i\omega) = 0$ — entonces la aproximación discreta no tiene transmisión de señal para una frecuencia ω' tal que

$$\omega = \frac{2}{h} \tan\left(\frac{\omega' h}{2}\right)$$

Esto es,

$$\omega' = \frac{2}{h} \tan^{-1}\left(\frac{\omega h}{2}\right) \approx \omega\left(1 - \frac{\omega^2 h^2}{12}\right) \quad (8.7)$$

Esta expresión da la distorsión de la escala de frecuencias (véase la Fig. 8.2). Se deduce de (8.7) que no hay distorsión de frecuencia para $\omega = 0$ y que la distorsión es pequeña si ωh es pequeña.

Es fácil introducir una transformación que elimine la distorsión de escala a una frecuencia específica ω_1 modificando la transformación de Tustin de (8.6) a la transformación

$$s' = \frac{\omega_1}{\tan(\omega_1 h/2)} \frac{z-1}{z+1} \quad (\text{Tustin con predesvío}) \quad (8.8)$$

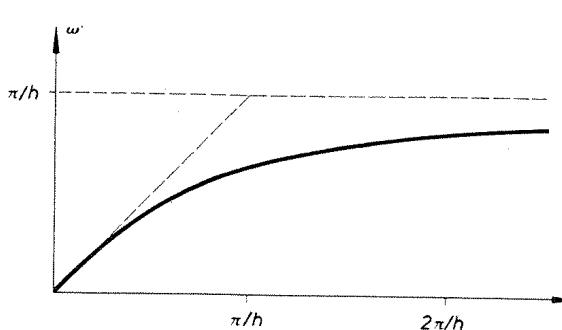


Figura 8.2 Distorsión de frecuencia (desvio) obtenida con la aproximación de Tustin.

De (8.8) se deduce que

$$H[\exp(i\omega_1 h)] = G(i\omega_1)$$

es decir, el filtro continuo y su aproximación tienen el mismo valor a la frecuencia ω_1 . Sin embargo, todavía hay distorsión a otras frecuencias.

Ejemplo 8.1 – Predesvío de frecuencias

Supongamos que el integrador

$$G(s) = \frac{1}{s}$$

debe realizarse como un filtro digital. Utilizando la transformación (8.6) sin predesvío da

$$H_T(z) = \frac{1}{\frac{2}{h} \cdot \frac{z-1}{z+1}} = \frac{h}{2} \cdot \frac{z+1}{z-1}$$

El predesvío da

$$H_p(z) = \frac{\tan(\omega_1 h/2)}{\omega_1} \cdot \frac{z+1}{z-1}$$

La función de frecuencia de H_p es

$$\begin{aligned} H_p(e^{i\omega h}) &= \frac{\tan(\omega_1 h/2)}{\omega_1} \cdot \frac{e^{i\omega h} + 1}{e^{i\omega h} - 1} \\ &= \frac{\tan(\omega_1 h/2)}{i\omega_1} \cdot \frac{1}{\tan(\omega h/2)} \end{aligned}$$

así $G(i\omega)$ y $H_p[\exp(i\omega h)]$ son iguales para $\omega = \omega_1$.

Selección del intervalo de muestreo

La elección del período de muestreo depende de muchos factores. Un método para determinar el período de muestreo consiste en utilizar argumentos continuos. El sistema de muestreo puede aproximarse por un circuito mantenedor, seguido de un

sistema continuo. Para períodos de muestreo cortos, la función de transferencia del circuito mantenedor puede aproximarse por

$$\frac{1 - e^{-sh}}{sh} \approx \frac{1 - 1 + sh - (sh)^2/2 + \dots}{sh} = 1 - \frac{sh}{2} + \dots$$

Los dos primeros términos corresponden al desarrollo en serie de $\exp(-sh/2)$. Esto es, para h pequeños, el mantenedor se puede aproximar por un retardo de tiempo de medio intervalo de muestreo. Supongamos que el margen de fase puede disminuirse de 5° a 15° . Esto proporciona la siguiente regla empírica:

$$h\omega_c \approx 0.15 - 0.5$$

donde ω_c es la frecuencia de cruce del sistema continuo en radianes por segundo. Esta regla da períodos de muestreo bastante cortos. La frecuencia de Nyquist será 5-20 veces mayor que la frecuencia de cruce.

Ejemplo 8.2 – Rediseño digital del compensador de adelanto

Considérese el sistema del Ejemplo A.2, el cual es el modelo normalizado de un motor. La función de transferencia en lazo cerrado

$$G_c(s) = \frac{4}{s^2 + 2s + 4} \quad (8.9)$$

se obtiene con el compensador de adelanto

$$G_k(s) = \frac{4s + 1}{s + 2} \quad (8.10)$$

El sistema en lazo cerrado tiene un amortiguamiento de $\zeta = 0.5$ y una frecuencia natural de $\omega_0 = 2$ rad/s. El objetivo actual es encontrar $H(z)$ de la Fig. 8.3 que se aproxime a (8.10).

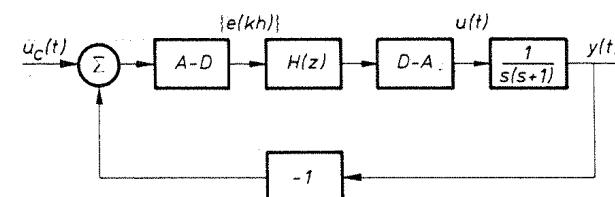


Figura 8.3 Control digital del ejemplo del motor.

El método de Euler da la aproximación

$$H_E(z) = 4 \frac{z - 1 + h}{z - 1 + 2h} = 4 \frac{z - (1 - h)}{z - (1 - 2h)} \quad (8.11)$$

mientras la aproximación de Tustin da

$$H_T(z) = 4 \frac{(2+h)z - 2 + h}{(2+2h)z - 2 + 2h} = 4 \frac{2+h}{2+2h} \cdot \frac{z - (2-h)/(2+h)}{z - (1-h)/(1+h)}$$

Finalmente, el muestreo con mantenedor de orden cero de (8.10) da

$$H_{ZOH}(z) = \frac{4z - 2(1 + e^{-2h})}{z - e^{-2h}} = 4 \frac{z - 0.5(1 + e^{-2h})}{z - e^{-2h}}$$

Todas las aproximaciones tienen la forma

$$H(z) = \frac{b_0 z + b_1}{z + a_1}$$

La frecuencia de cruce del proceso continuo situado en cascada con el compensador (8.10) es $\omega_c = 1,6$ rad/s. La regla empírica anterior da un período de muestreo entre 0,1-0,3 s.

La Figura 8.4 muestra la señal de control y la salida del proceso cuando se utiliza la aproximación de Euler para diferentes períodos de muestreo. La otra aproximación proporciona resultados similares. El sistema en lazo cerrado tiene un comportamiento satisfactorio para todos los compensadores cuando el período de muestreo es corto. La regla empírica también da valores razonables para el intervalo de muestreo. La sobreelongación para $h=0,5$ es aproximadamente el doble que para los compensadores continuos. En el ejemplo, los cambios en u_c ocurren en los instantes de muestreo. Esto no es cierto en la práctica, y puede haber un retardo en la respuesta de al menos un período de muestreo.

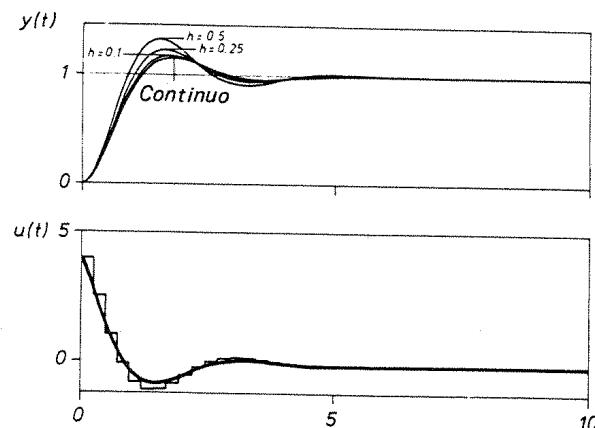


Figura 8.4 Salida del proceso, $y(t)$, cuando se controla el motor con el compensador de (8.11) cuando $h=0,1, 0,25$ y $0,5$. La señal de control corresponde a $h=0,25$. Como comparación se presentan las señales continuas.

8.3. Controladores digitales PID

Para algunas aplicaciones de control es suficiente utilizar controladores PID típicos. En esta sección se expondrán diferentes métodos de realización de controladores PID, además de algunos aspectos prácticos. En los libros de texto es

frecuente encontrar escrita la transformada de Laplace de un controlador PID continuo como

$$U(s) = K \left(1 + \frac{1}{T_I s} + \frac{T_D s}{1 + T_D s/N} \right) E(s) \quad (8.12)$$

donde $U(s)$ y $E(s)$ son las transformadas de Laplace de la salida del controlador y de la señal de error, respectivamente; K es la ganancia proporcional; T_I es la constante de integración y T_D es la constante de derivación. En el controlador hay un filtro, con la constante de tiempo T_D/N , para la parte derivadora. N está entre los valores 3-10 y suele fijarlo el fabricante del controlador.

Los métodos de aproximación para sistemas continuos de las secciones anteriores se pueden utilizar para transformar (8.12) en un controlador digital. El muestreo directo de (8.12) da

$$u(kh) = K \left(1 + \frac{\alpha}{(q-1)} + \frac{\beta(q-1)}{q+\gamma} \right) e(kh) \quad (8.13)$$

donde

$$\alpha = \frac{h}{T_I}$$

$$\beta = N$$

$$\gamma = -\exp\left(-\frac{hN}{T_D}\right)$$

También es posible aproximar el controlador PID utilizando otros métodos. El más común es realizar una aproximación de Euler para la parte integral y una aproximación en diferencias en retroceso para la parte derivadora. Esto proporciona

$$u(kh) = K \left(1 + \frac{h}{T_I(q-1)} + \frac{T_D}{h+T_D/N} \cdot \frac{q-1}{q-T_D/(Nh+T_D)} \right) e(kh) \quad (8.14)$$

Las aproximaciones (8.13) y (8.14) tienen la misma estructura principal; solamente hay una pequeña diferencia en los coeficientes. Cuando h decrece, las respuestas de (8.13) y (8.14) se aproximan más y más. Si la parte derivadora se aproxima utilizando la aproximación en diferencias adelantadas, el controlador sería inestable si $h > 2T_D/N$; es decir, no es posible utilizar $T_D = 0$. Sin embargo, las partes derivadoras de (8.13) y (8.14) son estables para todos los h posibles. A veces se utiliza la aproximación de diferencias en retroceso para la parte integral. En este caso, la única diferencia es que el retardo en el numerador de la parte integral desaparece, lo cual significa que el último error medido se utiliza en la parte integral sin retardo. Esto puede ser beneficioso si el período de muestreo es grande. Finalmente, existen muchas posibilidades para modificar la parte derivadora del controlador.

Para unificar la notación, las partes P, I y D se designan por

$$K_D, \quad \frac{K_D h}{T_I(q-1)}, \quad \text{and} \quad \frac{K_D T_D(q-1)}{h(q+\gamma)}$$

respectivamente, donde T_{ID} y T_{DD} son las constantes de tiempo discretas equivalentes a la constante de integración y la constante de derivación. Los parámetros del controlador tienen las mismas dimensiones que el controlador continuo. Por tanto, la forma típica de los libros de texto para el controlador digital PID puede escribirse en la forma

$$u(kh) = K_D \left(1 + \frac{h}{T_{ID}} \frac{1}{q - 1} + \frac{T_{DD}}{h} \frac{q - 1}{q + \gamma} \right) e(kh) \quad (8.15)$$

Los coeficientes K_D , T_{ID} y T_{DD} tendrán diferentes interpretaciones dependiendo de cómo se haya aproximado el controlador continuo. No habrá diferencia si el período de muestreo es corto.

Las expresiones (8.13) y (8.15) se denominan *formas de posición* para el controlador PID porque calculan la salida total. Si en su lugar se calcula la variación en la señal de control, $\Delta u(kh)$, la *forma de velocidad*, o *forma incremental* se determina a partir de (8.15):

$$\begin{aligned} \Delta u(kh) &= u(kh) - u(kh - h) \\ &= K_D \left[1 - q^{-1} + \frac{q^{-1}h}{T_{ID}} + \frac{T_{DD}(1 - 2q^{-1} + q^{-2})}{h(1 + \gamma q^{-1})} \right] e(kh) \end{aligned}$$

Con frecuencia, el retardo de la parte integral se elimina, lo cual corresponde a efectuar una aproximación en diferencias retardadas para la parte integral. Cuando se utiliza la forma de velocidad, el integrador se sitúa fuera del controlador. Por ejemplo, la salida pueden ser pulsos para un motor paso a paso, que está conectado a una válvula. Entonces, la parte integral la realiza el motor. Un inconveniente del algoritmo incremental es que no puede operar en el modo P o PD. Si se intenta esto, el integrador externo debe compensarse con una diferencia en la parte digital. Si se hace esto, se cancelará un modo inestable, lo cual puede producir dificultades. En la última sección se hace una nueva comparación de las dos formas.

Diferentes estructuras de los controladores PID

Existen muchas formas de modificar la estructura típica de los controladores PID de (8.12). La Figura 8.5 muestra diferentes estructuras PID, que se pueden utilizar tanto en sistemas continuos como en sistemas discretos. La estructura de la Figura 8.5b) tiene la ventaja de que el controlador no da una gran señal de control ante cambios en escalón en la señal de referencia. Esta es la estructura de controlador ante cambios en escalón en la señal de referencia. Esta es la estructura de controlador de la Figura 8.5c), es menos común. El filtro para la parte derivadora puede utilizarse de diferentes modos. El más común se muestra en (8.12). También es posible filtrar las tres partes o solamente las partes proporcional y derivadora. La última, por ejemplo, atenuará el ruido de medida de alta frecuencia.

El sintonizado de un controlador suele hacerse utilizando perturbaciones en escalón en la señal de referencia. Los valores de los parámetros adecuados para este

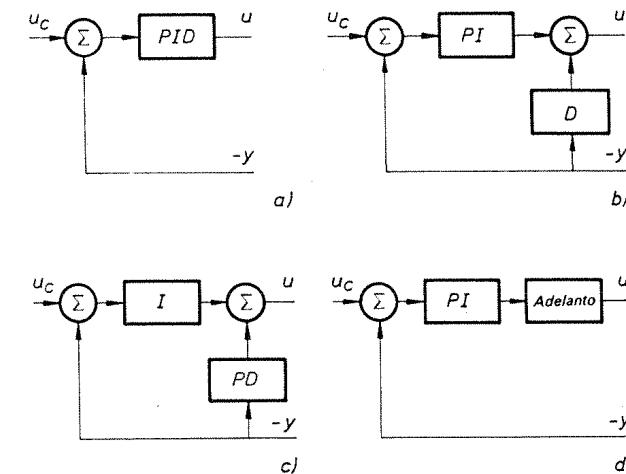


Figura 8.5 Diferentes formas de realización de controladores con funciones PID. P, I y D indican las diferentes partes del controlador.
a) Controlador de libro.
b) Controlador con derivada de la salida.
c) Controlador con integración sólo de la señal de referencia.
d) Controlador PI seguido de un circuito de adelanto.

tipo de perturbación pueden no ser buenos si las principales perturbaciones son perturbaciones del proceso. Sin embargo, debe tenerse en cuenta que la estructura de la Figura 8.5c) es la que presenta menores diferencias en los parámetros del controlador si se sintoniza para consignas o perturbaciones del proceso.

Las diferentes estructuras de la Figura 8.5 pueden reescribirse utilizando una forma común (véase la Fig. 8.6) como

$$R(q)u(kh) = T(q)u_c(kh) - S(q)y(kh)$$

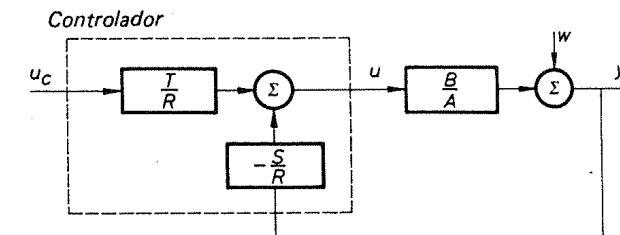


Figura 8.6 Forma general normal de los controladores PID de la Figura 8.5.

donde la interpretación de los polinomios T y S depende de la estructura. Los tres polinomios son de segundo orden y

$$R(q) = (q + \gamma)(q - 1)$$

en todos los casos. De la Figura 8.6, el sistema en lazo cerrado está dado por

$$y(kh) = \frac{TB}{AR + BS} u_c(kh) + \frac{AR}{AR + BS} w(kh)$$

Los polos en lazo cerrado pueden ser los mismos para todas las estructuras. Esto significa que los cuatro controladores se pueden sintonizar de manera que los sistemas en lazo cerrado tengan el mismo operador de transferencia discreta que relaciona las perturbaciones del proceso con la salida. Sin embargo, el polinomio T dependerá de la forma del controlador, y T introducirá dos ceros en el operador de transferencia discreta desde u_c hasta y . Los valores de los ceros dependerán de la forma del controlador y de los polinomios R y S . Para las formas de la Figura 8.5a), b) y d), T tendrá dos ceros no nulos, mientras que la forma de la Figura 8.5c) tiene un cero en el origen y otro cero no nulo. También es posible obtener un polinomio T con dos ceros en el origen. Esto puede realizarse utilizando el mismo filtro tanto en la parte integral como en la parte derivadora cuando se utiliza la estructura de la Figura 8.5c). Esta estructura resulta ventajosa si el método para sintonizar los parámetros del controlador se basa en la asignación de polos.

Aspectos operacionales

Cuando se diseñan controladores no sólo son importantes los aspectos de regulación, sino que también son importantes los aspectos operacionales como es la conmutación entre el control manual y el automático. Diferentes fabricantes de paquetes de programas DDC (del inglés Direct Digital Control) realizan sus algoritmos PID de muy diferentes modos, en parte debido a las diferentes soluciones de los aspectos operacionales. Cuando se analizan los algoritmos suelen encontrarse formas de posición o velocidad de las estructuras de la Figura 8.5.

En los paquetes de programas comerciales para DDC, es normal incluir versiones del controlador PID con diferentes no linealidades. Un ejemplo de esto lo constituye la sustitución de la parte proporcional, $k \cdot e$, por $K |e| \cdot e$.

Transferencia sin saltos. Como el regulador es un sistema dinámico, es necesario asegurarse de que el estado del sistema es correcto cuando se comutó el regulador entre el modo manual y el automático. Cuando el sistema está en modo manual, el controlador produce una señal de control que puede ser diferente de la señal de control generada manualmente. Es necesario asegurarse de que el valor del integrador es correcto en el instante de conmutación. Esto recibe el nombre de *transferencia sin saltos*.

La forma de velocidad tiene la ventaja de que no necesita inicializarse cuando se conmuta el control de manual a automático. En general, el operador fija el actuador en la posición deseada, antes de conmutar a control automático. El actuador no cambiará cuando se conmuta, hasta que ocurre un error. Para el algoritmo de posición tiene que inicializarse el valor de la salida si se quiere conseguir una transferencia sin saltos. Esto puede hacerse de varias formas; una de ellas consiste en tener una realimentación de la posición del actuador, lo que representa una solución costosa. Otro modo de lograr una transferencia sin saltos se muestra en el Ejemplo 8.3.

Ejemplo 8.3 – Controlador PI de transferencia sin saltos

La siguiente realización se puede utilizar para efectuar transferencias sin saltos en un controlador PI:

$$u(kh) = Ke(kh) + b(kh)$$

$$b(kh) = b(kh - h) + \alpha[u(kh - h) - b(kh - h)]$$

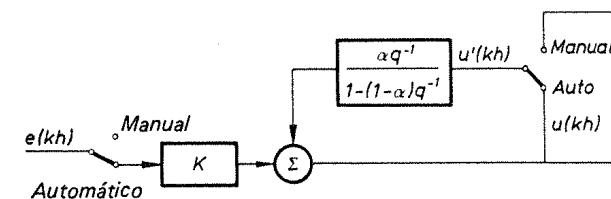


Figura 8.7 Una forma de conseguir la transferencia sin saltos con un controlador en posición.

Normalmente $u'(kh)$ es igual a $u(kh)$, pero $u'(kh)$ también puede ser la salida de una estación manual. Para comprender la acción del controlador, observemos su diagrama de bloques en la Figura 8.7. Si el controlador está en el modo manual, entonces la salida será igual a la señal manual en el estado estacionario, porque el filtro en el diagrama de bloques tiene ganancia unidad. La relación entre la salida del controlador y el error cuando $u'(kh) = u(kh)$ es

$$u(kh) = Ke(kh) + \frac{\alpha}{q - (1 - \alpha)} u(kh)$$

o bien

$$u(kh) = K \frac{q - (1 - \alpha)}{q - 1} e(kh)$$

$$= K \left[1 + \frac{\alpha}{q - 1} \right] e(kh)$$

es decir, un controlador PI.

Saturación del término integral. Un regulador con acción integral es un sistema inestable desde el punto de vista entrada-salida. Su modo inestable puede presentar

dificultades bajo determinadas circunstancias. Puede ocurrir una *saturación del integrador* si la salida se satura y el controlador continúa integrando el error. La salida del integrador puede tomar, entonces, valores muy grandes que le pueden llevar un tiempo considerablemente largo para volver otra vez a su valor normal. Este problema se elimina automáticamente cuando se utiliza un algoritmo basado en la forma de velocidad, porque la integración para automáticamente cuando la salida está limitada. Si se desea utilizar la forma de posición, deben tomarse algunas precauciones. Un método consiste en detener la actualización del integrador cuando se limita la salida. Otro método consiste en utilizar una *integración condicional*, en la cual la parte integral del controlador solamente se utiliza cuando el error es suficientemente pequeño.

Ejemplo 8.4 – Controlador PI con saturación del término integral

Un programa en lenguaje Simnon para un controlador PI discreto con saturación del término integral puede escribirse así

```
discrete system dpiaw
"Controlador PI discreto con saturación del término integral
input y uc
output u
state i
new ni
time t
tsamp ts
e=uc-y
v=k*i+e
u;if v<ulow then ulow else if v<uhigh then v else uhig
ni=itk*h/ti*e+u-v
ts=t+h
h:0.5      "Periodo de muestreo
k:0.4      "Ganancia proporcional
ti:5       "Constante de integración
ulow:-0.2   "Límite inferior de la salida
uhigh:0.2   "Límite superior de la salida
end
```

Cuando se limita la salida, la parte integral se inicia al valor *ulow* o *uhigh*, respectivamente.

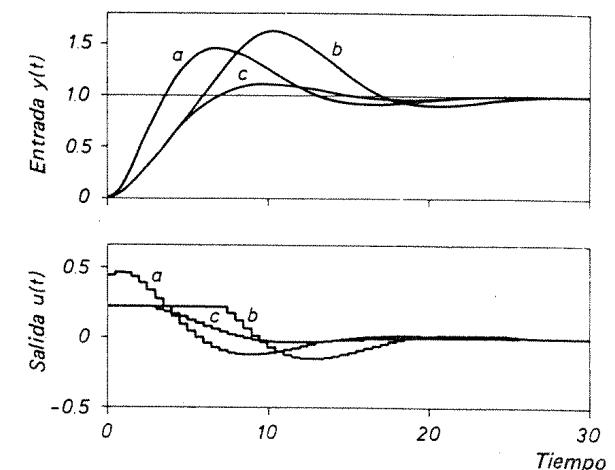
Consideremos el motor de la Figura 8.3 y sea $H(z)$ un controlador PI con salida limitada. La Figura 8.8 muestra la entrada y salida del proceso cuando se utilizan diferentes controladores.

Error de cuantificación en la integración. La parte integral en los controladores PID digitales se aproxima por una suma. Debido a la precisión finita en la representación de los números en un computador, pueden surgir problemas numéricos tales como *errores en la integración*. Supongamos que hay un error, $e(kh)$. Entonces el término integral se incrementa en cada período de muestreo en

$$\frac{K_p h}{T_{ID}} e(kh)$$

Figura 8.8 Salida y entrada del proceso cuando se controla el motor del Ejemplo 8.4 mediante diferentes controladores PI con $k = 0.4$, $t_i = 5$ y $h = 0.5$.

- a) No hay limitación en la señal de control.
- b) La señal de control está limitada a ± 0.2 sin limitación en la saturación del término integral en el controlador.
- c) La señal de control está limitada a ± 0.2 empleando el regulador con limitación en la saturación del término integral presentado en el programa Simnon dpiaw.



[véase (8.15)]. Supongamos que la ganancia es pequeña o que el tiempo de integración es grande comparando con el periodo de muestreo. Entonces, el cambio en la salida puede ser más pequeño que el paso de cuantificación del convertidor D-A. Por ejemplo, un convertidor de 12 bits D-A —es decir, una resolución de 1/4096— debería dar una resolución suficientemente buena para los propósitos del control. Además si $K_D = h = 1$ y $T_{ID} = 3.600$, cualquier error menor del 90 % del margen de variación del convertidor D-A produce un cambio en la parte integral menor que el paso de cuantificación. Si la parte integral se almacena con la misma precisión que tiene el convertidor D-A, entonces habrá un error en la salida. Un modo de evitar esto consiste en emplear mayor precisión en los cálculos internos de la parte integral. Otro procedimiento es redondear por exceso o por defecto todos los cálculos que sean menores que el nivel de cuantificación de la salida. Frecuentemente se utilizan al menos 24 bits para realizar la parte integral en un computador, con el fin de evitar el error en la integración.

Reglas de sintonía

Los controladores PID discretos tienen la ventaja de que se comportan como controladores PID continuos cuando el periodo de muestreo es corto. De este modo no hay problemas de formación de personal si se reconvierte un controlador a la forma digital, porque pueden utilizarse las mismas reglas heurísticas para sintonizar los controladores. Ziegler y Nichols (1942) dieron dos métodos de sintonía: el método de la respuesta transitoria y el método de sensibilidad final. Se supone que el controlador está parametrizado como en (8.15). Se recomienda la estructura de punto de referencia actuando solamente en la parte integral [véase la Fig. 8.5c)].

Método de la respuesta transitoria. Para utilizar el método de la respuesta transitoria, se miden la máxima pendiente, R , y el retardo, L , de la respuesta del sistema en lazo abierto ante una entrada escalón unitario (véase la Fig. 8.9). Los parámetros para el controlador PID se obtienen con ayuda de la Tabla 8.1.

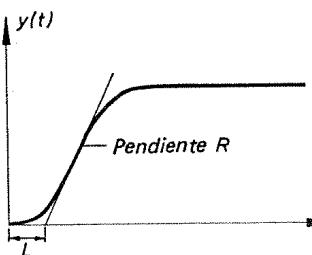


Figura 8.9 Medida de las variables R y L sobre la respuesta en lazo abierto a un salto unidad con el método de la respuesta transitoria.

TABLA 8.1 Parámetros del controlador cuando se emplea el método de la respuesta transitoria.

	K_D	T_{ID}	T_{DD}
P	$\frac{1}{RL}$		
PI	$\frac{0.9}{RL}$	$3L$	
PID	$\frac{1.2}{RL}$	$2L$	$0.5L$

Método de sensibilidad final. En el método de sensibilidad final se utiliza, en primer lugar, un controlador P para controlar el sistema. Se miden la ganancia del controlador, K_{\max} , y el período de tiempo, T_p , cuando el sistema en lazo cerrado está en el límite de estabilidad. Los parámetros del controlador se obtienen con ayuda de la Tabla 8.2.

TABLA 8.2 Parámetros del controlador cuando se emplea el método de la sensibilidad final.

	K_D	T_{ID}	T_{DD}
P	$0.5K_{\max}$		
PI	$0.45K_{\max}$	$\frac{T_p}{1.2}$	
PID	$0.6K_{\max}$	$\frac{T_p}{2}$	$\frac{T_p}{8}$

Las reglas de sintonía anteriores deberían utilizarse solamente como primeras aproximaciones. El sintonizado final suele realizarse manualmente. También existen

otros métodos para sintonizar controladores PID digitales. Algunos incluyen compensaciones para la longitud del intervalo de muestreo (véanse las referencias).

Selección del intervalo de muestreo

Cuando se introdujo el control DDC, los computadores no tenían la potencia de cálculo de los computadores actuales. Se necesitaban grandes períodos de muestreo para controlar varios lazos. Y se daban las siguientes recomendaciones para las variables de proceso más comunes para el control CDD.

Tipo de variable	Período de muestreo en s
Flujo	1-3
Nivel	5-10
Presión	1-5
Temperatura	10-20

Los controladores digitales comerciales para pocos lazos suelen tener intervalos de muestreo fijos del orden de los 200 ms. Esto implica que los controladores se pueden considerar como continuos, y por tanto, se pueden aplicar las reglas para controladores continuos.

En la literatura se pueden encontrar varias reglas que relacionan el intervalo de muestreo y el tiempo de derivación. Es típica la sugerencia de que el período de muestreo se elija de modo que

$$\frac{h}{T_{DD}} \approx 0.1 - 0.5$$

Para las reglas de sintonía de Ziegler-Nichols, esto implica que

$$\frac{h}{L} \approx 0.05 - 0.25$$

o

$$\frac{h}{T_p} \approx 0.01 - 0.05$$

Con un experimento simple sobre el proceso, es posible comprobar si el período de muestreo es correcto.

Predictor de Smith

El predictor de Smith se introdujo en la Sección 7.5. Un inconveniente de la versión continua del predictor de Smith es que contiene un retardo. Esto no causará problemas cuando se utilice la versión digital, porque el retardo de tiempo se realiza

como un vector que se desplaza en cada período de muestreo. Un ejemplo simple ilustra el diseño de un predictor de Smith.

Ejemplo 8.5 – Predictor de Smith

En el Ejemplo A.4 se describe un proceso con retardo. El proceso puede representar, por ejemplo, una máquina de papel. Supongamos que el proceso de (A.9) tiene un retardo de dos unidades de tiempo y que el período de muestreo es $h=1$. Entonces, el sistema se describe por el modelo

$$y(k+1) = 0.37y(k) + 0.63u(k-2)$$

(véase el Ejemplo 3.6). Si no hubiese retardos, un regulador PI con ganancia 0,4 y tiempo de integración $T_i=0,4$ sería un buen controlador (véase la Fig. 8.10). El punto de referencia

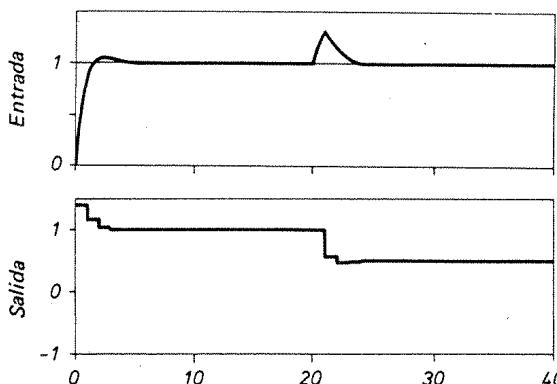


Figura 8.10 Control PI del proceso del Ejemplo 8.5 sin retardo.

se cambia en $t=0$ y se introduce una perturbación en escalón en la salida en $t=20$. Este regulador PI no realizará bien sus funciones si el proceso tiene un retardo. Para obtener una buena regulación PI, es necesario tener una ganancia de 0,1 y $T_i=0,5$. La respuesta a este regulador se ilustra en la Figura 8.11. Comparando con la Figura 8.10 se muestra que la respuesta es más lenta porque la ganancia es más pequeña.

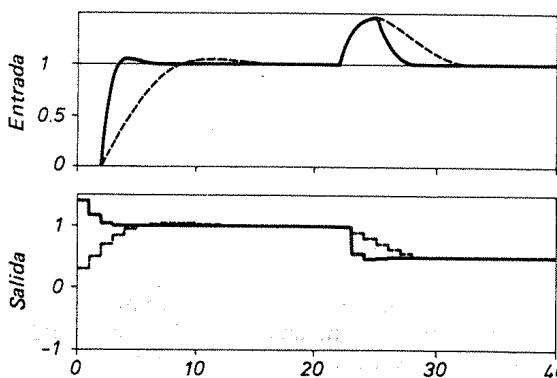


Figura 8.11 Control PI (línea a trazos) y control con predictor de Smith (línea continua) del proceso del Ejemplo 8.5 con retardo.

Si se utiliza el controlador PI como G , y el modelo muestreado como G_p , puede describirse el regulador con el siguiente programa en lenguaje Simonon:

```
Discrete system Smith
"Predictor de Smith para sistema de primer orden con retardo
input y uc
output u
state i u1 ym ym1 ym2
new ni nu1 nym nym1 nym2
time t
tsamp ts
e=uc-y+ym2-ym
u=k*(e+e*h/ti+i)
ni=i+e*h/ti
nym=am*y+bm1*u+bm2*u1 "modelo del proceso simulado
nu1=u
nym1=ym
nym2=ym1
ts=t+h
t1:1 "periodo de muestreo
ti:0.4 "tiempo de integración
k:0.4 "ganancia
am:0.37 "
bm1:0.63 "parámetros del proceso
bm2:0 "
end
```

Obsérvese lo fácil que es realizar el predictor para control con computador. La respuesta del sistema en lazo cerrado obtenida con el predictor de Smith se muestra en la Figura 8.11, la cual muestra la mejora obtenida frente al controlador PI. Puede observarse que la señal de control en el comienzo del escalón es idéntica a la señal de control con el controlador PI para un proceso sin retardo.

8.4. Rediseño de la realimentación de estado

En esta sección se traduce la realimentación de estados de los controladores continuos a los controladores discretos. Los controladores de realimentación de estado se pueden considerar como controladores P generalizados. La formulación del problema supone que el proceso está definido por las ecuaciones

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \quad (8.16)$$

donde se supone que todos los estados son medibles. Utilizando un controlador de la forma

$$u(t) = Mu_c(t) - Lx(t) \quad (8.17)$$



es posible situar arbitrariamente los polos del sistema en lazo cerrado siempre que el sistema sea controlable. El controlador de (8.17) puede realizarse de forma digital muestreando los estados y manteniendo constante la señal de control durante los intervalos de muestreo. Así es como se hace el control en el Ejemplo 1.3. Si el período de muestreo se incrementa, entonces el comportamiento del sistema en lazo cerrado comienza a deteriorarse. Sin embargo, es posible modificar el controlador para mejorar el rendimiento del sistema en lazo cerrado. Supongamos que el controlador discreto es

$$u(kh) = \tilde{M}u_e(kh) - \tilde{L}x(kh) \quad (8.18)$$

Un modo de solucionar el problema es diseñar el controlador de (8.18) utilizando la teoría de muestreo. Esto se hace en el Capítulo 9. Aquí se utiliza un método aproximado para convertir el controlador (8.18) a forma discreta.

Controlando (8.16) con el controlador continuo de (8.17) se obtiene el sistema en lazo cerrado

$$\begin{aligned}\dot{x} &= (A - BL)x + BMu_e = A_c x + BMu_e \\ y &= Cx\end{aligned}$$

Si $u_e(t)$ es constante sobre un período de muestreo, esta ecuación puede integrarse; esto da

$$x(kh + h) = \Phi_e x(kh) + \Gamma_e M u_e(kh) \quad (8.19)$$

donde

$$\begin{aligned}\Phi_e &= \exp(A_e h) \\ \Gamma_e &= \int_0^h \exp(A_e s) ds B\end{aligned}$$

Si se utiliza el controlador discreto (8.18) para controlar (8.16), entonces

$$x(kh + h) = (\Phi - \Gamma \tilde{L})x(kh) + \tilde{M}u_e(kh) \quad (8.20)$$

donde Φ y Γ son las matrices del sistema obtenidas muestreando (8.16). En general, no es posible elegir \tilde{L} tal que

$$\Phi_e = \Phi - \Gamma \tilde{L}$$

Sin embargo, se puede hacer un desarrollo en serie e igualar los términos de las diferentes potencias de h . Supongamos que

$$\tilde{L} = L_0 + L_1 h/2$$

entonces

$$\Phi_e \approx I + (A - BL)h + [A^2 - BLA -ABL + (BL)^2]h^2/2 + \dots$$

y

$$\Phi - \Gamma \tilde{L} \approx I + (A - BL_0)h + (A^2 - AB L_0 - AB L_1)h^2 + \dots$$

Los sistemas (8.19) y (8.20) tienen los mismos polos hasta el orden h^2 inclusive cuando

$$\tilde{L} = L[I + (A - BL)h/2] \quad (8.21)$$

Sin modificaciones de L los polos son los mismos hasta el orden h inclusive.

La modificación de M se determina suponiendo que los valores en estado estacionario son iguales para (8.19) y (8.20). Consideremos un valor de referencia constante y supongamos que el valor del estado en el estacionario es x^0 . Esto produce las relaciones

$$[I - \Phi_e]x^0 = \Gamma_e M u_e$$

y

$$[I - (\Phi - \Gamma \tilde{L})]x^0 = \Gamma \tilde{M} u_e$$

Los desarrollos en serie de los términos de la izquierda de las relaciones anteriores son iguales para potencias de h hasta h^2 inclusive. Ahora determinemos \tilde{M} de modo que los desarrollos en serie de los términos de la derecha sean iguales para h y h^2 . Supongamos que

$$\tilde{M} = M_0 + M_1 h/2,$$

entonces

$$\Gamma_e M \approx BMh + (A - BL)BMh^2/2 + \dots$$

y

$$\Gamma \tilde{M} \approx BM_0 h + (BM_1 + ABM_0)h^2/2 + \dots$$

Esto proporciona

$$\tilde{M} = (I - LBh/2)M \quad (8.22)$$

Las modificaciones de (8.21) y (8.22) se pueden calcular fácilmente del sistema continuo y del controlador continuo.

Ejemplo 8.6

El sistema del Ejemplo 1.3 es el integrador doble; es decir, el sistema está definido por las matrices

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad y \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

Consideremos el controlador continuo

$$u(t) = u_e(t) - [1 \ 1]x(t)$$

La Figura 8.12 muestra el comportamiento cuando el controlador muestreado

$$u(kh) = u_e(kh) - [1 \ 1]x(kh) \quad (8.23)$$

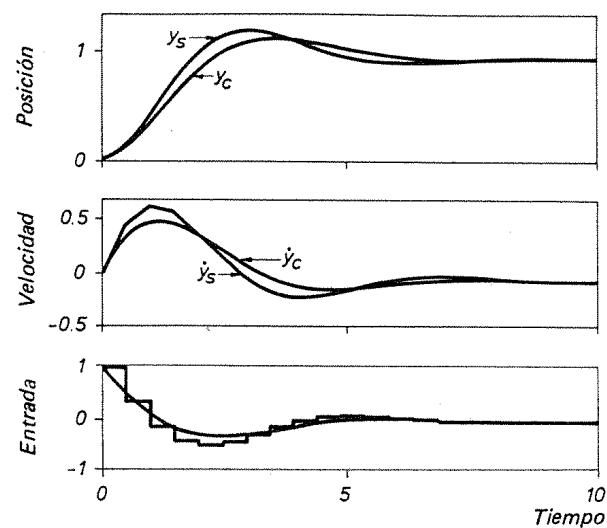


Figura 8.12 Control digital del integrador doble mediante la ley de control dada en (8.23) cuando $h=0,5$; y_c es la respuesta continua.

se utiliza para $h=0,5$. Utilizando las modificaciones de (8.21) y (8.22), obtenemos

$$\begin{aligned} L &= [1 - 0.5h \quad 1] \\ \tilde{M} &= 1 - 0.5h \end{aligned} \quad (8.24)$$

La Figura 8.13 muestra el comportamiento cuando el controlador modificado se utiliza para $h=0,5$; hay mejoras respecto al controlador no modificado. Sin embargo, el periodo de muestreo no puede incrementarse sin que el comportamiento del sistema en lazo cerrado comience a deteriorarse, incluso cuando se utiliza el controlador modificado.

8.5. Métodos de diseño basados en la respuesta en frecuencia

Hasta ahora este capítulo ha presentado cómo se pueden pasar los controladores continuos a discretos. Esta sección estudia cómo se pueden utilizar los métodos de diseño basados en la frecuencia para diseñar controladores discretos.

Los métodos de diseño de frecuencia basados en los diagramas de Bode y Nichols son útiles en el diseño de compensadores para sistemas descritos por funciones de transferencia. La utilidad de los métodos depende de la simplicidad de la representación de Bode y de las reglas de elección de compensadores. El diagrama de Bode es fácil de dibujar si las funciones de transferencia están expresadas como funciones de $i\omega$, excepto para los retardos puros.

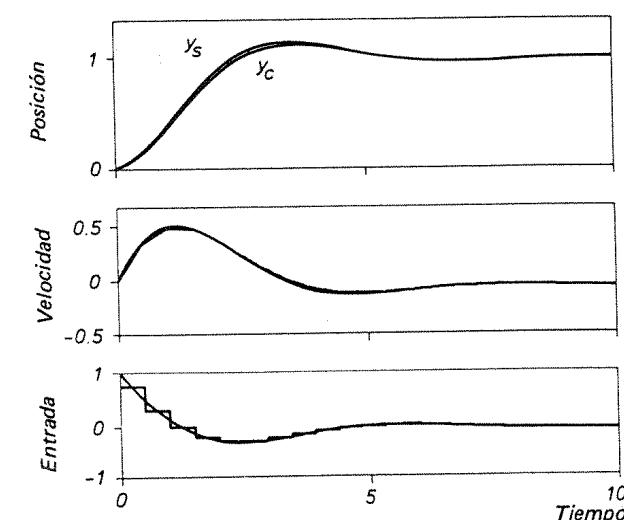


Figura 8.13 Control del integrador doble con el controlador modificado de (8.24) cuando $h=0,5$; y_s es la respuesta continua.

Las curvas de frecuencia para los sistemas discretos son más difíciles de dibujar, ya que las funciones de transferencia discreta no son funciones racionales en $i\omega$, sino en $\exp(i\omega h)$. El *método de transformada en w* es un modo de evitar esta dificultad. El método puede resumirse en los siguientes pasos:

1. Muestrear el sistema continuo que debería controlarse utilizando un circuito mantenedor de orden cero. Esto da $H(z)$.
2. Introducir la variable

$$w = \frac{2}{h} \frac{z - 1}{z + 1}$$

[compárese con (8.6)]. Transformar la función de transferencia discreta del proceso al plano w dando

$$H'(w) = H(z) \Big|_{z=\frac{1+wh/2}{1-wh/2}}$$

Para $z = \exp(i\omega h)$ se tiene

$$w = i \frac{2}{h} \tan(\omega h/2) = iv$$

(compárese con el predesvio de frecuencias de la Sección 8.2). La función de transferencia transformada $H'(iv)$ es una función racional en iv .

3. Dibujar el diagrama de Bode de $H'(iv)$ y utilizar métodos convencionales para diseñar un compensador $H_r(iv)$ que proporcione las características deseadas en el dominio de la frecuencia. La distorsión entre las escalas de v y ω debe tenerse en cuenta cuando se decide, por ejemplo, la frecuencia de cruce y el ancho de banda.
4. Transformar al plano z el compensador y realizar $H_r(z)$ como un sistema discreto.

La ventaja del método de la transformada w está en que pueden utilizarse las técnicas convencionales del diagrama de Bode para efectuar el diseño. Las dificultades están en manipular la distorsión de la escala de frecuencias y elegir el intervalo de muestreo.

8.6. Conclusión

Son necesarios diferentes métodos de conversión del control analógico de sistemas al control digital para acercar los universos analógico y digital. El problema tiene un interés sustancial si se dispone de un diseño analógico y se necesita encontrar una solución digital. El método es extremadamente simple: el algoritmo que se va a codificar en el computador se describe por su función de transferencia discreta, la cual se obtiene por sustitución del argumento s de la función de transferencia continua utilizando (8.4) a (8.6) u (8.8). El método de Tustin es una buena elección, pero también pueden utilizarse aproximaciones diferentes. Los sistemas digitales diseñados de este modo son siempre (ligeramente) inferiores a los sistemas analógicos debido al retardo de tiempo introducido por el circuito mantenedor. Este retardo es aproximadamente $h/2$.

Los métodos de conversión trabajan bien si los períodos de muestreo son cortos. Un buen método para elegir los períodos de muestreo consiste en observar que el retardo de tiempo extra decrementa el margen de fase en $\omega_c h/2$ radianes o $180 \omega_c / \omega_s$ grados, donde ω_c es la frecuencia de cruce.

Existe la posibilidad de encontrar mejores diseños que los discutidos en este capítulo, los cuales se discutirán en capítulos próximos.

8.7. Problemas

- 8.1. Encontrar cómo se transforma el semiplano izquierdo s en el plano z cuando se utilizan las aplicaciones (8.4)-(8.6).
- 8.2. Efectuar una aproximación de la función de transferencia

$$G(s) = \frac{a}{s+a}$$

utilizando a) el método de Euler, b) la aproximación de Tustin y c) aproximación de Tustin con predesvío de frecuencias si la frecuencia de desvío es $\omega_1 = a$ rad/s.

- 8.3. La red de adelanto (8.10) produce unos 20° de avance de fase a $\omega_c = 1,6$ rad/s. Aproximar la red para $h = 0,25$ utilizando a) el método de Euler, b) las diferencias de retroceso, c) la aproximación de Tustin, d) la aproximación de Tustin con predesvío utilizando $\omega_1 = \omega_c$ como frecuencia de desvío y e) el muestreador de orden cero.
Calcular la fase de la red aproximada en $z = \exp(i\omega_c h)$.
- 8.4. Verificar los cálculos de la Sección 8.2 que conducen a la regla empírica para la elección del intervalo de muestreo.
- 8.5. Demostrar que (8.14) se obtiene a partir de (8.12) approximando la parte integral con el método de Euler y la parte derivadora con diferencias de retroceso. Discutir las ventajas y desventajas de los siguientes casos.
 - a) La parte integral se aproxima utilizando diferencias retrasadas.
 - b) La parte derivadora se aproxima utilizando el método de Euler. (Sugerencia: Considerar el caso en el que T_D sea pequeño.)
- 8.6. Un controlador PI continuo está definido por la función de transferencia

$$K \left(1 + \frac{1}{T_I s} \right)$$

Utilizar la aproximación bilineal para encontrar un controlador discreto. Encontrar la relación entre los parámetros K y T_I del sistema continuo y los correspondientes parámetros del sistema discreto de (8.15).

- 8.7. Considerar el sistema del depósito del Problema 3.10. Suponer las siguientes especificaciones para el sistema en lazo cerrado:
 1. El error en estado estacionario después de un salto escalón en la referencia es cero.
 2. La frecuencia de cruce del sistema compensado es 0,025 rad/s.
 3. El margen de fase es aproximadamente 50° .
 - a) Diseñar un controlador PI que verifique las especificaciones.
 - b) Determinar los polos y ceros del sistema en lazo cerrado. ¿Cuál es el amortiguamiento correspondiente a los polos complejos?
 - c) Elegir un intervalo de muestreo apropiado y aproximar el controlador continuo empleando el método de Tustin con desvío de frecuencias. Utilizar la frecuencia de cruce como frecuencia de desvío.
 - d) Simular el sistema cuando se utiliza el controlador muestreado. Comparar con la respuesta deseada, es decir, cuando se utiliza el controlador continuo.
- 8.8. Efectuar una aproximación analógica a (8.21) y (8.22) tal que las modificaciones sean válidas hasta los términos h^3 inclusive.
- 8.9. El motor normalizado tiene una representación en el espacio de estados dada por (A.5). La ley de control

$$u(t) = Mu_c(t) - Lx(t)$$

con $M = 4$ y $L = [2 \quad 4]$ da la función de transferencia continua

$$\frac{4}{s^2 + 3s + 4}$$

entre la entrada u_c y la salida y . Esto corresponde a $\zeta = 0,75$ y $\omega_0 = 2$.



- a) Realizar una versión muestrada del controlador anterior.
 b) Modificar la ley de control utilizando (8.21) y (8.22).
 c) Simular los controladores de a) y b) para diferentes períodos de muestreo y comparar con el controlador continuo.
- 8.10. Dado el sistema continuo

$$\frac{dx}{dt} = \begin{bmatrix} -3 & 1 \\ 0 & -2 \end{bmatrix}x + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u$$

$$y = [1 \ 0]x$$

- a) Determinar un controlador continuo con realimentación de estado

$$u(t) = -Lx(t)$$

tal que el polinomio característico del sistema en lazo cerrado sea

$$s^2 + 8s + 32$$

Entonces, utilizar un computador para realizar el controlador del tipo

$$u(kh) = -Lx(kh)$$

- b) Modificar el controlador utilizando (8.21).
 c) Simular los controladores de a) y b) y elegir un intervalo de muestreo apropiado. Suponer que $x(0) = [1 \ 0]^T$.
- 8.11. Utilizar el método del plano w para diseñar un compensador para el motor del Ejemplo 8.2 cuando $h=0.25$. Diseñar un compensador para que el sistema transformado tenga una frecuencia de corte de 1,4 rad/s y un margen de fase de 50 grados. Comparar con el diseño continuo y la aproximación discreta del Ejemplo 8.2. Estudiar cuáles son los mayores intervalos de muestreo que pueden utilizarse para el método del plano w .

8.8. Referencias

El problema del diseño de filtros digitales que se aproximen a funciones de transferencia analógicas se trata en la literatura sobre filtros digitales

- RABINER, L. R., y B. GOLD (1975): *Theory and Application of Digital Signal Processing*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.
 OPPENHEIM, A. V., y R. W. SCHAFER (1975): *Digital Signal Processing*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.
 ANTONIOU, A. (1979): *Digital Filters: Analysis and Design*. New York: McGraw-Hill.

Revisiones interesantes de las similitudes y diferencias entre el procesamiento digital de señales y la teoría de control se presentan en

- WILSKY, A. S. (1979): *Digital Signal Processing and Control and Estimation Theory*. Cambridge, Mass.: MIT Press.

Una presentación más orientada al control de la aproximación en diferencias se encuentra en

- FRANKLIN, G. F., y J. D. POWELL (1980): *Digital Control of Dynamic Systems*. Reading, Mass.: Addison-Wesley.

Los controladores digitales PID y sus aspectos operacionales se discuten en
 GOFF, K. W. (1966): «A Systematic Approach to DDC Design», *ISA Journal*, Diciembre, 44-54.
 BRISTOL, E. H. (1977): «Design and Programming Control Algorithms for DDC Systems», *Control Engineering*, Enero, 24-26.
 SHINSKEY, F. G. (1979): *Process Control Systems* (2ed.). New York: McGraw-Hill.

La referencia básica para la sintonía de los controladores PID es

- ZIEGLER, J. G., y N. B. NICHOLS (1942): «Optimum Settings for Automatic Controllers». *Trans. ASME*, 64, 759-68.

Una modificación de las reglas de Ziegler y Nichols que tiene en cuenta el intervalo de muestreo se dan en

- TAKAHASHI, Y., C. S. CHAN, y D. M. AUSLANDER (1971): «Parametereinstellung bei linearen DDC-Algorithmen». *Regelungstechnik und Prozess-Datenverarbeitung*, 19, 237-44.

El predictor de Smith se describió en primer lugar en

- SMITH, O. J. M. (1957): «Closer Control of Loops with Deadtime», *Chem. Eng. Progr.*, 53, 217-19.

El rediseño de realimentación de estado se discute con detalle en

- KUO, B. C. (1980): *Digital Control Systems*. Tokyo: Holt-Saunders.

Métodos de diseño en el espacio de estados

Objetivo:

Desarrollar técnicas de diseño basadas en el enfoque del espacio de estados.

9.1. Introducción

Este capítulo presenta los métodos de diseño basados en modelos internos del sistema. A la luz de la exposición del Capítulo 7, los métodos presentados en este capítulo pueden verse como soluciones a problemas de control específicos idealizados. Las soluciones dan un conocimiento profundo de la naturaleza de los problemas de control. Este capítulo también muestra la utilidad de muchos de los conceptos introducidos anteriormente. Cuando se aplican estos métodos, siempre se ponen en juego juicios subjetivos. Esto queda expresado formalmente con los llamados parámetros de diseño que debe elegir el diseñador.

En la Sección 9.2 se presenta la realimentación de estados. La Sección 9.3 expone cómo obtener las variables de estado que no son medibles directamente. Esto conduce a la introducción de los observadores. El problema de la realimentación de la salida se resuelve en la Sección 9.4 combinando la realimentación de estado y los observadores. La Sección 9.5 trata el problema del servo, es decir, cómo introducir las señales de control. El capítulo se resume en la Sección 9.6.

9.2. Regulación basada en la asignación de polos por realimentación del estado

En esta sección se presenta un problema de regulación sencillo. Se desarrolla uno de los métodos fundamentales de diseño, la asignación de polos. Este método de

diseño puede verse como una extensión del método clásico del lugar de las raíces. El propósito es proponer una realimentación de forma que todos los polos del sistema en lazo cerrado tomen valores prefijados. En esta sección se resuelve este problema bajo la hipótesis restrictiva de que todas las variables de estado pueden medirse directamente. Esta restricción se levanta en la Sección 9.3.

Formulación del problema

Para especificar el problema de diseño, deben darse el proceso, las perturbaciones, el criterio y las señales de control admisibles.

El proceso. Se supone que el proceso a controlar puede describirse por el modelo

$$\frac{dx}{dt} = Ax + Bu \quad (9.1)$$

donde u representa las variables de control, x representa el vector de estado, y A y B son matrices constantes. Además sólo se va a presentar el caso de una entrada y una salida. Como estamos considerando el caso del control con computador, las señales de control serán constantes durante los períodos de muestreo, de longitud constante. El muestreo del sistema (9.1) da el sistema discreto

$$x(kh + h) = \Phi x(kh) + \Gamma u(kh) \quad (9.2)$$

donde las matrices Φ y Γ vienen dadas por

$$\Phi = e^{Ah}$$

$$\Gamma = \left(\int_0^h e^{As} ds \right) B$$

(véase la Sección 3.2).

Las perturbaciones. Se supone que las perturbaciones son impulsos que ocurren de forma irregular, pero que el espacio entre los impulsos es tan largo que el sistema puede recuperarse y volver al estacionario después de una perturbación antes de que se presente el siguiente impulso. Como el efecto de un impulso consiste solamente en cambiar el estado del proceso, se puede representar la perturbación por un estado inicial.

El criterio. El problema es un tema de regulación, en el que se intenta llevar el estado a cero después de una perturbación en las condiciones iniciales. En la formulación de la asignación de polos, se da indirectamente la velocidad de anulación del estado especificando los polos del sistema en lazo cerrado.

Controles admisibles. Puesto que se desean soluciones de realimentación, hay que especificar la información disponible para generar la señal de control. El sistema

en lazo cerrado debe ser lineal porque las propiedades del sistema se especifican por los polos en lazo cerrado. Entonces la realimentación también debe ser lineal. Se supone que todas las variables de estado pueden medirse directamente, con lo que los controles admisibles pueden expresarse como una realimentación lineal:

$$u(kh) = -Lx(kh) \quad (9.3)$$

Parámetros de diseño. En la especificación formal del problema, los parámetros de diseño son el período de muestreo y los polos en lazo cerrado deseados. No es corriente que el usuario de un sistema de control sea capaz de dar las especificaciones en términos de estos parámetros. Por tanto, el diseñador debe ser capaz de relacionar los parámetros de diseño con las cantidades que son más significativas para el usuario. Para esto a menudo es útil considerar la historia temporal de las variables de estado y de las variables de control. Es particularmente útil exponer los compromisos entre la magnitud de las señales de control y la velocidad a la que el sistema se recupera de una perturbación (véase la exposición en la Sección 7.6).

Un ejemplo

Para introducir el método de diseño e ilustrar la influencia de los parámetros de diseño, se va a exponer en primer lugar un caso especial.

Ejemplo 9.1 – Asignación de polos para un sistema con un integrador doble

La planta del integrador doble muestreado se representa por

$$x(kh + h) = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} x(kh) + \begin{bmatrix} h^2/2 \\ h \end{bmatrix} u(kh) \quad (9.4)$$

Una realimentación lineal de tipo general se representa por

$$u = -l_1 x_1 - l_2 x_2 \quad (9.5)$$

Con esta realimentación, el sistema en lazo cerrado resulta

$$x(kh + h) = \begin{bmatrix} 1 - l_1 h^2/2 & h - l_2 h^2/2 \\ -l_1 h & 1 - l_2 h \end{bmatrix} x(kh) \quad (9.6)$$

La ecuación característica del sistema en lazo cerrado es

$$z^2 + \left(\frac{l_1 h^2}{2} + l_2 h - 2\right)z + \left(\frac{l_1 h^2}{2} - l_2 h + 1\right) = 0$$

Supongamos que la ecuación característica deseada es

$$z^2 + p_1 z + p_2 = 0 \quad (9.7)$$

Esto conduce a las siguientes ecuaciones lineales para l_1 y l_2 :

$$\frac{l_1 h^2}{2} + l_2 h - 2 = p_1$$

$$\frac{l_1 h^2}{2} - l_2 h + 1 = p_2$$

Estas ecuaciones tienen la solución

$$\begin{cases} l_1 = \frac{1}{h^2}(1 + p_1 + p_2) \\ l_2 = \frac{1}{2h}(3 + p_1 - p_2) \end{cases} \quad (9.8)$$

El caso general

Ahora se da la solución del problema de asignación de polos en el caso general para sistemas con una sola señal de entrada.

Sea el sistema descrito por la ecuación (9.2), donde u es un escalar. Supongamos que el sistema muestreado es alcanzable. Esto implica que el sistema (9.2) puede transformarse a la forma canónica controlable:

$$z(kh + h) = \begin{bmatrix} -a_1 & -a_2 & \cdots & -a_{n-1} & -a_n \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} z(kh) + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ u(kh) \\ \vdots \\ 0 \end{bmatrix} \quad (9.9)$$

donde

$$z^n + a_1 z^{n-1} + \cdots + a_n$$

es el polinomio característico de la matriz Φ . Cuando las ecuaciones del sistema están en esta forma, el problema de la asignación de polos es fácil de resolver. Supongamos que se desea obtener un sistema en lazo cerrado con el polinomio característico

$$P(z) = z^n + p_1 z^{n-1} + \cdots + p_n \quad (9.10)$$

De (9.9) se deduce que la ley de realimentación

$$u = -\tilde{L}z \quad (9.11)$$

donde

$$\tilde{L} = [p_1 - a_1 \ p_2 - a_2 \ \cdots \ p_n - a_n] \quad (9.12)$$

da un sistema en lazo cerrado con el polinomio característico deseado. La ley de realimentación puede escribirse como

$$u = -\tilde{L}z = -\tilde{L}Tx = -Lx$$

donde T define la transformación de x a z . Queda por determinar la matriz de transformación T . Sean W_c y \tilde{W}_c las matrices de controlabilidad de (9.2) y (9.9). Estas están relacionadas por la expresión

$$\tilde{W}_c = TW_c$$

Además,

$$\tilde{W}_c^{-1} = \begin{bmatrix} 1 & a_1 & \cdots & a_{n-1} \\ 0 & 1 & \cdots & a_{n-2} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (9.13)$$

(compárese con la Sección 5.3 y el Ejemplo 5.6). La matriz de transformación T puede expresarse por

$$T = \tilde{W}_c W_c^{-1} \quad (9.14)$$

Empleando (9.10) puede obtenerse la matriz polinómica

$$\begin{aligned} P(\tilde{\Phi}) &= \tilde{\Phi}^n + p_1 \tilde{\Phi}^{n-1} + \cdots + p_n I \\ &= (p_1 - a_1) \tilde{\Phi}^{n-1} + \cdots + (p_n - a_n) I \end{aligned}$$

donde $\tilde{\Phi}$ es la matriz del sistema transformado (9.9). La segunda igualdad se obtiene empleando el teorema de Cayley-Hamilton. Además la última fila de $\tilde{\Phi}^k$ es cero excepto en la posición $n-k$, que es 1; por tanto, de (9.12),

$$\tilde{L} = [0 \ \dots \ 0 \ 1]P(\tilde{\Phi})$$

Pero

$$\tilde{\Phi} = T\Phi T^{-1}$$

y entonces

$$\begin{aligned} L &= \tilde{L}T = [0 \ \dots \ 0 \ 1]P(T\Phi T^{-1})T \\ &= [0 \ \dots \ 0 \ 1]TP(\Phi) \end{aligned}$$

La matriz T , sin embargo, está dada por (9.14). Además, de (9.13)

$$[0 \ \dots \ 0 \ 1]\tilde{W}_c = [0 \ \dots \ 0 \ 1]$$

con lo que se obtiene que

$$L = \tilde{L}\tilde{W}_c W_c^{-1} = [0 \ \dots \ 0 \ 1]W_c^{-1}P(\Phi) \quad (9.15)$$

La ecuación (9.15) a veces recibe el nombre de *fórmula de Ackermann*. El resultado se resume en el Teorema 9.1.

Teorema 9.1. Consideremos el sistema dado en (9.2) con una entrada. Existe una realimentación que da un sistema en lazo cerrado con polos especificados por $P(z) = 0$ si y sólo si el sistema muestrado es alcanzable. La realimentación es

$$u(kh) = -Lx(kh)$$

[ecuación (9.3)], donde

$$L = \tilde{L}\tilde{W}_c W_c^{-1} = [0 \ \dots \ 0 \ 1]W_c^{-1}P(\Phi) \quad (9.15)$$

\tilde{L} viene dada por (9.12), y W_c y \tilde{W}_c son las matrices de controlabilidad del sistema en (9.2) y (9.9), respectivamente. \square

Nota 1. La construcción de un contraejemplo muestra que la alcanzabilidad es una condición necesaria.

Nota 2. Nótese que el problema de asignación de polos puede formularse como el problema abstracto de seguimiento. Dadas las matrices Φ y Γ , encontrar una matriz L tal que la matriz $\Phi - \Gamma L$ tenga unos valores propios prefijados.

Nota 3. Nótese que de (9.13), (9.14), y de la definición de W_c se deduce que

$$T^{-1} = [\Gamma \ \Phi\Gamma + a_1\Gamma \ \dots \ \Phi^{n-1}\Gamma + a_1\Phi^{n-2}\Gamma + \cdots + a_{n-1}\Gamma]$$

El teorema se ilustra con un ejemplo.

Ejemplo 9.2

Consideremos el sistema con el integrador doble del Ejemplo 9.1. Supongamos que el polinomio característico deseado es el dado por (9.7). Tenemos

$$W_c = [\Gamma \ \Phi\Gamma] = \begin{bmatrix} h^2/2 & 3h^2/2 \\ h & h \end{bmatrix}$$

y el polinomio característico de Φ es

$$z^2 - 2z + 1$$

Por tanto,

$$W_c^{-1} = \begin{bmatrix} -1/h^2 & 1.5/h \\ 1/h^2 & -0.5/h \end{bmatrix}$$

Tenemos

$$P(\Phi) = \Phi^2 + p_1\Phi + p_2I = \begin{bmatrix} 1 + p_1 + p_2 & 2h + p_1h \\ 0 & 1 + p_1 + p_2 \end{bmatrix}$$

Ahora la fórmula (9.15) nos da

$$\begin{aligned} L &= [0 \ 1]W_c^{-1}P(\Phi) = [1/h^2 \ -0.5/h]P(\Phi) \\ &= \begin{bmatrix} 1 + p_1 + p_2 & 3 + p_1 - p_2 \\ h^2 & 2h \end{bmatrix} \end{aligned}$$

que es el mismo resultado que el obtenido por cálculo directo en el Ejemplo 9.1 [compárese con (9.8)].

Aspectos prácticos

Es fácil resolver el problema del diseño de asignación de polos explicitamente. Nótese que la alcanzabilidad es una condición necesaria y suficiente para resolver el problema. Para aplicar en la práctica el método del diseño con asignación de polos es necesario comprender cómo influyen los parámetros de diseño —es decir, los polos en lazo cerrado y el período de muestreo— en las propiedades del sistema en lazo cerrado. Esto puede verse con un ejemplo.

Ejemplo 9.3

Consideremos el sistema de integrador doble. En lugar de emplear los parámetros p_1 y p_2 en (9.7), introduciremos otros dos parámetros, que admiten una interpretación física más inmediata. Si el sistema discreto deseado se obtiene muestreando un sistema de segundo orden, tenemos que

$$p_1 = -2e^{-\zeta\omega h} \cos(\omega h \sqrt{1 - \zeta^2})$$

$$p_2 = e^{-2\zeta\omega h}$$

donde ω es la frecuencia natural y ζ es el amortiguamiento relativo (compárese con el Ejemplo 3.13). El parámetro ζ influye sobre el amortiguamiento relativo de la respuesta y ω influye sobre la velocidad de la respuesta.

Para estudiar la magnitud de la señal de control se supone que el sistema tiene una posición inicial x_0 y una velocidad inicial v_0 . El valor inicial de la señal de control es entonces

$$u(0) = -l_1 x_0 - l_2 v_0$$

Si el período de muestreo es corto, las expresiones para p_1 y p_2 pueden aproximarse mediante un desarrollo en serie. Para (9.3) y (9.8) se obtiene la siguiente aproximación:

$$u(0) \approx -\omega^2 x_0 + 2\zeta\omega v_0$$

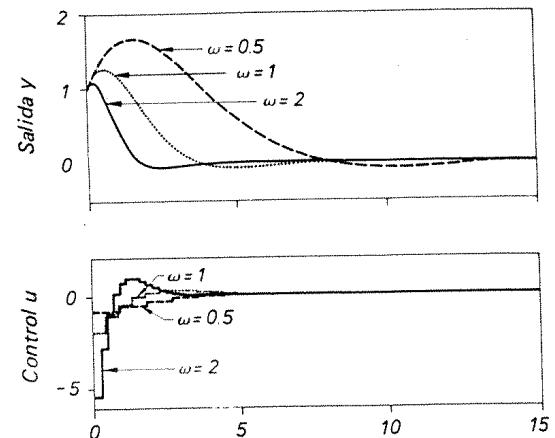
Esta expresión muestra claramente que el módulo de la señal de control aumenta al aumentar ω . Entonces un aumento en la velocidad de la respuesta del sistema requerirá un aumento en las señales de control. Si se conocen los límites de la señal de control y de las perturbaciones típicas, es posible determinar los valores razonables de ω . Las consecuencias de diferentes elecciones de ω cuando $x_0 = 1$ y $v_0 = 0$ se muestran en la Figura 9.1.

Si los parámetros ζ y ω están fijados, todavía hay que seleccionar el período de muestreo. La elección del período de muestreo influirá sobre la rapidez con que se detectarán las perturbaciones. También influirá sobre las curvas de respuesta. La Sección 3.7 expone cómo seleccionar el período de muestreo para un sistema en lazo abierto. Los mismos argumentos pueden emplearse para el sistema en lazo cerrado. El período de muestreo se puede escoger de tal forma que

$$N_r \approx 2 - 4$$

donde N_r es el número de muestras por tiempo de subida del sistema en lazo cerrado. Esto significa que se debe escoger el período de muestreo en relación con el comportamiento deseado del sistema en lazo cerrado. El período de muestreo también se puede relacionar

Figura 9.1 Respuesta del sistema en lazo cerrado del Ejemplo 9.3 obtenida para diferentes valores de ω con $\omega h = 0,44$ y $\zeta = 0,707$.



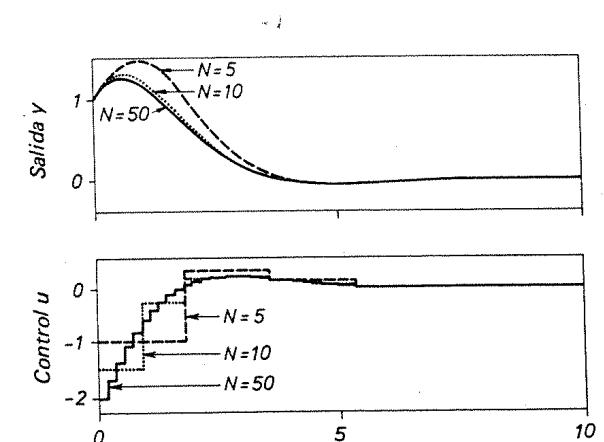
con la frecuencia amortiguada del sistema en lazo cerrado. Conviene introducir el parámetro N definido por

$$N = \frac{2\pi}{\omega h \sqrt{1 - \zeta^2}} \quad (9.16)$$

Este parámetro es la relación del período amortiguado y el período de muestreo. La Figura 9.2 muestra las consecuencias de diferentes elecciones de N en el caso especial. Está claro que mientras $N > 10$, hay muy pocas diferencias en las respuestas. Las respuestas obtenidas para $N > 50$ no se pueden diferenciar en la gráfica.

Una regla empírica razonable es escoger N alrededor de 20. Esto da $\omega h \approx 0,44$ para $\zeta = 0,707$; $N_r = 2 - 4$ corresponde a $N \approx 10 - 20$ para el amortiguamiento $\zeta = 0,707$.

Figura 9.2 Respuesta del sistema del Ejemplo 9.3 para diferentes períodos de muestreo cuando $\omega = 1$ y $\zeta = 0,707$.



Control de tiempo finito

Si en el problema de asignación de polos se escogen los polos deseados para que estén en el origen, la ecuación característica para el sistema en lazo cerrado resulta

$$z^n = 0$$

Entonces se deduce del teorema de Cayley-Hamilton que la matriz del sistema $\Phi_c = \Phi - \Gamma L$ del sistema en lazo cerrado satisface

$$\Phi_c^n = 0$$

Esta estrategia tiene la propiedad de que lleva todos los estados a cero en n pasos como máximo después de una perturbación impulsiva en el estado del proceso. Esta estrategia de control recibe el nombre de *control de tiempo finito* (en inglés: *deadbeat*).

En el control de tiempo finito sólo hay un parámetro de diseño: el período de muestreo. Puesto que el error va a cero en n pasos como máximo, el tiempo de establecimiento es, como máximo, nh . Sin embargo, el período de muestreo influirá drásticamente sobre la magnitud de la señal de control. Cuando se emplea el control de tiempo finito, la magnitud de la señal de control normalmente aumenta de forma drástica cuando se disminuye el período de muestreo. Este hecho ha dado al control de tiempo finito una inmerecidamente mala reputación. Por eso es importante escoger cuidadosamente el período de muestreo cuando se usa el control de tiempo finito.

La estrategia de tiempo finito es única de los sistemas muestreados. No hay una característica equivalente para los sistemas continuos. Un ejemplo muestra las propiedades del control de tiempo finito.

Ejemplo 9.4 – Control de tiempo finito de un integrador doble

Consideremos una planta con un integrador doble. De (9.8) se deduce que la estrategia de control de tiempo finito viene dada por (9.5) con

$$\begin{cases} l_1 = \frac{1}{h^2} \\ l_2 = \frac{3}{2h} \end{cases}$$

Si el proceso tiene el estado inicial $x(0) = \text{col}[x_0, v_0]$, resulta que

$$u(0) = -\frac{x_0}{h^2} - \frac{3v_0}{2h}$$

$$u(h) = \frac{x_0}{h^2} + \frac{v_0}{2h}$$

En la Tabla 9.1 se listan las señales de control obtenidas para $x_0 = 1$ y $v_0 = 1$. Las señales de salida y de control pueden verse en la Figura 9.3.

TABLA 9.1 Señales de control para el control de tiempo finito de un integrador doble con $x(0) = \text{col}[1, 1]$ y diferentes períodos de muestreo.

h	100	10	1	0.1	0.01
$u(0)$	-0.0151	-0.16	-2.5	-115	-10,150
$u(h)$	0.0051	0.06	1.5	105	10,050

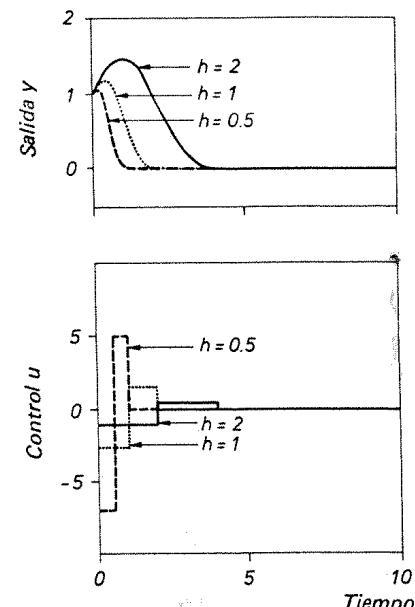


Figura 9.3 Simulación del control de tiempo finito de un integrador doble. La condición inicial es $x(0) = \text{col}[1, 1]$.

Perturbaciones más generales

Hay muchas formas de generalizar el problema del diseño con asignación de polos. En la práctica es importante ser capaz de manejar perturbaciones más generales que los impulsos. Una forma de hacerlo es emplear la idea descrita en el Capítulo 6. Allí se considera que una perturbación se genera enviando un impulso a través de un sistema lineal.

Por ejemplo, supongamos que el sistema está descrito por

$$\frac{dx}{dt} = Ax + Bu + v$$

donde v es una perturbación descrita por

$$\frac{d\xi}{dt} = A_v \xi$$

$$v = C_v \xi$$

con unas condiciones dadas. También se supone que ξ puede medirse. Introduciendo el vector de estado aumentado

$$z = \begin{bmatrix} x \\ \xi \end{bmatrix}$$

el sistema se puede describir por

$$\frac{d}{dt} \begin{bmatrix} x \\ \xi \end{bmatrix} = \begin{bmatrix} A & C_v \\ 0 & A_v \end{bmatrix} \begin{bmatrix} x \\ \xi \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u \quad (9.17)$$

Entonces tenemos un problema de la misma forma que el problema básico de asignación de polos. Nótese, sin embargo, que existe una diferencia importante: el sistema de (9.17) no es completamente alcanzable. La realimentación no puede influir sobre los polos asociados con la descripción de la perturbación; es decir, los valores propios de A_v . Esto es completamente natural.

Cuando se muestrea el sistema se obtiene el siguiente sistema discreto:

$$\begin{bmatrix} x(kh + h) \\ \xi(kh + h) \end{bmatrix} = \begin{bmatrix} \Phi & \Phi_{xv} \\ 0 & \Phi_v \end{bmatrix} \begin{bmatrix} x(kh) \\ \xi(kh) \end{bmatrix} + \begin{bmatrix} \Gamma \\ 0 \end{bmatrix} u(kh)$$

La realimentación lineal y general del estado viene dada por

$$u(kh) = -Lx(kh) - L_v \xi(kh) \quad (9.18)$$

El sistema en lazo cerrado está descrito por

$$x(kh + h) = [\Phi - \Gamma L]x(kh) + [\Phi_{xv} - \Gamma L_v]\xi(kh)$$

$$\xi(kh + h) = \Phi_v \xi(kh)$$

Si el par (Φ, Γ) es alcanzable, la matriz L puede escogerse de tal forma que la matriz $\Phi - \Gamma L$ tenga los valores propios prefijados. La matriz Φ_v no puede ser influenciada por la realimentación. Para reducir el efecto de las perturbaciones es útil escoger L_v de forma que la matriz $\Phi_{xv} - \Gamma L_v$ sea pequeña en el sentido de que el acople con la perturbación sea débil.

Nótese que la ley de realimentación dada en (9.18) puede interpretarse como una combinación de un término de realimentación Lx y un término de prealimentación $L_v \xi$. El término de prealimentación surge del modelo de la perturbación. Se ve claramente que la solución requiere que ambos, x y ξ , sean directamente medibles.

Aspectos de cálculo

Para sistemas sencillos de bajo orden, a menudo lo más fácil es calcular la realimentación del estado, como se hizo en el Ejemplo 9.1. El procedimiento general es introducir una realimentación general del estado con coeficientes indeterminados, determinar el polinomio característico, e igualar esto al polinomio característico deseado. Entonces se obtiene un conjunto de ecuaciones lineales en los coeficientes de la realimentación. Estas ecuaciones siempre se pueden resolver si el sistema es alcanzable.

Por supuesto, también se puede emplear la fórmula general dada en (9.15) para calcular la realimentación del estado. Es fácil escribir un programa de computador para este problema. Sin embargo, nótese que la fórmula no es muy apropiada para cálculos numéricos muy precisos. Como norma, debe evitarse emplear métodos que utilicen el cálculo de potencias de matrices. Hay otras formas de calcular la matriz de realimentación L que son más apropiadas para el cálculo numérico.

9.3. Observadores

Normalmente no es realista suponer que pueden medirse todos los estados de un sistema y todas las perturbaciones. Si se dispone de un modelo matemático del sistema, se puede intentar calcular el estado a partir de las entradas y salidas medidas. En esta sección se exponen algunas formas de hacerlo.

Se supone que el sistema está descrito por el modelo muestreado

$$\begin{cases} x(k+1) = \Phi x(k) + \Gamma u(k) \\ y(k) = C x(k) \end{cases} \quad (9.19)$$

A continuación se considera el problema de calcular el estado $x(k)$ a partir de las secuencias de entrada y salida $y(k), y(k-1), \dots, u(k), u(k-1), \dots$

Para determinar el estado del sistema a partir de las entradas y salidas, el sistema debe ser completamente observable (compárese con la Sección 5.3). Hay muchas formas de determinar el estado y predecir sus valores. En esta sección se exponen las alternativas al cálculo directo, empleando modelos dinámicos y observadores de Luenberger. El método más adecuado para un sistema en particular depende de la naturaleza de las perturbaciones y del ruido de medida. Si se conocen los modelos de estos dos factores, es posible diseñar observadores y predictores óptimos. Esto se expone en el Capítulo 11.

Cálculo directo de las variables de estado

Lo primero que se va a hacer es calcular el vector de estado directamente de las entradas y salidas. Para el caso general con una salida escalar, de (9.19) se tiene que



$$\begin{aligned}
 y(k-n+1) &= Cx(k-n+1) \\
 y(k-n+2) &= C\Phi x(k-n+1) + C\Gamma u(k-n+1) \\
 &\vdots \\
 &\vdots \\
 y(k) &= C\Phi^{n-1}x(k-n+1) + C\Phi^{n-2}\Gamma u(k-n+1) + \cdots + C\Gamma u(k-1)
 \end{aligned}$$

Estas ecuaciones pueden escribirse como

$$\begin{bmatrix} y(k-n+1) \\ y(k-n+2) \\ \vdots \\ y(k) \end{bmatrix} = W_o x(k-n+1) + \Omega \begin{bmatrix} u(k-n+1) \\ u(k-n+2) \\ \vdots \\ u(k-1) \end{bmatrix}$$

donde W_o es la matriz de observabilidad

$$W_o = \begin{bmatrix} C \\ C\Phi \\ \vdots \\ C\Phi^{n-1} \end{bmatrix} \quad (9.20)$$

$$\Omega = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ C\Gamma & 0 & \cdots & 0 \\ \vdots & & & \\ C\Phi^{n-2}\Gamma & C\Phi^{n-3}\Gamma & \cdots & C\Gamma \end{bmatrix}$$

y

Si el sistema (9.19) es observable, el vector de estado $x(k-n+1)$ puede determinarse a partir de

$$x(k-n+1) = W_o^{-1} \begin{bmatrix} y(k-n+1) \\ y(k-n+2) \\ \vdots \\ y(k) \end{bmatrix} - W_o^{-1}\Omega \begin{bmatrix} u(k-n+1) \\ u(k-n+2) \\ \vdots \\ u(k-1) \end{bmatrix}$$

El empleo repetido de (9.19) da

$$x(k) = \Phi^{n-1}W_o^{-1} \begin{bmatrix} y(k-n+1) \\ y(k-n+2) \\ \vdots \\ y(k) \end{bmatrix} + \Psi \begin{bmatrix} u(k-n+1) \\ u(k-n+2) \\ \vdots \\ u(k-1) \end{bmatrix} \quad (9.21)$$

donde

$$\Psi = [\Phi^{n-2}\Gamma \quad \Phi^{n-3}\Gamma \quad \cdots \quad \Gamma] - \Phi^{n-1}W_o^{-1}\Omega$$

Entonces el vector de estado viene dado como una combinación lineal de $y(k)$, $y(k-1)$, ..., $y(k-n+1)$ y de $u(k-1)$, $u(k-2)$, ..., $u(k-n+1)$. La ecuación (9.21) puede escribirse en la forma

$$x(k) = F(q^{-1})y(k) + G(q^{-1})u(k-1) \quad (9.22)$$

donde F y G son vectores polinomiales de grados $n-1$ y $n-2$.

Las ecuaciones (9.21) y (9.22) muestran que después de una perturbación, la estimación correcta se obtendrá después de al menos n pasos. El observador de (9.21) o (9.22) puede llamarse observador de tiempo finito. El resultado se sintetiza en el Teorema 9.2.

Teorema 9.2. Considérese el sistema de (9.19). Supóngase que es completamente observable. Entonces el vector de estado puede calcularse mediante (9.21). \square

La fórmula general se ilustra con un ejemplo.

Ejemplo 9.5 – El integrador doble

Para el integrador doble tenemos

$$\Phi = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} h^2/2 \\ h \end{bmatrix}, \quad C = [1 \quad 0]$$

La matriz de observabilidad viene dada por

$$W_o = \begin{bmatrix} 1 & 0 \\ 1 & h \end{bmatrix} \quad y \quad W_o^{-1} = \frac{1}{h} \begin{bmatrix} h & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1/h & 1/h \end{bmatrix}$$

Esto conduce a

$$\Phi^{n-1}W_o^{-1} = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1/h & 1/h \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1/h & 1/h \end{bmatrix}$$

Además

$$W_o^{-1} \begin{bmatrix} 0 \\ C\Gamma \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1/h & 1/h \end{bmatrix} \begin{bmatrix} 0 \\ h^2/2 \end{bmatrix} = \begin{bmatrix} 0 \\ h/2 \end{bmatrix}$$

y

$$\Psi = \Gamma - \Phi W_0^{-1} \begin{bmatrix} 0 \\ C\Gamma \end{bmatrix} = \begin{bmatrix} h^2/2 \\ h \end{bmatrix} - \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ h/2 \end{bmatrix} = \begin{bmatrix} 0 \\ h/2 \end{bmatrix}$$

La ecuación (9.21) nos da entonces

$$\begin{aligned} x(kh) &= \begin{bmatrix} 0 & 1 \\ -1/h & 1/h \end{bmatrix} \begin{bmatrix} y(kh-h) \\ y(kh) \end{bmatrix} + \begin{bmatrix} 0 \\ h/2 \end{bmatrix} u(kh-h) \\ &= \begin{bmatrix} 1 \\ 1/h \end{bmatrix} y(kh) - \begin{bmatrix} 0 \\ 1/h \end{bmatrix} y(kh-h) + \begin{bmatrix} 0 \\ h/2 \end{bmatrix} u(kh-h) \end{aligned}$$

La primera componente se mide directamente mientras que la segunda se obtiene mediante la expresión

$$x_2(kh) = \frac{y(kh) - y(kh-h)}{h} + \frac{h}{2} \cdot u(kh-h)$$

Como se esperaba, $x_2 \rightarrow y$ cuando $h \rightarrow 0$. Sin embargo, en este caso el cálculo directo es mucho más sencillo que empleando la fórmula general.

Reconstrucción mediante un sistema dinámico

El cálculo directo del vector de estado tiene la ventaja de que la variable de estado se obtiene como máximo en n pasos. La desventaja del método está en que puede ser sensible a las perturbaciones. Por tanto, es útil disponer de otras alternativas.

Considérese el sistema de (9.19). Supóngase que se quiere aproximar el estado x por el estado \hat{x} del modelo

$$\hat{x}(k+1) = \Phi \hat{x}(k) + \Gamma u(k) \quad (9.23)$$

que tiene la misma entrada que el sistema de (9.19). Si el modelo de (9.23) es perfecto en el sentido que los parámetros son idénticos a los del sistema en (9.19) y si las condiciones iniciales de (9.19) y (9.23) son las mismas, entonces el estado \hat{x} del modelo de (9.23) será idéntico al estado x del sistema real en (9.19). Si las condiciones iniciales de (9.19) y (9.23) son diferentes, entonces \hat{x} convergerá hacia x sólo si el sistema (9.19) es asintóticamente estable.

Nótese que la reconstrucción en (9.23) no utiliza la salida medida. La reconstrucción en (9.23) se puede mejorar introduciendo la diferencia entre las salidas medida y estimada, $y - C\hat{x}$, como una realimentación, para obtener

$$\hat{x}(k+1|k) = \Phi \hat{x}(k|k-1) + \Gamma u(k) + K[y(k) - C\hat{x}(k|k-1)] \quad (9.24)$$

Aquí K es una matriz que se ha escogido de la forma adecuada. La notación $\hat{x}(k+1|k)$ se emplea para indicar que se trata de una estimación, o predicción, de $x(k+1)$ basada en las medidas disponibles en el instante k . Introduzcamos el error de reconstrucción

$$\tilde{x} = x - \hat{x} \quad (9.25)$$

Restando (9.24) de (9.19) tenemos

$$\begin{aligned} \tilde{x}(k+1|k) &= \Phi \tilde{x}(k|k-1) - K[y(k) - C\hat{x}(k|k-1)] \\ &= [\Phi - KC]\tilde{x}(k|k-1) \end{aligned} \quad (9.26)$$

Por tanto, si se escoge K de forma que el sistema en (9.26) sea asintóticamente estable, el error de reconstrucción siempre convergerá a cero. Además, al introducir una realimentación de las medidas en la reconstrucción es posible hacer que el error vaya a cero incluso si el sistema (9.19) es inestable. El sistema de la expresión (9.24) recibe el nombre de *observador* del sistema (9.19) porque genera el estado del sistema a partir de las medidas de las entradas y salidas.

Hay muchas variaciones del observador de (9.24). El observador tiene un retardo, porque $\hat{x}(k|k-1)$ depende solamente de las medidas hasta el instante $k-1$. Para obviar esto se puede emplear el siguiente observador:

$$\begin{aligned} \hat{x}(k|k) &= \Phi \hat{x}(k-1|k-1) + \Gamma u(k-1) \\ &\quad + K[y(k) - C\{\Phi \hat{x}(k-1|k-1) + \Gamma u(k-1)\}] \\ &= [I - KC][\Phi \hat{x}(k-1|k-1) + \Gamma u(k-1)] + Ky(k) \end{aligned} \quad (9.27)$$

Cuando se usa (9.27), el error de reconstrucción viene dado por

$$\tilde{x}(k|k) = x(k) - \hat{x}(k|k) = (\Phi - KC\Phi)\tilde{x}(k-1|k-1)$$

Esta ecuación es similar a (9.26) y de la definición de matriz de observabilidad W_0 se obtiene que el par $(\Phi, C\Phi)$ es observable si el par (Φ, C) es detectable. Esto implica que a $\Phi - KC\Phi$ se le pueden dar valores propios arbitrarios seleccionando K adecuadamente. Además,

$$\begin{aligned} y(k) - C\hat{x}(k|k) &= C\tilde{x}(k|k) \\ &= (C\Phi - CKC\Phi)\tilde{x}(k-1|k-1) \\ &= (I - CK)C\Phi\tilde{x}(k-1|k-1) \end{aligned}$$

Si el sistema tiene p salidas, entonces $I - CK$ es una matriz $p \times p$; K puede escogerse de tal forma que $CK = I$ si $\text{rango}(C) = p$. Esto implica que $C\hat{x}(k|k) = y(k)$; es decir, las salidas del sistema se estiman sin error. Esto hará posible eliminar p ecuaciones de (9.27), y se reducirá el orden del observador. Los observadores de orden reducido de este tipo se les suele llamar *observadores de Luenberger*.

Cómo decidir cuándo se puede encontrar una matriz K adecuada

Queda por encontrar la forma apropiada de escoger la matriz K para que el sistema (9.26) sea estable. Dadas las matrices Φ y C , encontraremos una matriz K tal que la matriz $\Phi - KC$ tenga unos valores propios prefijados. Ya que una matriz y su transpuesta tienen los mismos valores propios, el problema es el mismo que encontrar una matriz K^T tal que $\Phi^T - C^T K^T$ tenga los valores propios pedidos. Sin embargo,



este problema se ha resuelto en la Sección 9.2 en conexión con la exposición del problema de la asignación de polos. Si trasladamos aquellos resultados, el problema se podrá resolver si la matriz

$$W_o^T = [C^T \quad \Phi^T C^T \quad \dots \quad (\Phi^{n-1})^T C^T]$$

es de rango completo. Nótese que la matriz es la transpuesta de la matriz de observabilidad del sistema (9.19). Esto se sintetiza en el Teorema 9.3.

Teorema 9.3. Supóngase que el sistema (9.19) es completamente observable. Entonces es posible encontrar una matriz K tal que el observador dado en (9.24) tiene el polinomio característico $P(z)$ definido por (9.10). \square

Determinación de K

Se ve que la determinación de la matriz K en el observador (9.24) es un problema matemático idéntico al de determinar la matriz de realimentación L en el problema de asignación de polos. Los aspectos prácticos están también íntimamente relacionados. La determinación de los polos del observador es un compromiso entre la sensibilidad a los errores de medida y la rápida recuperación de los errores iniciales. Un observador rápido convergerá rápidamente, pero también será muy sensible a los errores de medida.

Un procedimiento normal para determinar la matriz K es decidir los polos adecuados para el observador y entonces buscar la matriz K que da los polos deseados. Para problemas sencillos lo más fácil es fijar la matriz K y determinar las ecuaciones para que $\Phi - KC$ tenga los valores propios prefijados.

Ya se ha visto anteriormente que la selección de K está relacionada con la selección de L para la asignación de polos. La elección de K puede determinarse a partir del Teorema 9.1 usando las conversiones

$$L \longrightarrow K^T \quad W_o \longrightarrow W_o^T \quad \Phi \longrightarrow \Phi^T$$

De (9.15) se deduce que K viene dada por

$$K^T = [0 \quad \dots \quad 0 \quad 1](W_o^T)^{-1}P(\Phi^T)$$

o

$$K = P(\Phi)W_o^{-1}[0 \quad \dots \quad 0 \quad 1]^T \quad (9.28)$$

Entonces el polinomio característico de $\Phi - KC$ es $P(z)$, como se definió en (9.10). La dualidad con la asignación de polos implica también que K es especialmente fácil de determinar si el sistema está en la forma observable.

Ejemplo 9.6 – Observador de orden completo para el integrador doble

Consideremos un sistema con un integrador doble. La matriz $\Phi - KC$ viene dada por

$$\Phi_o = \Phi - KC = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -k_2 & 1 \end{bmatrix} = \begin{bmatrix} 1 - k_1 & h \\ -k_2 & 1 \end{bmatrix}$$

Entonces la ecuación característica es

$$z^2 - (2 - k_1)z + 1 - k_1 + k_2h = 0$$

Si el observador (9.24) debe tener la ecuación característica

$$z^2 + p_1z + p_2 = 0$$

se obtienen las siguientes ecuaciones:

$$2 - k_1 = -p_1$$

$$1 - k_1 + k_2h = p_2$$

Estas ecuaciones lineales dan

$$k_1 = 2 + p_1$$

$$k_2 = (1 + p_1 + p_2)/h$$

Ejemplo 9.7 – Observador de orden reducido para el integrador doble

El observador de (9.27) aplicado al integrador doble da las ecuaciones

$$\hat{x}(k|k) = \begin{bmatrix} 1 - k_1 & h(1 - k_1) \\ -k_2 & 1 - hk_2 \end{bmatrix} \hat{x}(k-1|k-1) + \begin{bmatrix} (1 - k_1)h^2/2 \\ h(1 - hk_2/2) \end{bmatrix} u(k-1) + \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} y(k)$$

Si $I - CK = 0$ —es decir, si $k_1 = 1$ — entonces la primera ecuación se reduce a

$$\hat{x}_1(k|k) = y(k)$$

El observador de orden reducido viene dado por la segunda ecuación, que se puede simplificar en

$$\hat{x}_2(k|k) = (1 - hk_2)\hat{x}_2(k-1|k-1) + k_2[y(k) - y(k-1)] + h(1 - hk_2/2)u(k-1)$$

Escogiendo k_2 se puede dar al observador de orden reducido un valor propio arbitrario. Por ejemplo, si $k_2 = 1/h$, la respuesta de tiempo finito, entonces se obtiene el mismo resultado que cuando se hizo el cálculo directo en el Ejemplo 9.5.

9.4. Realimentación de la salida

Se puede obtener una solución al problema de la asignación de polos con realimentación de la salida combinando los resultados de las Secciones 9.2 y 9.3, lo que lleva al sistema descrito por

$$\begin{aligned} x(k+1) &= \Phi x(k) + \Gamma u(k) \\ y(k) &= Cx(k) \end{aligned} \quad (9.29)$$

Se desea una ley de realimentación lineal que relacione u e y de forma que el sistema en lazo cerrado tenga unos polos dados. Se supone que las perturbaciones son impulsos o, lo que es equivalente, que los estados iniciales son desconocidos.

La ley de control admisible es tal que $u(k)$ es función de $y(k-1), y(k-2), \dots, u(k-1), u(k-2), \dots$. Si todas las variables de estado son medibles, se ha visto en la Sección 9.2 que la realimentación

$$u(k) = -Lx(k)$$

dado los polos deseados. Cuando el estado no puede medirse, parece razonable de forma intuitiva emplear la ley de control

$$u(k) = -L\hat{x}(k|k-1) \quad (9.30)$$

donde \hat{x} se obtiene mediante el observador

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k-1) + \Gamma u(k) + K[y(k) - C\hat{x}(k|k-1)] \quad (9.31)$$

Entonces la realimentación es un sistema dinámico. Nótese que la dinámica se debe a la dinámica del observador. En la Figura 9.4 se presenta un diagrama de bloques de la realimentación.

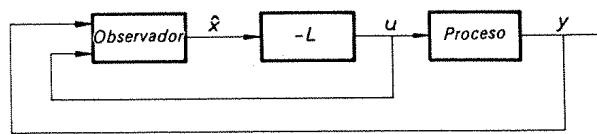


Figura 9.4 Diagrama de bloques del regulador que se obtiene combinando la realimentación del estado con un observador.

La dinámica descrita por (9.30) y (9.31) puede representarse por una función de transferencia discreta de orden n entre y y u :

$$G_r(z) = -L[zI - \Phi + \Gamma L + KC]^{-1}K$$

Análisis del sistema en lazo cerrado

El sistema en lazo cerrado tiene las propiedades deseadas. Para verlo, introduzcamos

$$\tilde{x} = x - \hat{x}$$

Empleando (9.29), (9.30) y (9.31) se puede describir el sistema en lazo cerrado por las ecuaciones

$$\begin{cases} x(k+1) = [\Phi - \Gamma L]x(k) + \Gamma L\hat{x}(k|k-1) \\ \tilde{x}(k+1|k) = [\Phi - KC]\tilde{x}(k|k-1) \end{cases} \quad (9.32)$$

El sistema en lazo cerrado es de orden $2n$. Los valores propios del sistema en lazo cerrado son los de las matrices $\Phi - \Gamma L$ y $\Phi - KC$. Nótese que los valores propios de

$\Phi - \Gamma L$ son los polos en lazo cerrado deseados obtenidos en la Sección 9.2 y que los valores propios de $\Phi - KC$ son los polos del observador dados en la Sección 9.3.

Ejemplo 9.8 – Realimentación de salida de un integrador doble

Consideremos la planta con integrador doble. Supongamos que el vector de realimentación L está determinado como en los Ejemplos 9.2 y 9.3 con la frecuencia natural en lazo cerrado $\omega = 1$, el amortiguamiento $\zeta = 0.7$ y $h = 0.44$. Esto da $L = [0.73, 1.21]$. Supongamos primero que se diseña un observador como en el Ejemplo 9.6 con los polos del observador en $z = 0.75$. La Figura 9.5a) muestra los estados reales y estimados cuando en la ley de control se emplean los estados estimados. Por su parte la Figura 9.5b) presenta los estados y la estimación del segundo de ellos cuando se emplea el observador de orden reducido del Ejemplo 9.7. El polo del observador está en $z = 0.75$.

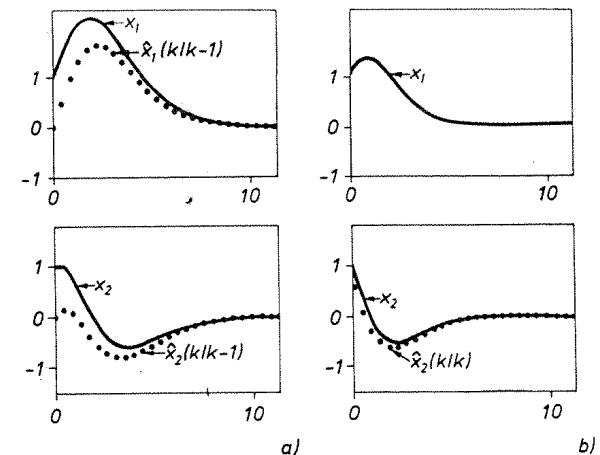


Figura 9.5 Control del integrador doble empleando estimaciones de los estados. Se presentan los estados y las estimaciones de los estados para:
a) Un observador de segundo orden.
b) Un observador de orden reducido.

Extensiones

El problema se puede ampliar a otras perturbaciones introduciendo modelos de las perturbaciones del proceso y del ruido de medida. Entonces el modelo se puede representar como sigue (compárese con la Sección 9.2):

$$\begin{aligned} x(k+1) &= \Phi x(k) + \Phi_{xu}\xi(k) + \Gamma u(k) \\ \xi(k+1) &= \Phi_v\xi(k) \\ \eta(k+1) &= \Phi_w\eta(k) \\ y(k) &= Cx(k) + C_w\eta(k) \end{aligned} \quad (9.33)$$

La ley de control viene dada por

$$u(k) = -L\hat{x}(k) - L_v\hat{\xi}(k) - L_w\hat{\eta}(k) \quad (9.34)$$

donde

$$\begin{bmatrix} \hat{x}(k+1) \\ \hat{\xi}(k+1) \\ \hat{\eta}(k+1) \end{bmatrix} = \begin{bmatrix} \Phi & \Phi_{xv} & 0 \\ 0 & \Phi_v & 0 \\ 0 & 0 & \Phi_w \end{bmatrix} \begin{bmatrix} \hat{x}(k) \\ \hat{\xi}(k) \\ \hat{\eta}(k) \end{bmatrix} + \begin{bmatrix} \Gamma \\ 0 \\ 0 \end{bmatrix} u(k) + \begin{bmatrix} K \\ K_v \\ K_w \end{bmatrix} \epsilon(k) \quad (9.35)$$

$$\epsilon(k) = y(k) - C\hat{x}(k) - C_w\hat{\eta}(k)$$

Nótese que (9.34) es un ejemplo típico de una combinación de realimentación y prealimentación.

9.5. El problema del servo

Hasta ahora sólo se ha expuesto el problema del regulador. El criterio ha sido eliminar perturbaciones tipo impulso y llevar los estados del sistema a cero. Otro problema importante es el problema del servo. En éste el objetivo es hacer que los estados y las salidas del sistema sigan unas trayectorias deseadas. Esta sección estudia cómo hacer que el sistema siga las señales de mando y al mismo tiempo sea capaz de resolver el problema del regulador. El diseño se hace resolviendo primero el problema del regulador empleando los métodos de asignación de polos expuestos anteriormente y resolviendo entonces el problema del servo en esencia introduciendo ceros en el sistema en lazo cerrado.

Una interpretación de los polos y los ceros

Antes de exponer cómo pueden introducirse las señales de mando vamos a hacer una digresión sobre la interpretación de los polos y ceros. Los polos, o los valores propios, de un sistema lineal reflejan los acoplos internos del sistema. Nos dicen cómo se comporta el sistema cuando se le deja evolucionar por sí mismo. Los ceros nos dicen cómo está acoplado el sistema a su entorno.

Básicamente, la introducción de las señales de mando en un sistema significa acoplar el sistema a su entorno. Es evidente que esto supone el manipular los ceros del sistema. Para ver cómo influye la forma en que se introducen las señales de mando sobre los ceros de un sistema, considérense los diagramas de flujo de señal de un sistema de una sola entrada y una sola salida de la Figura 9.6. Los coeficientes del polinomio característico, que determinan los polos del sistema, se designan por a_i . Los coeficientes del polinomio que determina los ceros del sistema se designan por b_i .

En la Figura 9.6b) se ve claramente que los polos del sistema pueden cambiarse por realimentación del estado. De la misma forma se ve en la Figura 9.6a) que se puede influir sobre todos los ceros del sistema cambiando la prealimentación a todas las variables de estado.

En general es muy difícil modificar las conexiones de prealimentación a las variables de estado de un proceso. Sólo se puede influir sobre el proceso a través de la

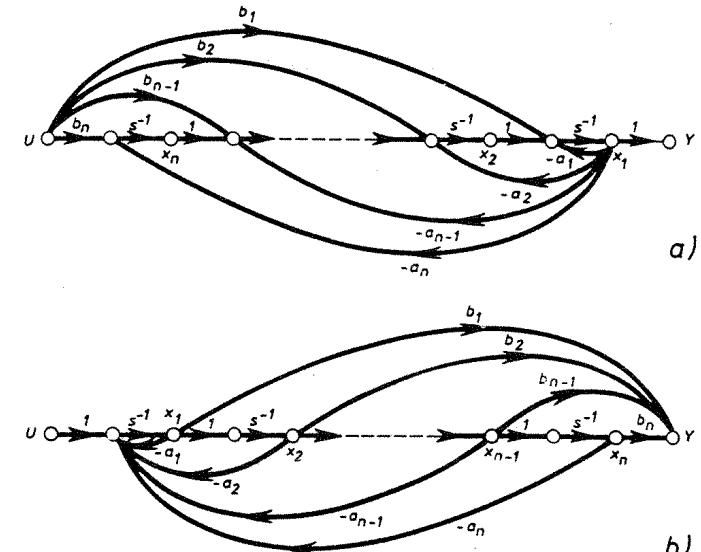


Figura 9.6 Diagramas de flujo de señal de un sistema con función de transferencia

$$G(s) = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \cdots + b_n}{s^n + a_1 s^{n-1} + \cdots + a_n}$$

variable de control. Se necesitan nuevos actuadores para influir sobre la totalidad del estado del proceso directamente. Sin embargo, se puede influir libremente sobre las variables de estado asociadas con el compensador para modificar algunos de los ceros.

Una estructura de regulador

Para el problema del servo es natural dar las especificaciones en términos de un modelo que dice cómo debería responder idealmente el modelo a las señales de mando. Sea el sistema muestrado descrito por el sistema de orden n

$$\begin{aligned} x_m(k+1) &= \Phi_m x_m(k) + \Gamma_m u_c(k) \\ y_m(k) &= C_m x_m(k) \end{aligned} \quad (9.36)$$

Es natural organizar el sistema como se ve en la Figura 9.7. En primer lugar se determina el modelo del proceso inverso. La salida del modelo del proceso inverso se designa por u_m . A la entrada del proceso se le suma la señal que viene del modelo del proceso inverso. Para estar seguro que la realimentación L no se opone a la respuesta deseada del servo, es necesario restar x_m de \hat{x} antes de aplicar la realimentación. Si el

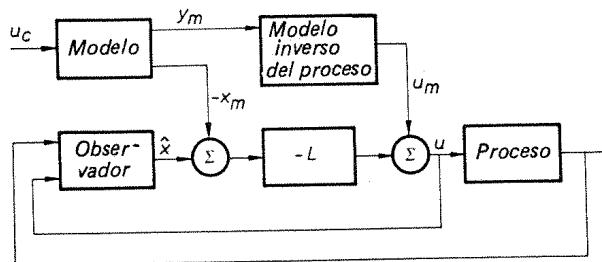


Figura 9.7 Introducción de la señal de mando en un sistema con realimentación de la salida.

inverso es correcto y no hay perturbaciones, la salida del proceso responderá en la forma deseada. La realimentación L contribuirá a reducir cualquier tipo de desviación.

También, aunque el modelo del proceso inverso no sea causal, la combinación del modelo y del modelo del proceso inverso puede que sí lo sea. Si el proceso no tiene inverso estable, entonces se puede utilizar en su lugar una aproximación estable.

El modelo puede utilizarse para introducir polos y ceros en la función de transferencia discreta entre u_e e y . Esto permite disponer de diferentes funciones de transferencia discretas de la salida respecto del valor de referencia y de las perturbaciones.

Análisis

El sistema en lazo cerrado está descrito por las ecuaciones

$$\begin{cases} x(k+1) = \Phi x(k) + \Gamma u(k) \\ u(k) = L[x_m(k) - \hat{x}(k)] + u_m(k) \\ \hat{x}(k+1|k) = \Phi \hat{x}(k|k-1) + \Gamma u(k) + K[y(k) - C\hat{x}(k|k-1)] \end{cases} \quad (9.37)$$

Introduciendo

$$\begin{aligned} \tilde{x} &= x - \hat{x} \\ \tilde{u}_e(k) &= u_m(k) + Lx_m(k) \end{aligned} \quad (9.38)$$

las ecuaciones se pueden escribir como

$$\begin{cases} x(k+1) = [\Phi - \Gamma L]x(k) + \Gamma L\tilde{x}(k|k-1) + \Gamma \tilde{u}_e(k) \\ \tilde{x}(k+1|k) = [\Phi - KC]\tilde{x}(k|k-1) \\ u(k) = -L\tilde{x}(k) + \tilde{u}_e(k) \end{cases} \quad (9.39)$$

De (9.39) se deduce que las señales de mando no influirán sobre los errores del observador. Esto es normal, ya que no tendría sentido introducir las señales de mando

de tal forma que ellas mismas generaran errores de estimación. Como resultado, los polos del observador no aparecerán en la función de transferencia, que relaciona la salida con la señal de mando. Esto puede verse en el diagrama de bloques de la Figura 9.8.

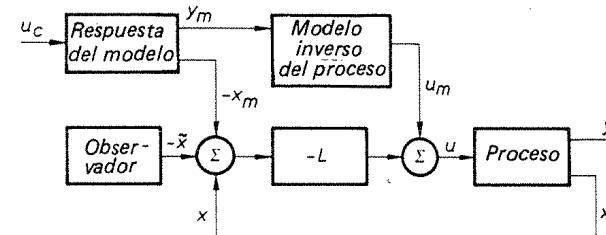


Figura 9.8 Diagrama de bloques de un regulador y un servo combinados.

Nótese que el regulador puede describirse por la relación entrada-salida

$$u = -L[qI - \Phi + \Gamma L + KC]^{-1}[\Gamma \tilde{u}_e + Ky] + \tilde{u}_e \quad (9.40)$$

En el caso de medidas escalares, esta ecuación se puede escribir en la forma

$$R(q)u(k) = T(q)\tilde{u}_e(k) - S(q)y(k) \quad (9.41)$$

donde R , S y T son polinomios.

Introducción de integradores

Los métodos expuestos en este capítulo utilizan toda la información disponible sobre el sistema. Esto significa que todas las perturbaciones están modeladas y que los modelos son exactos. Sin embargo, a menudo interesa que el sistema sea robusto frente a pequeños errores en el modelo del proceso. Por ejemplo, las salidas en estado estacionario deben ser iguales a unos valores deseados dados. Una forma de eliminar las perturbaciones de carga es estimarlas y utilizar una prealimentación de las perturbaciones estimadas (compárese con la Sección 9.4). En el diseño clásico, los errores estacionarios se eliminan introduciendo integradores. Esto mismo también puede hacerse en la formulación de la asignación de polos expuesta en este capítulo.

Supóngase que el sistema es el dado por (9.29). Por simplicidad se supone que la entrada y la salida son escalares. Sea u_e el nivel deseado para y en estado estacionario. Una forma de introducir un integrador es introducir un nuevo estado que integre el error $u_e - y$; es decir,

$$x_{n+1}(k+1) = x_{n+1}(k) + u_e(k) - Cx(k)$$



El sistema puede aumentarse con el nuevo estado, lo que da

$$\begin{bmatrix} x(k+1) \\ x_{n+1}(k+1) \end{bmatrix} = \begin{bmatrix} \Phi & 0 \\ -C & 1 \end{bmatrix} \begin{bmatrix} x(k) \\ x_{n+1}(k) \end{bmatrix} + \begin{bmatrix} \Gamma \\ 0 \end{bmatrix} u(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_e(k)$$

Se puede utilizar el diseño de la asignación de polos para el nuevo sistema. Si es alcanzable, entonces $x_{n+1} \rightarrow 0$ si u_e es constante. La ley de control tendrá la estructura

$$u(k) = -Lx(k) - l_{n+1}x_{n+1}(k) + l_e u_e(k)$$

donde el último término puede interpretarse como una prealimentación de una perturbación medible (véase la Fig. 9.9). Si los polos del sistema aumentado en lazo cerrado son estables, entonces —en estado estacionario— $y = u_e$. Ya que los valores propios de una matriz dependen de sus elementos de forma continua, el error estacionario será cero mientras el sistema en lazo cerrado sea estable, incluso aunque haya pequeños errores en el modelo del proceso.

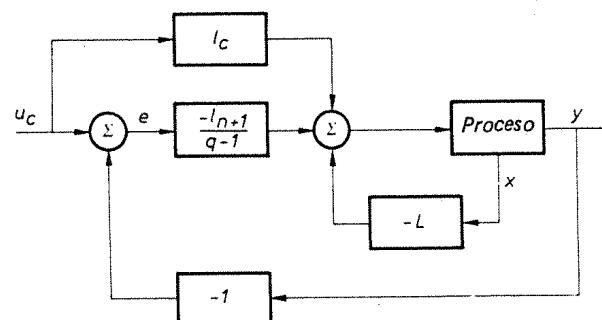


Figura 9.9 Diagrama de bloques mostrando la introducción de un integrador en el diseño de asignación de polos.

9.6. Conclusiones

El capítulo muestra cómo se pueden resolver los problemas del diseño del regulador y del servo empleando la asignación de polos y los observadores. La solución tiene tres componentes principales: la matriz de realimentación L , el observador y el modelo de la respuesta. La matriz de realimentación L se escoge empleando las técnicas expuestas en la Sección 9.2 de tal forma que las perturbaciones de carga disminuyan de forma apropiada. El observador se diseña considerando las perturbaciones de carga y el ruido de medida, como se ha expuesto en la Sección 9.3. El mayor compromiso está entre una rápida convergencia y la sensibilidad a los errores de medida. La matriz L y el observador cuidan de las propiedades de la regulación. El modelo de la respuesta y el modelo del proceso inverso se escogen

entonces para obtener el comportamiento deseado del servo en respuesta a las señales de mando.

El diseño de la asignación de polos se ha hecho aquí para el caso de una entrada y una salida. Con n parámetros en el vector de realimentación de estado es posible colocar n polos arbitrariamente, si el sistema es alcanzable. En el caso multivariable hay más grados de libertad. Esto hace posible determinar no sólo los polos, sino también algunos valores propios del sistema en lazo cerrado. En las referencias se pueden encontrar más detalles.

9.7. Problemas

- 9.1. Un sistema general discreto de segundo orden puede escribirse como

$$\begin{aligned} x(k+1) &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} x(k) + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(k) \\ y(k) &= [c_1 \quad c_2] x(k) \end{aligned}$$

Determine un controlador de realimentación de estado de la forma

$$u(k) = -Lx(k)$$

tal que la ecuación característica del sistema en lazo cerrado sea

$$z^2 + p_1 z + p_2 = 0$$

Utilícese el resultado para verificar el controlador de tiempo finito para el integrador doble dado en el Ejemplo 9.4.

- 9.2. Dado el sistema

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 1.0 & 0.1 \\ 0.5 & 0.1 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(k) \\ y(k) &= [1 \quad 1] x(k) \end{aligned}$$

determine un controlador de realimentación de estado lineal

$$u(k) = -Lx(k)$$

tal que los polos en lazo cerrado sean 0,1 y 0,25.

- 9.3. Determine un controlador de tiempo finito para el motor normalizado del Ejemplo A.2. Supóngase que $x(0) = [1, 1]^T$. Determine el intervalo de muestreo tal que la señal de control sea de módulo menor que la unidad. Puede suponerse que el valor máximo de $u(kh)$ se tiene para $k=0$.

- 9.4. Considérese el sistema del Problema 8.10. Cuando se muestrea el sistema con período $h = 0,2$ se tiene

$$x(kh+h) = \begin{bmatrix} 0.55 & 0.12 \\ 0 & 0.67 \end{bmatrix} x(kh) + \begin{bmatrix} 0.01 \\ 0.15 \end{bmatrix} u(kh)$$



- a) Determinese una ley de control de realimentación del estado tal que el polinomio característico en lazo cerrado sea

$$z^2 - 0.63z + 0.21$$

Esto corresponde a las especificaciones dadas en el Problema 8.10.

- b) Simúlese el sistema en lazo cerrado y compárese con los resultados anteriores.

9.5. El sistema

$$\begin{aligned}x(k+1) &= \begin{bmatrix} 0.78 & 0 \\ 0.22 & 1 \end{bmatrix}x(k) + \begin{bmatrix} 0.22 \\ 0.03 \end{bmatrix}u(k) \\y(k) &= [0 \ 1]x(k)\end{aligned}$$

representa el motor normalizado para el periodo de muestreo $h=0.25$. Determinense los observadores basados en la salida empleando cada uno de los siguientes métodos.

- a) Cálculo directo mediante la expresión (9.21).
b) Un sistema dinámico que dé $\hat{x}(k+1|k)$ empleando (9.24).
c) El observador de orden reducido de (9.27).

Use el observador de tipo de tiempo finito; es decir, los polos del observador deben estar en el origen.

- 9.6. Determinese un observador del estado completo basado en (9.24) para el sistema del tanque del Problema 3.10. Escójase la ganancia del observador de tal forma que el observador sea dos veces más rápido que el sistema en lazo abierto.

- 9.7. Considerese el observador de (9.27) y sea la ley de control dada por

$$u(k) = -L\hat{x}(k|k)$$

Mostrar que el controlador resultante puede escribirse como

$$\begin{aligned}\xi(k+1) &= \Phi_o\xi(k) + \Gamma_oy(k) \\u(k) &= C_o\xi(k) + D_oy(k)\end{aligned}$$

donde

$$\begin{aligned}\Phi_o &= (I - KC)(\Phi - \Gamma L) & \Gamma_o &= (I - KC)(\Phi - \Gamma L)K \\C_o &= -L & D_o &= -LK\end{aligned}$$

9.8. Dado el sistema discreto

$$\begin{aligned}x(k+1) &= \begin{bmatrix} 0.5 & 1 \\ 0.5 & 0.7 \end{bmatrix}x(k) + \begin{bmatrix} 0.2 \\ 0.1 \end{bmatrix}u(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix}v(k) \\y(k) &= [1 \ 0]x(k)\end{aligned}$$

donde v es una perturbación constante. Determinense unos controladores tales que se pueda eliminar la influencia de v en estado estacionario en cada uno de los siguientes casos.

- a) Se pueden medir el estado y v .
b) Se puede medir el estado.
c) Sólo se puede medir la salida.

- 9.9. Considerese el sistema de dos tanques del Problema 3.10 para $h=12$ s.

- a) Determinese un controlador de realimentación del estado tal que los polos en lazo cerrado estén dados por la ecuación característica

$$z^2 - 1.55z + 0.64 = 0$$

Esto corresponde a $\zeta=0.7$ y $\omega=0.027$ rad/s.

- b) Introdúzcase una señal de mando y determinese un controlador tal que el error en estado estacionario entre la señal de mando y la salida sea cero en estado estacionario; es decir, introdúzcase un integrador en el sistema.
c) Simúlese el sistema con los reguladores a) y b). Compárense los resultados con los obtenidos en el Problema 8.7.

- 9.10. Considerese el integrador doble con una perturbación de carga actuando sobre la entrada del proceso. La perturbación se puede describir como una sinusode con frecuencia ω_0 , pero con amplitud y fase desconocidas.

Disíñese un controlador de realimentación de estado y un observador tales que no haya error estacionario debido a la perturbación sinusoidal.

9.8. Referencias

La asignación de polos fue una de las primeras aplicaciones de la aproximación en el espacio de estados. Uno de los primeros en resolver el problema fue J. Bertram en 1959. La primera solución publicada está en

Rissanen, J. (1960): «Control System Synthesis by Analogue Computer Based on 'Generalized Linear Feedback' Concept,» *Proceedings of the Symposium on Analog Computation Applied to the Study of Chemical Processes*, Brussels, November 21-23, 1-13.

El tratamiento del caso multivariable de la asignación de polos puede encontrarse, por ejemplo, en

Rosenbrock, H. H. (1970): *State-Space and Multivariable Theory*. London: Nelson.

Wolowich, W. A. (1974): *Linear Multivariable Systems*. New York: Springer-Verlag.

Kailath, T. (1980): *Linear Systems*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.

También se tratan los observadores en los libros anteriores. El observador de orden reducido lo expuso por primera vez Luenberger en su Tesis Doctoral. Referencias más fáciles de encontrar son

Luenberger, D. G. (1964): «Observing the State of a Linear System,» *IEEE Trans. Mil. Electron.*, MIL-8, 74-80.

Luenberger, D. G. (1971): «An Introduction to Observers,» *IEEE Trans. Autom. Control*, AC-16, 596-603.

Los aspectos numéricos del cálculo de la realimentación de estado y de la ganancia del observador se exponen en

Miminis, G. S. y C. C. Paige (1982): «An Algorithm for Pole Assignment of Time Invariant Linear Systems,» *Int. Journal Control*, 35, No. 2, 341-54.

Diseño por asignación de polos basado en modelos de entrada-salida

Objetivo:

Solucionar el problema de la asignación de polos utilizando realimentación de la salida y diseño polinómico.

10.1. Introducción

En este capítulo se resuelve el problema del servo utilizando modelos de entrada-salida. El diseño se basa en la manipulación de polinomios. La idea fundamental consiste en comenzar con un regulador lineal y determinar sus parámetros para que el sistema en lazo cerrado tenga las propiedades deseadas. El problema se formula como un problema de asignación de polos; el método de diseño es simple de utilizar y comprender. Resulta fácil introducir restricciones, como altas ganancias de lazo a determinadas frecuencias y bajas ganancias a otras frecuencias.

El problema se formula en la Sección 10.2, y la solución se da en la Sección 10.3. La solución conduce a ecuaciones polinómicas lineales, que deben resolverse. En la Sección 10.4 se dan las condiciones para la existencia de soluciones al problema de diseño. Un procedimiento de diseño con algunos ejemplos se dan en la Sección 10.5. El diseño generalmente se basa en modelos aproximados. La Sección 10.6 trata la sensibilidad frente a errores de modelado. Las relaciones con otros métodos de diseño se dan en la Sección 10.7. En la Sección 10.8 se hace especial énfasis para relacionar el procedimiento de diseño con las especificaciones físicas naturales; por ejemplo, se analiza la sensibilidad frente a perturbaciones.

10.2. Formulación del problema

El problema del diseño se especifica habitualmente dando el modelo, los criterios y las leyes admisibles de control. El diseño considera un problema de servo, el

cual se formula como un problema de seguimiento del modelo. Por tanto, se desea encontrar una ley de control para obtener la respuesta apropiada a la señal de entrada. Esta formulación incluye la asignación de polos como un caso especial. Por tanto, una solución alternativa al problema del diseño se obtuvo en la Sección 9.5.

Dinámica del proceso

Se supone que el proceso tiene una entrada, u , y una salida medible, y . La relación entre u e y está dada por la función de transferencia discreta

$$H(z) = \frac{B(z)}{A(z)} \quad (10.1)$$

donde $A(z)$ y $B(z)$ son polinomios. Obsérvese que la función de transferencia discreta representa la dinámica del proceso, incluyendo el circuito mantenedor, el actuador, el sensor y el filtro anti-enmascaramiento. Se supone que A y B no tienen factores comunes. Recuérdese de la Sección 3.2 que el modelo de (10.1) puede representar un modelo discreto de un sistema continuo con una función de transferencia racional y un retardo de tiempo arbitrario.

El criterio

Las especificaciones del servo se expresan en términos de un modelo que da la respuesta deseada a las señales de referencia. La función de transferencia discreta deseada en lazo cerrado es

$$H_m(z) = \frac{B_m(z)}{A_m(z)} \quad (10.2)$$

donde A_m y B_m no tienen factores comunes. En general no es suficiente con especificar H_m . Con realimentación de la salida habrá una dinámica adicional que no se excita con la señal de referencia. Los resultados de la Sección 9.4 muestran que también es necesario especificar la dinámica del observador. Esto se realiza especificando el polinomio característico A_o del observador.

Perturbaciones

Como estudiaremos en la Sección 6.3, existen tres tipos fundamentales de perturbaciones: perturbaciones en la carga, errores de medida e incertidumbres de la planta. En el problema del servo también puede ser importante especificar la naturaleza de las señales de referencia.

Los modelos específicos de perturbaciones no se utilizan en el método de diseño de asignación de polos. Las perturbaciones se consideran indirectamente introduciendo restricciones en el modelo H_m , el polinomio del observador A_o , y la ley de control admisible.

La sensibilidad del sistema en lazo cerrado frente a los errores de modelado y el ruido de medida de alta frecuencia está influenciada por la función de transferencia del modelo H_m . Los requisitos de sensibilidad se contraponen con los requisitos de un modelo de seguimiento rápido. El polinomio del observador también influye en la sensibilidad a las perturbaciones en la carga y en el ruido de medida.

Leyes admisibles de control

El regulador tiene una salida, u , y dos entradas, la señal de referencia, u_c , y la medida de la salida, y . Una estructura lineal general para un regulador con estas entradas y salidas puede representarse por

$$u(k) = \frac{T_1(q)}{R_1(q)}u_c(k) - \frac{S_1(q)}{R_2(q)}y(k) \quad (10.3)$$

donde R_1 , R_2 , T_1 y S_1 son polinomios en el operador desplazamiento hacia adelante q . La ley de control de (10.3) puede escribirse como

$$R(q)u(k) = T(q)u_c(k) - S(q)y(k) \quad (10.4)$$

donde $R = R_1 R_2$, $T = T_1 R_2$ y $S = S_1 R_1$. Se supone que R es mónico; es decir, el coeficiente de la mayor potencia de R es la unidad. La ley de control de (10.4) representa la combinación de una prealimentación de la señal de referencia u_c con la función de transferencia discreta

$$H_{ff}(z) = \frac{T(z)}{R(z)} \quad (10.5)$$

y una realimentación de la medida de la salida y con la función de transferencia

$$H_{fb}(z) = \frac{S(z)}{R(z)} \quad (10.6)$$

Algunos requisitos de la ley de control se discutirán más adelante.

Causalidad. Las condiciones

$$\text{gra } R \geq \text{gra } T \quad (10.7)$$

$$\text{gra } R \geq \text{gra } S \quad (10.8)$$

aseguran que las funciones de transferencia de la prealimentación y realimentación son causales. Si el tiempo de cálculo de la señal de control en el computador es solamente una pequeña fracción del período de muestreo, entonces es normal exigir que

$$\text{gra } R = \text{gra } T = \text{gra } S \quad (10.9)$$

Esto significa que no hay retardo de tiempo en el regulador.

Si el tiempo de cálculo se aproxima al período de muestreo, la restricción correspondiente es

$$\text{gra } R = 1 + \text{gra } T = 1 + \text{gra } S \quad (10.10)$$

Esto significa que hay un retardo en la ley de control de un período de muestreo.

La restricción (10.9) se usa normalmente como un caso típico. Los posibles retardos por cálculos en el computador se pueden incluir en el modelo del proceso en lugar de incluirlos en el regulador. (Compárese con la Sección 3.2.)

Perturbaciones e incertidumbres de la planta. Las perturbaciones se manipulan introduciendo especificaciones adicionales en las leyes de control admisibles. Las restricciones se expresan adecuadamente en términos de la ganancia del lazo

$$H_{lg}(z) = \frac{S(z)B(z)}{R(z)A(z)} \quad (10.11)$$

Para estar seguros que las perturbaciones de baja frecuencia producen errores pequeños, la ganancia del lazo $H_{lg}(\exp i\omega)$ debe ser grande para las bajas frecuencias. Esto puede hacerse exigiendo que el polinomio R tenga la forma

$$R(z) = (z - 1)^{\ell} R'_1(z) \quad (10.12)$$

con un ℓ adecuado. Este es el principio clásico del control integral. La condición también asegura que el sistema en lazo cerrado es insensible a incertidumbres en el modelado que influyen en la transmisión de señales de baja frecuencia.

La influencia de incertidumbres en la dinámica de alta frecuencia del proceso y el ruido de medida de alta frecuencia pueden reducirse eligiendo S y R de modo que la ganancia del lazo disminuya rápidamente para altas frecuencias (véase el Teorema 5.4). Obsérvese que el muestreo es muy útil en este caso porque la combinación de los filtros anti-enmascaramiento y el circuito mantenedor reducen la transmisión efectiva de la señal para las frecuencias superiores a la frecuencia de Nyquist. De este modo la elección del período de muestreo es muy importante para la sensibilidad frente a perturbaciones de alta frecuencia y errores de modelado.

También hay situaciones en las cuales la ganancia del lazo debe ser baja para algunas frecuencias, las cuales son menores que el ancho de banda del servo. Esto puede ocurrir cuando hay perturbaciones periódicas muy grandes que no puede reducirlas el servo o cuando hay cambios rápidos de fase en la transmisión de la señal. Ejemplos típicos son las perturbaciones debidas a ondas en mantenimiento de satélites en órbita, barcos de prospección y modos flexibles de baja frecuencia en vehículos aeroespaciales. Estos tipos de perturbaciones se manipulan clásicamente introduciendo filtros de eliminación de banda que bloquean la transmisión de señal a ciertas frecuencias. En el diseño con asignación de polos, se pueden obtener los mismos efectos introduciendo un factor en el polinomio S .

10.3. Solución

El problema del diseño formulado en la sección anterior puede resumirse del siguiente modo: Sea un modelo del proceso especificado por la relación entrada-salida

$$A(q)y(k) = B(q)u(k) \quad (10.13)$$

donde u es la señal de control e y es la señal medida en la salida. Encontrar una ley de control admisible de la forma de (10.4) tal que el sistema en lazo cerrado tenga la relación entrada-salida definida por la función de transferencia discreta (10.2), y un observador con el polinomio característico $A_o(z)$.

En la Figura 10.1 se muestra un diagrama de bloques del sistema en lazo cerrado. En la figura, también se representan las perturbaciones. Para simplificar la escritura en el análisis que se realiza a continuación, se suprime los argumentos de los polinomios y de las funciones temporales. La relación entrada-salida para el sistema en lazo cerrado se obtiene eliminando u entre las ecuaciones (10.4) y (10.13). Por tanto,

$$(AR + BS)y = BTu_c \quad (10.14)$$

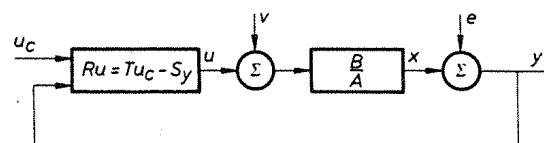


Figura 10.1 Diagrama de bloques del sistema en lazo cerrado.

Exigiendo que la relación entrada-salida sea equivalente a (10.2) se obtiene

$$\frac{BT}{AR + BS} = \frac{B_m}{A_m} \quad (10.15)$$

El problema del diseño se reduce, de este modo, a encontrar los polinomios R , S y T que satisfagan (10.15) y los requisitos adicionales de los controles admisibles.

Ejemplos

Existen algunas posibilidades para elegir los polinomios R , S y T . Se dan dos ejemplos a modo de ilustración.

Ejemplo 10.1 – Solución con prealimentación pura

Es fácil verificar que los polinomios

$$R = BA_m \quad S = 0 \quad T = AB_m$$

satisfacen (10.15). Con esta elección de polinomios se deduce que la función de transferencia discreta de la realimentación es

$$H_{fb} = \frac{S}{R} = 0$$

y que la función de transferencia discreta de la prealimentación es

$$H_{ff} = \frac{T}{R} = \frac{B_m A}{A_m B}$$

De este modo, el controlador es un compensador con prealimentación pura sin realimentación. El compensador simplemente cancela la dinámica del proceso y añade la dinámica deseada. Aunque la solución satisface (10.15), no satisface las otras restricciones de la ley de control. Puesto que la realimentación es nula, la condición de que la ganancia del lazo sea grande para frecuencias bajas, obviamente, no se satisface. Todos los polos y ceros del sistema se cancelan, de modo que el sistema será inestable si el proceso tiene polos o ceros fuera del círculo unidad.

Ejemplo 10.2 – Error de realimentación

Si los polinomios S y T son iguales, la ley de control (10.4) puede escribirse como

$$Ru = S(u_c - y) = Se$$

donde e es el error de control. Esto significa que la ley de control se basa solamente en la realimentación del error. Es fácil de verificar que los polinomios

$$R = B(A_m - B_m) \quad S = T = AB_m$$

satisfacen (10.15). Esta elección representa una solución de realimentación. Sin embargo, no hay garantía de que la realimentación tenga altas ganancias a bajas frecuencias. Obsérvese que todos los polos y ceros del sistema están cancelados. El sistema en lazo cerrado será inestable si hay polos o ceros fuera de círculo unidad. La necesidad de cancelar polos y ceros se debe a la restricción impuesta por la realimentación del error.

Los ejemplos muestran claramente que hay algunas elecciones posibles para R , S y T que satisfacen (10.15). A continuación se presentan métodos racionales para la elección de los parámetros que satisfacen (10.15) y las restricciones de las leyes admisibles de control.

Cancelación de polos y ceros

Se deduce de (10.15) que los polos del sistema en lazo cerrado son las soluciones de la ecuación característica

$$AR + BS = 0 \quad (10.16)$$

Los ceros del sistema en lazo cerrado son los ceros de los polinomios B y T . El orden del sistema en lazo cerrado generalmente es mayor que el orden del modelo. Para satisfacer la condición (10.15), tiene que haber cancelación de polos y ceros.

En primer lugar consideremos los ceros en lazo abierto, es decir, los ceros del polinomio B . Si un factor de B no lo es de B_m , entonces, tiene que ser un factor de $AR + BS$, por tanto tiene que ser cancelado por un polo en lazo cerrado. Puesto que el sistema en lazo cerrado tiene que ser estable, se deduce que sólo pueden cancelarse los ceros estables. Por tanto, factorizamos B como

$$B = B^+ B^- \quad (10.17)$$

donde B^- tiene todos sus ceros fuera del círculo unidad y B^+ tiene todos sus ceros dentro del círculo unidad. Para obtener una factorización única el coeficiente de la mayor potencia de B^+ se hace igual a la unidad. Entonces se dice que el polinomio B^+ es mónico.

Puesto que B^- no puede ser un factor de $AR + BS$, se deduce que tiene que dividir a B_m ; es decir,

$$B_m = B^- B'_m \quad (10.18)$$

Esto implica que los ceros inestables del proceso no pueden cambiarse, pero deben ser incluidos en B_m . Puesto que B^+ es un factor de $AR + BS$, se deduce que también es un factor de R . Por tanto,

$$R = B^+ R' \quad (10.19)$$

La ecuación (10.15) puede escribirse como

$$\frac{B^+ B^- T}{B^+(AR' + B^- S)} = \frac{B^- B'_m}{A_m}$$

Esta condición se reduce a

$$\frac{T}{AR' + B^- S} = \frac{B'_m}{A_m}$$

y se deduce que A_m es un factor de $AR' + B^- S$. Además, se sigue de la Sección 9.5 que el polinomio del observador se cancela en la función de transferencia que relaciona la señal de referencia con la salida. Por tanto, se deduce que A_o es un factor de $AR + BS$. De esto se obtienen las siguientes condiciones:

$$AR' + B^- S = A_o A_m \quad (10.20)$$

y

$$T = B'_m A_o \quad (10.21)$$

La ecuación característica en lazo cerrado llega a ser

$$AR + BS = B^+ A_o A_m \quad (10.22)$$

De este modo, los polos en lazo cerrado son los ceros estables cancelados del proceso, B^+ , los polos del modelo, A_m y los polos del observador, A_o .

Nota. También puede desecharse la cancelación de los polos estables del proceso. Esto debe incluirse en A_o . Entonces, el polinomio A_o pierde su interpretación como polinomio del observador. De (10.20) y (10.21) se ve cómo los polos cancelados son factores de S y T .

Restricciones prácticas

En la práctica conviene disponer de exigencias estrictas para las cancelaciones permitidas. Algunas veces, la cancelación puede no ser deseable. En otros casos puede ser razonable la cancelación de ceros que están suficientemente bien amortiguados. Un método para expresar esto formalmente consiste en introducir una región D en el plano complejo correspondiente a los modos con amortiguamiento absoluto y relativo suficiente. Solamente los ceros dentro de D pueden cancelarse. Un ejemplo de una región se muestra en la Figura 10.2. Recordemos de la Sección 3.6 que las líneas con amortiguamiento relativo constante son espirales logarítmicas en el plano z y que las líneas con amortiguamiento absoluto constante son circunferencias.

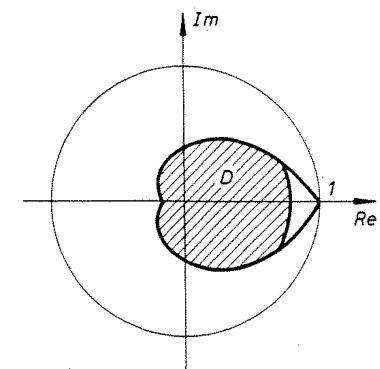


Figura 10.2 Región D tal que sus puntos tienen amortiguamientos relativos y absolutos que son mínimos.

Resumen

En resumen, el problema de la asignación de polos se resuelve encontrando R' y S que satisfagan (10.20) y tales que B está factorizado como $B^+ B^-$. El polinomio B^+ es mónico y tiene todos sus ceros dentro de D ; B^- tiene todos sus ceros fuera de D . Además, B^- , debe ser un factor de B_m . Finalmente, los polinomios R y T deberían elegirse por (10.19) y (10.21).

10.4. Un problema algebraico

El estudio de la sección anterior da alguna idea sobre la elección de los polinomios R , S y T . Pero todavía se mantiene un problema, esto es, determinar dos polinomios R' y S que satisfagan (10.20) y los otros requisitos sobre la realimentación



admisible. El problema matemático básico es determinar dos polinomios X e Y que satisfagan la ecuación polinómica lineal

$$AX + BY = C \quad (10.23)$$

donde A , B y C son polinomios conocidos. Este es un problema de álgebra elemental bien conocido.

Una digresión

La ecuación (10.23) resulta extraña a primera vista porque deben determinarse dos incógnitas a partir de una ecuación. Un problema simple, pero afín, de álgebra superior aclara este caso.

Ejemplo 10.3 – La ecuación diofántica

Considérese la ecuación

$$3x + 2y = 5 \quad (10.24)$$

donde x e y son enteros. Esta ecuación se denomina *ecuación diofántica* y se debe a Diofanto (siglo v a. de J.) que fue uno de los inventores originarios del álgebra.

Es obvio que $x=1$ e $y=1$ es una solución. Otra solución puede encontrarse incrementando x en 2 y decrementando y en 3. Por tanto, si x_0 e y_0 satisfacen (10.24), entonces otra solución es

$$\begin{aligned} x &= x_0 + 2n \\ y &= y_0 - 3n \end{aligned} \quad (10.25)$$

donde n es un entero. A continuación se listan algunas soluciones

$$\begin{array}{ccccccc} x: & -5 & -3 & -1 & 1 & 3 & 5 & 7 \\ y: & 10 & 7 & 4 & 1 & -2 & -5 & -8 \end{array}$$

Se deduce de (10.25) que si se conoce una solución x_0 , y_0 , es posible añadir o sustraer 2 a x_0 hasta encontrar una solución única con

$$0 \leq x < 2$$

Análogamente, también hay una única solución tal que

$$0 \leq y < 3$$

La ecuación (10.24) está relacionada estrechamente con la ecuación (10.23) porque los enteros y los polinomios de coeficientes reales obedecen a las mismas reglas algebraicas. Los enteros y los polinomios se pueden sumar y multiplicar con las reglas usuales. Sin embargo, el cociente de dos enteros (o polinomios) no es necesariamente un entero (o polinomio). En la terminología algebraica esto se expresa diciendo que los enteros y polinomios con coeficientes reales son *anillos*.

Otro ejemplo muestra cómo una ecuación de la forma de (10.24) puede no tener solución.

Ejemplo 10.4

Consideremos la ecuación

$$4x + 6y = 1$$

donde x e y son enteros. Como el término de la izquierda es un número par y el de la derecha es un número impar, es evidente que la ecuación no tiene solución. La dificultad de encontrar una solución se debe a que los números 4 y 6 tienen el factor común 2, mientras que el término de la derecha no lo tiene.

Resultado fundamental

Los ejemplos 10.3 y 10.4 revelan esencialmente los problemas que encierra la ecuación (10.23). Ahora lo que importa es realizar un análisis formal de la ecuación. Las soluciones están caracterizadas por el siguiente resultado.

Teorema 10.1. Sean A , B y C polinomios de coeficientes reales. Entonces la ecuación (10.23) tiene una solución si y sólo si el máximo común divisor de A y B divide a C . \square

La demostración del teorema se da en el Apéndice D.

Se deduce de (10.23) que si X_0 , Y_0 es una solución, entonces

$$\begin{aligned} X &= X_0 + QB \\ Y &= Y_0 - QA \end{aligned} \quad (10.26)$$

es también una solución, donde Q es un polinomio arbitrario. Si existe una solución pueden encontrarse infinitas soluciones añadiendo o sustrayendo un múltiplo apropiado de A y B .

Corolario. Existe una solución única de (10.23) tal que

$$\text{gra } X < \text{gra } B \quad (10.27)$$

o bien

$$\text{gra } Y < \text{gra } A \quad (10.28)$$

10.5. Un procedimiento de diseño

La comprensión de las propiedades algebraicas de (10.23) conduce de nuevo al problema de diseño por asignación de polos. La realimentación en (10.21) está determinada únicamente por los polinomios R , S y T . En la Sección 10.3 se muestra que T está definido por (10.21). También R' y S deben satisfacer (10.20). Como los polinomios A y B no tiene ningún factor común, se deduce del teorema 10.1 que existen polinomios R' y S que satisfacen (10.20). También existen infinitos polinomios



que satisfacen la ecuación. Todas las soluciones producen la misma función de transferencia discreta entre la señal de referencia y la señal de salida. A continuación se estudiará si puede encontrarse una solución que también satisface las otras restricciones de las leyes admisibles de control.

Causalidad

La ley de control (10.4) es causal si se satisfacen las inecuaciones (10.7) y (10.8). Las consecuencias de estas inecuaciones se resumen en el Teorema 10.2.

Teorema 10.2. Existe una solución causal del diseño por asignación de polos si

$$\text{gra } A_m - \deg B_m \geq \text{gra } A - \text{gra } B \quad (10.29)$$

$$\text{y} \quad \text{gra } A_o \geq 2\text{gra } A - \text{gra } A_m - \text{gra } B^+ - 1 \quad (10.30)$$

Demostración. Las ecuaciones (10.19) y (10.20) producen

$$AR + BS = B^+ A_o A_m$$

Puesto que $\text{gra } S \leq \text{gra } R$ y $\text{gra } B < \text{gra } A$, se deduce que

$$\text{gra } AR = \text{gra } (AR + BS) = \text{gra } B^+ A_o A_m$$

Por tanto,

$$\text{gra } R = \text{gra } A_o + \text{gra } A_m + \text{gra } B^+ - \text{gra } A$$

Además, se demuestra de (10.21) que

$$\text{gra } T = \text{gra } A_o + \text{gra } B'_m$$

Ahora estudiaremos las consecuencias de (10.7), lo cual implica que

$$\text{gra } A_o + \text{gra } A_m + \text{gra } B^+ - \text{gra } A \geq \text{gra } A_o + \text{gra } B'_m$$

Por tanto,

$$\text{gra } A_m - \text{gra } B'_m \geq \text{gra } A - \text{gra } B^+$$

Restando B^- en ambos lados y teniendo en cuenta (10.18) se obtiene (10.29). Se deduce del corolario del Teorema 10.1 que existe una solución de (10.20) tal que

$$\text{gra } S < \text{gra } A$$

La elección de $\text{gra } S = \text{gra } A - 1$ y la inecuación (10.8) implica que

$$\text{gra } A_o + \text{gra } A_m + \text{gra } B^+ - \text{gra } A \geq \text{gra } A - 1$$

Reagrupando los términos se obtiene (10.30). □

Nota 1. Es fácil de comprender intuitivamente la condición (10.29). Esto implica que el retardo en el modelo H_m debe ser al menos tan grande como el retardo en el sistema H .

Nota 2. La condición (10.30) implica que el grado del polinomio observador A_o debe ser lo suficientemente alto para obtener una ley de control causal.

Alta ganancia a bajas frecuencias

Es útil disponer de altas ganancias de realimentación para bajas frecuencias para obtener un sistema que sea insensible a los errores de modelado a bajas frecuencias y a las perturbaciones de baja frecuencia. Esto puede llevarse a cabo exigiendo que $(z-1)^l$ sea un factor de R . Introduciendo

$$R = (z-1)^l R_1 \quad (10.31)$$

la ecuación de diseño (10.20) se convierte en

$$A(z-1)^l R'_1 + B^- S = A_o A_m \quad (10.32)$$

Con cálculos sencillos se demuestra que la condición de causalidad (10.7) y (10.8) se satisface si (10.29) se mantiene y si

$$\text{gra } A_o \geq 2\text{gra } A - \text{gra } A_m - \text{gra } B^+ + l - 1 \quad (10.33)$$

El grado del polinomio observador debe incrementarse en l en relación con (10.30).

Un algoritmo de diseño

Los resultados obtenidos pueden resumirse en un procedimiento directo para resolver el problema del diseño por asignación de polos:

Algoritmo 10.1 – Diseño por asignación de polos

Datos. Un modelo del proceso especificado por su función de transferencia discreta B/A , un polinomio del observador A_o y especificaciones en términos de la función de transferencia deseada B_m/A_m . Además, se especifica la región de estabilidad D .

Condiciones. Se supone que los datos satisfacen (10.18), (10.29) y (10.23).

Paso 1. Factorizar B y B_m como

$$B = B^- B^+ \quad B_m = B^- B'_m$$

donde B^+ es mónico y tiene todos sus ceros dentro de la región de estabilidad D , y B^- tiene todos sus ceros fuera de D .

Paso 2. Resolver la ecuación

$$(z - 1)^l A R'_1 + B^- S = A_o A_m$$

respecto de R'_1 y S . Elegir una solución tal que

$$\text{gra } S < l + \text{gra } A$$

y

$$\text{gra } R'_1 = \text{gra } A_o + \text{gra } A_m - \text{gra } A - l$$

Paso 3. Entonces, la ley de control es

$$R u = T u_c - S y$$

donde

$$R = B^+ R' \quad T = B'_m A_o \quad R' = (z - 1)^l R'_1.$$

Nota. De la identidad de polinomios se deduce que R' es mónico. Esto implica que R también es mónico.

Solución de ecuaciones polinómicas lineales

Para efectuar el diseño es necesario resolver una ecuación polinómica lineal. Hay varios modos de hacerlo.

Una posibilidad consiste en introducir los polinomios R' y S con coeficientes desconocidos y órdenes determinados e identificar los coeficientes de potencias iguales de z . Entonces se obtiene un sistema de ecuaciones lineales. Del Teorema 10.1 se deduce que la ecuación lineal puede resolverse siempre si A y B^- no tienen factores comunes. Las ecuaciones obtenidas pueden simplificarse si A y B están dadas en forma factorial. Este método se ilustra en los siguientes ejemplos.

Ejemplos

Se presentan algunos ejemplos para ilustrar el diseño de sistemas de control con el algoritmo de asignación de polos.

Ejemplo 10.5—Motor con cancelación de ceros del proceso

La función de transferencia discreta de un motor CC puede escribirse como

$$H(z) = \frac{K(z - b)}{(z - 1)(z - a)} \quad (10.34)$$

(véase el Ejemplo A.2) donde

$$K = e^{-h} - 1 + h$$

$$a = e^{-h}$$

$$b = 1 - \frac{h(1 - e^{-h})}{e^{-h} - 1 + h}$$

Observemos que $b < 0$; es decir, el cero está sobre el eje real negativo. Se supone, en primer lugar, que el sistema deseado en lazo cerrado está caracterizado por la función de transferencia discreta

$$H_m(z) = \frac{z(1 + p_1 + p_2)}{z^2 + p_1 z + p_2} \quad (10.35)$$

La función de transferencia discreta H tiene un cero $z = b$ el cual no está incluido en H_m . Con la especificación dada es necesario cancelar el cero $z = b$. Factorizando B como

$$B^+ = z - b$$

$$B^- = K$$

Entonces

$$B'_m = \frac{B_m}{K} = \frac{z(1 + p_1 + p_2)}{K}$$

El grado del polinomio observador está definido por (10.30), el cual verifica que $\text{gra } A_o \geq 0$. Eligiendo

$$A_o(z) = 1$$

El grado de los polinomios R' y S están definido por

$$\text{gra } R' = \text{gra } A_o + \text{gra } A_m - \text{gra } A = 0$$

$$\text{gra } S = \text{gra } A - 1 = 1$$

Introduzcamos R' como un polinomio de orden cero y S como un polinomio de primer orden en la ecuación de diseño. Entonces se obtiene la siguiente identidad de polinomios.

$$(z - 1)(z - a)r_0 + K(s_0 z + s_1) = z^2 + p_1 z + p_2$$

Igualando los coeficientes de potencias iguales de z se obtienen las ecuaciones

$$r_0 = 1$$

$$-(1 + a)r_0 + Ks_0 = p_1$$

$$ar_0 + Ks_1 = p_2$$

Por tanto,

$$r_0 = 1$$

$$s_0 = \frac{1 + a + p_1}{K}$$

$$s_1 = \frac{p_2 - a}{K}$$

La ecuación (10.21) da

$$T(z) = A_o(z)B'_m(z) = \frac{z(1 + p_1 + p_2)}{K} = t_0 z$$

La ley de control (véase paso 3 del Algoritmo 10.1) puede escribirse como

$$u(k) = t_0 u_c(k) - s_0 y(k) - s_1 y(k - 1) + bu(k - 1) \quad (10.36)$$

En la Figura 10.3 se incluye una simulación de la respuesta frente a un salto en escalón. Obsérvese el rizado en la señal de control, causada por la cancelación del cero en el eje real negativo. El rizado no es observable en la salida en los instantes de muestreo. Sin embargo, se observa un rizado en la salida entre los instantes de muestreo. La amplitud del rizado en la salida depende del período de muestreo. Decae rápidamente cuando el período de muestreo decrece.

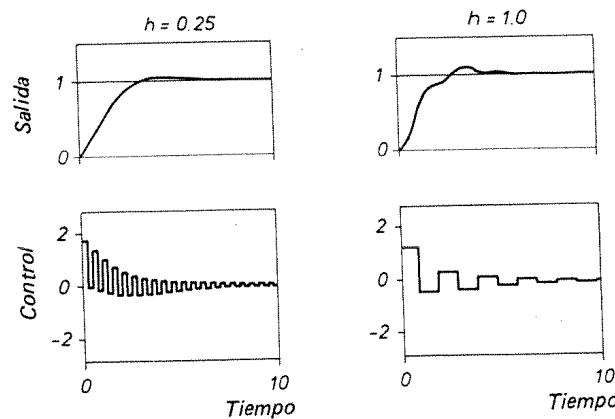


Figura 10.3 Respuesta salto de un motor con control de asignación de polos. Las especificaciones son $\zeta=0.7$ y $\omega=1$. Los períodos de muestreo son $h=0.25$ y $h=1.0$. El cero del proceso se cancela.

Ejemplo 10.6—Motor sin cancelación de ceros de proceso

Consideremos el mismo motor del Ejemplo 10.5, pero suponiendo que la función de transferencia en lazo cerrado deseada es

$$H_m(z) = \frac{1 + p_1 + p_2}{1 - b} \frac{z - b}{z^2 + p_1 z + p_2} \quad (10.37)$$

Observemos que el cero del proceso sobre el eje negativo, ahora es también un cero de la función de transferencia deseada en lazo cerrado. Esto significa que el cero no debe ser cancelado por el regulador. Factoricemos B como

$$B^+ = 1$$

$$B^- = K(z - b)$$

Por tanto,

$$B'_m = \frac{1 + p_1 + p_2}{K(1 - b)}$$

El grado del polinomio observador es

$$\text{gra } A_o \geq 2 \text{ gra } A - \text{gra } A_m - \text{gra } B^+ - 1 = 1$$

Por consiguiente, el polinomio del observador debería ser al menos de primer grado. Se elige un observador de tiempo finito:

$$A_o(z) = z$$

Entonces, el grado mínimo de los polinomios R y S es

$$\begin{aligned} \text{gra } R &= \text{gra } A_m + \text{gra } A_o - \text{gra } A = 1 \\ \text{gra } S &= \text{gra } A - 1 = 1 \end{aligned}$$

La ecuación de diseño de (10.20) puede escribirse como

$$(z - 1)(z - a)(z + r_1) + K(z - b)(s_0 z + s_1) = z^3 + p_1 z^2 + p_2 z \quad (10.38)$$

Para determinar r_1 , hagamos $z = b$ en (10.38). Se obtiene

$$(b - 1)(b - a)(b + r_1) = b^3 + p_1 b^2 + p_2 b$$

lo cual produce

$$r_1 = -b + \frac{b(b^2 + p_1 b + p_2)}{(b - 1)(b - a)}$$

Ahora hagamos $z = 1$ y $z = a$ en (10.38). Esto da

$$\begin{aligned} K(1 - b)(s_0 + s_1) &= 1 + p_1 + p_2 \\ K(a - b)(s_0 a + s_1) &= a^3 + p_1 a^2 + p_2 a \end{aligned}$$

de donde se puede determinar s_0 y s_1 . La ecuación (10.21) da

$$T(z) = A_o B'_m = z \frac{1 + p_1 + p_2}{K(1 - b)} = t_0 z$$

Entonces, la ley de control es

$$u(k) = t_0 u_o(k) - s_0 y(k) - s_1 y(k - 1) - r_1 u(k - 1)$$

Obsérvese que esta ley de realimentación tiene la misma forma que (11.36). Sin embargo, los coeficientes son diferentes. La Figura 10.4 muestra la respuesta del sistema ante un salto escalón.

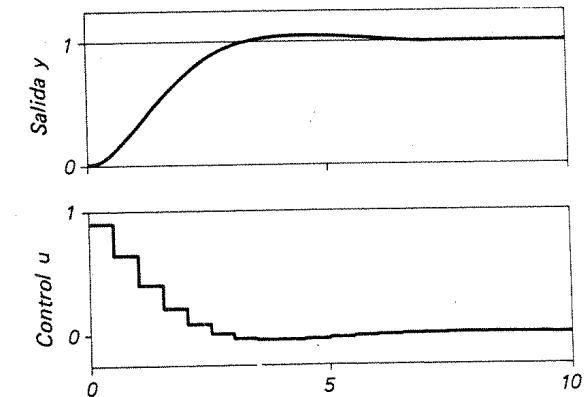


Figura 10.4 Respuesta salto de un motor con control de asignación de polos para $h=0.5$. Las especificaciones son $\zeta=0.7$ y $\omega=1$. El cero del proceso no se cancela.



Comparando con la Figura 10.3 se observa que la señal de control es mucho más suave, no hay rizado. La respuesta comienza también más lentamente, ya que A_o es de grado superior al del Ejemplo 10.5.

10.6. Sensibilidad a los errores de modelado

Es un error suponer que los modelos del proceso utilizados en un diseño de control son fiables. Por tanto, es muy importante comprender cómo influyen los errores de modelado en las propiedades del sistema en lazo cerrado, lo cual discutiremos en esta sección. Se supone que el diseño se basa en el modelo $H = B/A$, mientras que el verdadero modelo es $H^0 = B^0/A^0$.

Estabilidad

El siguiente resultado describe la influencia de los errores de modelado en la estabilidad del sistema en lazo cerrado.

Teorema 10.3. Consideremos un diseño por asignación de polos basado en el modelo aproximado

$$H = B/A$$

Sea H^0 la función de transferencia discreta del sistema que se va a controlar. Supongamos que H y H^0 tienen el mismo número de polos fuera del círculo unidad y que H_m es estable. Entonces el sistema en lazo cerrado referido a $H^0(z)$ es estable si

$$|H(z) - H^0(z)| < \left| \frac{H(z)T(z)}{H_m(z)S(z)} \right| = \left| \frac{H(z)}{H_m(z)} \right| \left| \frac{H_{ff}(z)}{H_{fb}(z)} \right| \quad (10.39)$$

para $|z|=1$, donde H_{ff} y H_{fb} están definidas por (10.5) y (10.6).

Democión. La demostración se basa en los resultados generales sobre sensibilidad de la Sección 5.2. Introduzcamos la ganancia de lazo

$$H_{lg} = \frac{BS}{AR} \quad (10.40)$$

Se deduce de (10.22) que

$$1 + H_{lg} = 1 + \frac{BS}{AR} = \frac{AR + BS}{AR} = \frac{B^+ A_o A_m}{AR}$$

De (10.21) se deduce que

$$A_o = \frac{T}{B'_m} = \frac{TB^-}{B'_m B^-} = \frac{TB^-}{B_m}$$

Por tanto,

$$1 + H_{lg} = \frac{B^+ B^- T A_m}{A R B_m} = \frac{B A_m T}{A B_m R} = \frac{HT}{H_m R}$$

Después de multiplicar por S/R , la condición dada por (10.39) puede escribirse como

$$\left| \frac{S}{R} H - \frac{S}{R} H^0 \right| < \left| \frac{HT}{H_m R} \right| = |1 + H_{lg}|$$

o

$$|H_{lg} - H_{lg}^0| < |1 + H_{lg}|$$

El resultado puede deducirse ahora con el Teorema 5.4. \square

Nota 1. Es sencilla la aplicación de este teorema. Cuando se realiza un diseño, el término de la derecha de (10.39) puede calcularse fácilmente para $z = \exp(i\omega h)$, y no depende de la función de transferencia discreta verdadera. Entonces, los requisitos sobre la precisión del sistema pueden expresarse en términos de las condiciones en el dominio de la frecuencia.

Nota 2. Obsérvese que la inecuación (10.39) se satisface automáticamente si

$$|H_{lg}| = \left| \frac{SB}{RA} \right| < \frac{1}{3}$$

y

$$|H_{lg}^0| = \left| \frac{SB^0}{RA^0} \right| < \frac{1}{3}$$

Por tanto, es suficiente comprobar (10.39) para frecuencias donde la ganancia del lazo es mayor que $1/3$.

Nota 3. Obsérvese que la precisión relativa es

$$\frac{|H(z) - H^0(z)|}{|H(z)|} \leq \frac{1}{|H_m(z)|} \left| \frac{H_{ff}(z)}{H_{fb}(z)} \right|$$

La condición (10.39) tiene una interpretación física. Consideremos en primer lugar la relación H/H_m . La función de transferencia discreta H del proceso normalmente es grande para bajas frecuencias y decrece para altas frecuencias (véase la Fig. 10.5). La función de transferencia discreta H_m deseada para el sistema en lazo cerrado es, normalmente, la unidad para bajas frecuencias. Hay un pequeño incremento en torno a la frecuencias de cruce y H_m decrece para altas frecuencias (compárese con la Fig. 7.9). También se muestra en la Figura 10.5 la respuesta en frecuencia de H_m . La relación H/H_m es fácil de obtener a partir de la figura. De la figura resulta evidente que es suficiente tener buena precisión en el modelo solamente en determinados rangos de frecuencias. También pueden determinarse las consecuencias del cambio del ancho de banda deseado para el sistema en lazo cerrado. Los

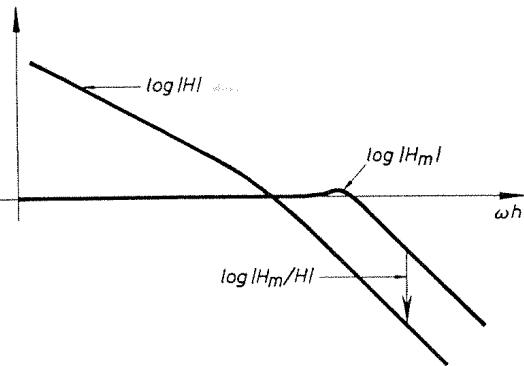


Figura 10.5 Diagramas de Bode para H y H_m . En la figura se puede encontrar fácilmente la relación H/H_m que aparece en (10.39).

requisitos sobre la precisión del modelo pueden hacerse menos estrictos si se decrementa el ancho de banda del sistema en lazo cerrado. Si se desea incrementar el ancho de banda, se necesitará un modelo más preciso. La razón

$$\left| \frac{H_{ff}}{H_{fb}} \right| = \left| \frac{T}{S} \right|$$

es el cociente entre las funciones de transferencia de prealimentación y realimentación. El coeficiente será la unidad si el procedimiento de diseño proporciona solamente realimentación del error. Los requisitos sobre la precisión del modelo son menores para frecuencias donde la ganancia de prealimentación es mayor que la ganancia de realimentación.

Precisión de la función de transferencia en lazo cerrado

Con algunos cálculos, la función de transferencia discreta del sistema en lazo cerrado puede expresarse por

$$H_{cl} = \frac{B^0 T}{A^0 R + B^0 S} = \frac{T/R}{A^0/B^0 + S/R} = H_m \frac{1}{1 + (RB^-/A_o A_m)(1/H^0 - 1/H)} \quad (10.41)$$

donde A_o es el polinomio observador. Esto muestra cómo los errores en el modelo se reflejan en errores de la función de transferencia pulsante en lazo cerrado. Se observa claramente de la expresión que los errores son pequeños cuando las funciones de transferencia discreta en lazo abierto H y H^0 son grandes.

10.7. Relación con otros métodos de diseño

La asignación de polos es un método general de diseño de sistemas de una entrada y una salida. Otros muchos métodos de diseño se pueden interpretar como diseño por asignación de polos.

Lugar de las raíces

El método del lugar de las raíces es una técnica clásica de diseño de sistemas de control. El método se basa en la idea de intentar situar los polos en lazo cerrado en las posiciones deseadas. Por tanto, está estrechamente relacionado con la asignación de polos. En este método, los polinomios R y S se eligen, en primer lugar, como $R=1$ y $S = K$, lo cual corresponde al control proporcional. Entonces se varía la ganancia K y se estudian las raíces de la ecuación característica

$$A + KB = 0 \quad (10.42)$$

Las raíces de (10.42) pueden esbozarse fácilmente variando K . Si no se puede obtener una situación razonable para los polos, se incrementan los órdenes de los polinomios R y S utilizando reglas heurísticas. En este caso, se repite el procedimiento.

Es evidente que el procedimiento de diseño por asignación de polos dado en la Sección 10.5 es más fácil de usar que el diseño por el método del lugar de las raíces. Todos los polos se posicionan en una operación. Por otro lado, también es evidente que el método del lugar de las raíces proporciona un regulador simple, siempre que el problema tenga solución. Sin embargo, el método general de diseño dará un regulador cuya complejidad está determinada por la complejidad del modelo del proceso utilizado en el diseño.

Compensación dipolar

En el diseño clásico de servos con realimentación del error es costumbre introducir un dipolo en la función de transferencia en lazo cerrado deseada; es decir, un polo y un cero muy próximos y además próximos al origen. El dipolo tiene solamente una influencia marginal en la transmisión de las señales de alta frecuencia, aunque influye en la transmisión de señales de baja frecuencia. No es inmediato comprender intuitivamente la utilidad de introducir un modo dinámico en el regulador, el cual está prácticamente cancelado, es decir, es casi inalcanzable, o inobservable. La compensación dipolar puede interpretarse a la luz del diseño por asignación de polos, y la dinámica introducida por el dipolo puede interpretarse como un observador dinámico.

Cuando solamente se utiliza realimentación del error, no es posible hacer inalcanzable el error del observador desde la señal de referencia. De este modo, la dinámica del observador aparecerá en la función de transferencia discreta entre la salida y la señal de referencia. El método de diseño utilizado en este capítulo permite evitar totalmente la excitación de los modos asociados con el observador introduciendo una función de transferencia discreta de prealimentación diferente. Entonces, el dipolo desaparecerá completamente en la función de transferencia discreta que relaciona la salida del proceso con la señal de referencia.



Predictor de Smith

El predictor de Smith se menciona en la Sección 7.5. Es una ley de control particular que ha sido propuesta para sistemas con retardo de tiempo. El predictor de Smith puede describirse de la siguiente forma.

Consideremos un proceso con la función de transferencia discreta

$$H(z) = \frac{1}{z^d} H'(z) = \frac{B'(z)}{z^d A'(z)} \quad (10.43)$$

donde $\text{gra } A' = \text{gra } B'$. El número d representa el retardo de tiempo en el sistema. En primer lugar, diseñaremos una realimentación

$$H_{fb}(z) = \frac{S'(z)}{R'(z)}$$

la cual define las propiedades deseadas en lazo cerrado si no existen retardos en el sistema. Para un sistema con retardo de tiempo, el predictor de Smith da la ley de control

$$u(k) = H_{fb}(u_c(k) - y(k)) - H' H_{fb}(1 - z^{-d}) u(k) \quad (10.44)$$

Entonces, la función de transferencia discreta en lazo cerrado es

$$\frac{B_m}{A_m} = \frac{B'S'}{z^d(A'R' + B'S')} \quad (10.45)$$

El predictor de Smith puede originarse fácilmente utilizando el diseño con asignación de polos. Sea (10.43) el proceso dado y (10.45) el modelo deseado. Consideremos, además, que el regulador solamente utiliza realimentación del error. Obsérvese que los ceros de proceso se conservan en el modelo. Con esta restricción, se deduce del Ejemplo 10.2 que la ley de control viene dada por

$$u(k) = \frac{S}{R}[u_c(k) - y(k)] = \frac{AB_m}{B(A_m - B_m)}[u_c(k) - y(k)]$$

El controlador es

$$\frac{S}{R} = \frac{z^d A' S'}{z^d(A'R' + B'S') - B'S'} = \frac{H_{fb}}{1 + H' H_{fb}(1 - z^{-d})}$$

O

$$u(k) = H_{fb}(u_c(k) - y(k)) - H' H_{fb}(1 - z^{-d}) u(k)$$

el cual es el predictor de Smith de (10.44). Es necesario cancelar todos los polos del proceso porque se utiliza solamente realimentación del error. Obsérvese que no es evidente de (10.44) que los polos del proceso estén cancelados. Desde luego es posible deducir controles similares al predictor de Smith utilizando la estructura más general

de realimentación-prealimentación de (10.4). Entonces, puede evitarse la cancelación de polos del proceso.

Seguimiento del modelo

La ley de control obtenida del diseño de asignación de polos puede interpretarse como un controlador de seguimiento del modelo. Para ver esto se utiliza la ecuación de diseño (10.20) con el fin de reescribir la ley de control. Se deduce de (10.21) que

$$\begin{aligned} \frac{T}{R} &= \frac{A_o B'_m}{B^+ R'} = \frac{A_o A_m B'_m}{A_m B^+ R'} = \frac{(AR' + B^- S) B'_m}{A_m B^+ R'} \\ &= \frac{AB'_m}{B^+ A_m} + \frac{B^- B'_m S}{A_m B^+ R'} = \frac{AB'_m B^-}{B^- B^+ A_m} + \frac{B^- B'_m S}{A_m B^+ R'} \\ &= \frac{AB_m}{BA_m} + \frac{B_m S}{A_m R} \end{aligned}$$

La ley de realimentación de (10.4) ahora es

$$\begin{aligned} u &= \frac{T}{R} u_c - \frac{S}{R} y = \frac{AB_m}{BA_m} u_c + \frac{SB_m}{RA_m} u_c - \frac{S}{R} y \\ &= \frac{AB_m}{BA_m} u_c - \frac{S}{R} (y - y_m) \end{aligned} \quad (10.46)$$

El primer término de la ley de control puede interpretarse como un término de prealimentación y el segundo como un término de realimentación. El término de prealimentación es una combinación de la respuesta deseada y de un modelo del proceso inverso. El término de realimentación tiene la función de transferencia discreta S/R . La señal de realimentación se obtiene de la diferencia entre la salida actual y la salida deseada y_m , donde

$$y_m = H_m u_c = \frac{B_m}{A_m} u_c$$

Una representación en diagrama de bloques de la ley de realimentación de (10.46) se presenta en la Figura 10.6. Obsérvese la similitud con la Figura 9.7.

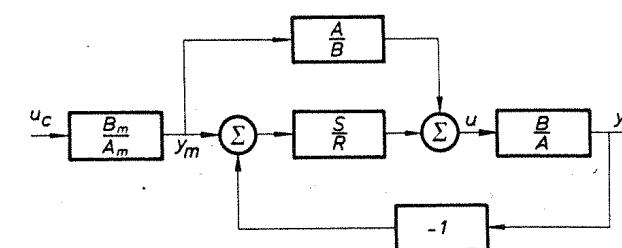


Figura 10.6 Diagrama de bloques de la ley de control (10.46).

Algoritmo de Dahlin-Higham

El método de Dahlin-Higham fue un método de diseño muy popular en los primeros diseños digitales de control de procesos porque los cálculos que requiere son muy simples. Se supone que la dinámica del proceso puede describirse por (10.43). En primer lugar, los polinomios A' y B' se cancelan con el regulador. Después de la cancelación, la dinámica del proceso es un retardo puro de tiempo

$$H''(z) = z^{-d}$$

Para reducir los errores en estado estacionario, es necesario disponer de una acción integral. Esto se asegura con $\ell=1$ en (10.31). El modelo deseado para el diseño es

$$H_m(z) = \frac{1-a}{z^{d-1}(z-a)}$$

El controlador se determina como en el Ejemplo 10.2, es decir, utilizando realimentación del error. El controlador obtenido también puede interpretarse como un controlador con un predictor de Smith. Puesto que el algoritmo está basado en la cancelación de todos los polos y ceros del proceso, no pueden permitirse polos o ceros fuera del círculo unidad. Habrá problemas de rizado debido a la cancelación de ceros estables pero poco amortiguados; compárese con la Figura 10.3.

Control algorítmico del modelo

El control algorítmico del modelo es un método particular de diseño que también se ha utilizado en el control de procesos. Este método puede interpretarse como un diseño por asignación de polos con

$$H(z) = z^{-d}B(z)$$

y

$$H_m(z) = \frac{1-a}{z^r(z-a)}$$

De este modo, el proceso se modela como respuesta pulsante finita.

10.8. Aspectos prácticos

El algoritmo de diseño por asignación de polos es un procedimiento directo de diseño si se conoce el modelo, H , la función de transferencia discreta deseada en lazo cerrado, H_m , y el polinomio del observador, A_0 . Sin embargo, no es precisamente fácil determinar los parámetros de diseño H_m y A_0 para que se cumplan las especificaciones físicas.

Como guía para el uso práctico del método de diseño por asignación de polos, estudiaremos la influencia de las especificaciones sobre la magnitud de la señal de

control, la sensibilidad frente a las perturbaciones en la carga, los errores de medida y los errores de modelado.

Rendimiento deseado para el servo

El rendimiento que se desea del servo se especifica para la función de transferencia discreta en lazo cerrado H_m . Se deduce del estudio de la Sección 10.3 que los ceros inestables o mal amortiguados del proceso deben corresponder a factores de B_m . Si esto no puede llevarse a cabo se obtendrán rizados como ilustra la Figura 10.3 y el algoritmo de Dahlin-Higham. Por tanto, es muy importante conocer los ceros inestables del proceso. El muestreo rápido de sistemas continuos con un exceso de polos superior a dos siempre produce ceros inestables (véase la Sección 3.6). Por tanto, es necesario conocer el retardo de tiempo y los ceros inestables del sistema que se va a controlar. Obsérvese que frecuentemente las dinámicas de orden elevado se pueden aproximar por retardos de tiempo.

La especificación de todos los polos y ceros para un sistema de orden elevado requiere muchos parámetros. En la práctica generalmente no tiene sentido definir tantos parámetros. Sería mucho más conveniente dar algunas características globales de la dinámica que se desea, tales como el ancho de banda y el pico de resonancia (véase Fig. 7.9). Para sistemas discretos, esto puede hacerse especificando los polos dominantes. Para sistemas oscilatorios, puede hacerse especificando un factor de seguro orden

$$P(z) = z^2 + p_1z + p_2 \quad (10.47)$$

dónde

$$\begin{aligned} p_1 &= -2e^{-\zeta\omega h} \cos(\omega h\sqrt{1-\zeta^2}) \\ p_2 &= e^{-2\zeta\omega h} \end{aligned}$$

También ζ corresponde al amortiguamiento relativo deseado y ω es la frecuencia natural.

Para sistemas sin rebalse, la dinámica dominante puede darse por un factor de primer orden

$$P(z) = z - a \quad (10.48)$$

dónde

$$a = \exp(-h/T)$$

y T es la constante de tiempo deseada.

Para sistemas discretos, el resto de los polos deseados en lazo cerrado se pueden situar en el origen. De este modo, la dinámica deseada en lazo cerrado tendrá la forma típica

$$H_m(z) = \frac{P(1)}{B^-(1)} \frac{B^-(z)}{z^m P(z)} \quad (10.49)$$

donde B^- corresponde a los ceros inestables o poco amortiguados y m es un número que considera los retardos de tiempo y las dinámicas de grado superior del proceso en lazo abierto.

Magnitud de las señales de control

Puede comprenderse intuitivamente que se necesitan grandes señales de control si se especifica un gran ancho de banda en lazo cerrado. Algunos análisis dan resultados cuantitativos. Anulando todas las perturbaciones, se deduce de (10.1) y (10.2) que

$$y = \frac{B_m}{A_m} u_c = H_m u_c$$

y

$$y = \frac{B}{A} u = Hu$$

Por tanto,

$$u = \frac{H_m}{H} u_c \quad (10.50)$$

Esta fórmula muestra con precisión las señales de control que se obtienen como respuesta a señales específicas de referencia. Obsérvese que la razón H_m/H también aparece en el análisis de robustez. Compárese este resultado con el Teorema 10.3.

Los órdenes de magnitud pueden estimarse fácilmente de los diagramas de Bode de las funciones de transferencia discreta H y H_m . El análisis de la Sección 4.4 indica el uso de los correspondientes diagramas continuos de Bode, incluyendo posiblemente el circuito mantenedor, lo cual es una buena aproximación, al menos, para altas velocidades de muestreo. Como muestra la Figura 10.5, la relación $|H_m/H|$ se obtiene fácilmente del diagrama. También puede estimarse de la figura el efecto de los cambios del ancho de banda deseado en el sistema en lazo cerrado sobre la magnitud de la señal de control. Si la pendiente de las curvas es -20 dlog/década para altas frecuencias, al duplicar el ancho de banda deseado la señal de control se cuadriplica. De este modo, la magnitud de la señal de control puede tener una gran sensibilidad ante cambios del ancho de banda.

Sensibilidad ante perturbaciones

La Figura 10.1 muestra un diagrama de bloques de un sistema, incluyendo perturbaciones y errores de medida. Con cálculos sencillos se obtiene

$$\begin{aligned} x &= \frac{BT}{AR + BS} u_c + \frac{BR}{AR + BS} v - \frac{BS}{AR + BS} e \\ &= \frac{B_m}{A_m} u_c + \frac{H_{lg}}{1 + H_{lg} H_{fb}} \frac{1}{v} - \frac{H_{lg}}{1 + H_{lg}} e \end{aligned} \quad (10.51)$$

donde x es la salida del proceso y H_{lg} es la ganancia del lazo definida por (10.11). Obsérvese la semejanza con la función de transferencia discreta que relaciona x con u_c , v y e . La única diferencia es la presencia de los polinomios R , S y T en los numeradores. Todos los polinomios R y S que satisfagan la ecuación de diseño (10.20) tendrán los mismos efectos en la transmisión de la señal desde u_c hasta x . La solución

arbitraria de (10.20) puede utilizarse para reducir la influencia de las perturbaciones en la carga y los errores de medida.

Se deduce de (10.51) que los errores de medida de baja frecuencia pueden propagarse a través del sistema con ganancia unidad, porque los errores pueden recogerse en la señal medida y . A través del mecanismo de realimentación, generarán variaciones en la salida del proceso. Sin embargo, los errores de medida de alta frecuencia no se propagan a través del sistema porque la dinámica del regulador los atenúa. También se deduce de (10.51) que la transmisión de la señal se puede bloquear a determinadas frecuencias exigiendo que $S(\exp i\omega h)$ sea pequeño para estas frecuencias. Por ejemplo, se obtiene un filtro de eliminación de banda exigiendo que

$$Q(z) = z^2 - 2ze^{-\alpha h} \cos \omega h + 1 \quad (10.52)$$

sea un factor de S .

También se deduce de (10.51) que las perturbaciones de baja frecuencia en la carga se atenuarán para las frecuencias en las que la ganancia de realimentación H_{fb} sea alta o el polinomio R pequeño. Un ejemplo típico consiste en eliminar los errores en estado estacionario debidos a perturbaciones constantes en la carga exigiendo que $R(1)=0$, lo cual es equivalente a una compensación integral. Para eliminar las perturbaciones en la carga de frecuencia ω puede ser necesario que $Q(z)$ —definida por (10.52)— sea un factor de R .

Cuando se requiere que R y S tengan factores específicos para llevar a cabo un filtro de eliminación de banda o control integral, el grado del polinomio observador A_0 debe incrementarse. La elección del polinomio observador también influirá en la transmisión de señal porque R y S están relacionados con A_0 a través de (10.20).

Influencia del polinomio del observador

El efecto del polinomio del observador en la transmisión de perturbaciones se ilustra con dos ejemplos.

Ejemplo 10.7

Consideremos un sistema con la función de transferencia

$$H(z) = \frac{0.1}{z - 1} \quad (10.53)$$

Supóngase que la función de transferencia discreta deseada entre la referencia y la salida es

$$H_m(z) = \frac{0.2}{z - 0.8}$$

Es fácil verificar que $A_0=1$, y utilizando la realimentación proporcional

$$u(k) = 2(u_c(k) - y(k))$$

se obtiene la función de transferencia en lazo cerrado deseada. Entonces se obtiene la siguiente expresión para la salida del proceso

$$x = \frac{0.2}{z - 0.8} u_e + \frac{0.1}{z - 0.8} v - \frac{0.2}{z - 0.8} e$$

Los diagramas de Bode para la transmisión de perturbaciones en la carga y de los errores de medida se muestran en la Figura 10.7.

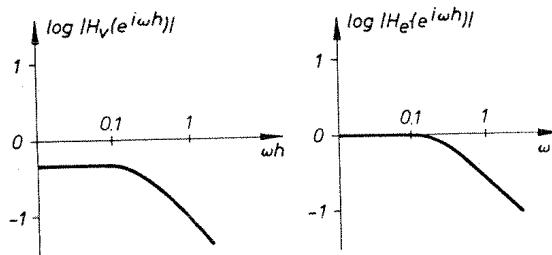


Figura 10.7 Diagramas de Bode de la transmisión de las perturbaciones de carga y del ruido de medida en el sistema del Ejemplo 10.7.

La Figura 10.7 muestra que la realimentación proporcional produce un sistema en lazo cerrado que es sensible a las perturbaciones en la carga. Se puede obtener un sistema menos sensible introduciendo un polinomio observador de mayor grado y restringiendo el polinomio R , como se muestra en el siguiente ejemplo.

Ejemplo 10.8

Consideremos el mismo sistema y la misma respuesta en lazo cerrado deseada que en el Ejemplo 10.7. Sea el polinomio del observador

$$A_o(z) = z - a$$

La ecuación de diseño (10.20) será

$$(z - 1)(z + r_1) + 0.1(s_0 z + s_1) = (z - a)(z - 0.8)$$

Por tanto, se obtienen las siguientes condiciones

$$\begin{aligned} -1 + r_1 + 0.1s_0 &= -a - 0.8 \\ -r_1 + 0.1s_1 &= 0.8a \end{aligned}$$

Como hay dos ecuaciones lineales y tres incógnitas, debe introducirse una condición extra. Elegimos $r_1 = -1$ para estar seguros que $R(1) = 0$ (acción integral). Por tanto,

$$s_0 = 12 - 10a$$

$$s_1 = 8a - 10$$

Se obtiene la siguiente expresión para la salida del proceso.

$$x = \frac{0.2}{z - 0.2} u_e + \frac{0.1(z - 1)}{(z - a)(z - 0.8)} v - \frac{(1.2 - a)z - 1 + 0.8a}{(z - a)(z - 0.8)} e$$

Los diagramas de Bode para la transmisión de señal procedente de perturbaciones en la carga y errores de medida se muestran en la Figura 10.8; compárense los resultados con la Figura 10.7.

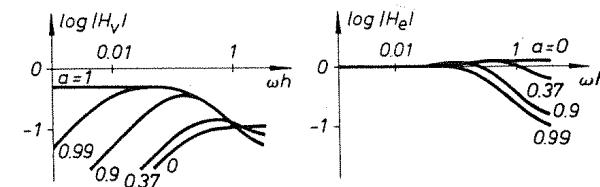


Figura 10.8 Diagrama de Bode de la transmisión de señal desde las perturbaciones de carga y desde el error de medida a la salida del proceso para diferentes polinomios de observador para el controlador del Ejemplo 10.8.

Selección del intervalo de muestreo

La elección del intervalo de muestreo se discutió en la Sección 9.2 para el diseño por asignación de polos basado en la realimentación de estado. El mismo argumento se puede utilizar para el método presentado en este capítulo. Esto significa que el intervalo de muestreo debería elegirse en relación con el comportamiento deseado en lazo cerrado. Una regla empírica consiste en tomar de 2 a 4 muestras por tiempo de subida del sistema en lazo cerrado, o de 8 a 16 muestras por período.

Ejemplo 10.9

Consideremos el mismo sistema y el mismo diseño que en el Ejemplo 10.6. Sean las especificaciones $\zeta = 0.7$ y $\omega = 1$. La regla empírica da un período de muestreo de $h \approx 0.5 - 1.0$. La Figura 10.9 muestra la salida del sistema para diferentes períodos de muestreo. El sistema es bastante insensible a la elección del período de muestreo.

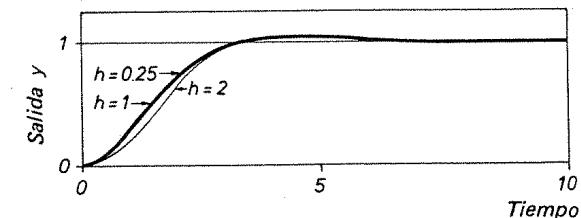


Figura 10.9 Influencia del intervalo de muestreo para el ejemplo del motor CC cuando $\zeta = 0.7$ y $\omega = 1$.

10.9. Un ejemplo de diseño

Un ejemplo de diseño más complejo ilustrará el uso del método de diseño polinómico.

Proceso

Consideremos un motor que mueve una carga formada por dos masas acopladas por un muelle de constante K (véase la Fig. 10.10). La señal de entrada es la corriente del motor I . Las velocidades angulares y los ángulos de las masas son ω_1 , ω_2 , φ_1 y φ_2 . Los momentos de inercia son J_1 y J_2 . Se supone que hay un amortiguamiento, d , en el muelle y que la primera de las masas puede ser perturbada por un par, v . Finalmente, la salida del proceso es la velocidad angular ω_2 .

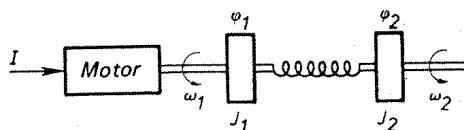


Figura 10.10 Proceso de dos masas acopladas con un muelle de amortiguamiento.

Se introducen los estados

$$x_1 = \varphi_1 - \varphi_2$$

$$x_2 = \omega_1/\omega_0$$

$$x_3 = \omega_2/\omega_0$$

donde

$$\omega_0 = \sqrt{k(J_1 + J_2)/(J_1 J_2)}$$

El proceso entonces se describe por

$$\frac{dx}{dt} = \omega_0 \begin{bmatrix} 0 & 1 & -1 \\ \alpha - 1 & -\beta_1 & \beta_1 \\ \alpha & \beta_2 & -\beta_2 \end{bmatrix} x + \begin{bmatrix} 0 \\ \gamma \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ \delta \\ 0 \end{bmatrix} v \quad (10.54)$$

$$y = [0 \ 0 \ \omega_0] x$$

donde

$$\alpha = J_1/(J_1 + J_2)$$

$$\beta_1 = d/(J_1 \omega_0)$$

$$\beta_2 = d/(J_2 \omega_0)$$

$$\gamma = k/J_1 \omega_0$$

$$\delta = 1/(J_2 \omega_0)$$

En el ejemplo se han utilizado los siguientes valores: $\omega_0 = 1$, $J_1 = 10/9$, $J_2 = 10$, $k = 1$, $d = 0,1$ y $k_f = 1$. Con estos valores el proceso (10.54) tiene tres polos $p_1 = 0$ y $p_{23} = -0,05 \pm 0,999i$ y un cero en $z_1 = -10$. El polo complejo tiene un amortiguamiento de $\zeta_p = 0,05$ y una frecuencia natural $\omega_p = 1$ rad/s. El diagrama de Bode del proceso se representa en la Figura 10.11 y la respuesta impulsional en la Figura 10.12.

Figura 10.11 Gráfico de Bode del proceso y el filtro antienmascaramiento. Sólo se presentan los módulos.

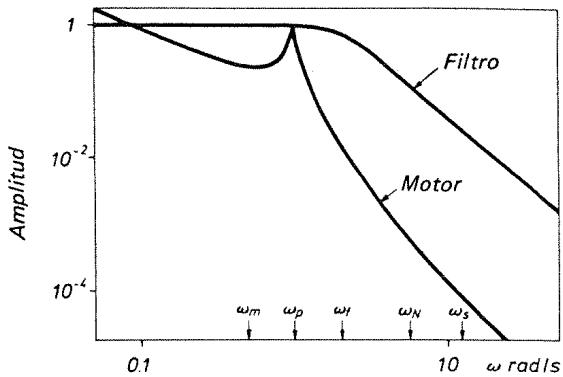
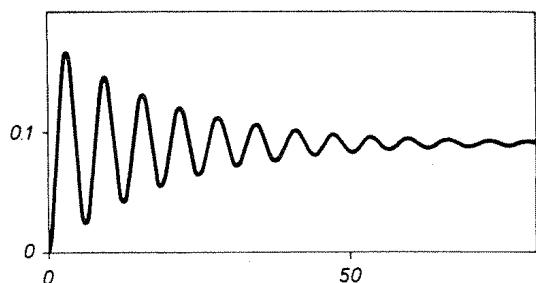


Figura 10.12 Respuesta impulsional del proceso.



Especificaciones

Se desea que el sistema en lazo cerrado tenga una respuesta a la señal de referencia tal que los modos dominantes tengan una frecuencia natural $\omega_m = 0,5$ rad/s y un amortiguamiento $\zeta_m = 0,7$.

Diseño

Elección del intervalo de muestreo. El modelo deseado tiene una frecuencia natural $\omega_m = 0,5$ rad/s. Utilizando la regla de la Sección 8.2 se obtiene $h = 0,5$ como una elección razonable de intervalo de muestreo. Esto produce una frecuencia de Nyquist $\omega_N = \pi/h \approx 6$ rad/s.

Se necesita un filtro para evitar problemas con las frecuencias enmascaradas. Se elige un filtro de segundo orden de la forma

$$\frac{\omega_f^2}{s^2 + 1.4\omega_f s + \omega_f^2}$$

con $\omega_f = 2$ rad/s. El filtro tiene una ganancia próxima a 0,1 para la frecuencia de Nyquist.

Muestreo del proceso. Las frecuencias ω_p y ω_f son de la misma magnitud. Entonces es necesario tener en cuenta el filtro antienmascaramiento de frecuencias para el diseño.

Muestreando el proceso y el filtro con $h=0,5$ se obtiene el modelo discreto

$$y(k) = \frac{B(q)}{A(q)} u(k)$$

con

$$A(q) = (q^2 - 0.7497q + 0.2432)(q^2 - 1.7125q + 0.9513)(q - 1)$$

$$B(q) = 0.1417(q + 12.1127)(q + 1.3395)(q + 0.2227)(q - 0.0024)$$

Los polos y ceros del sistema muestreado se representan en la Figura 10.13.

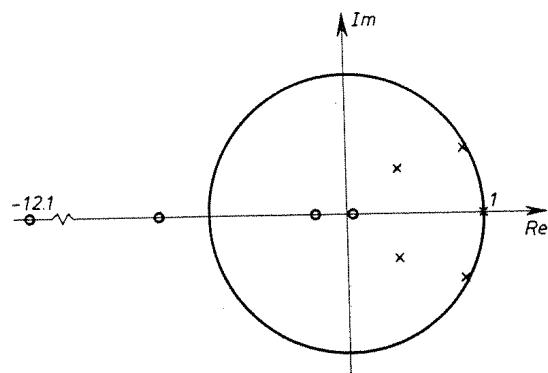


Figura 10.13 Diagrama de polos y ceros para el proceso y el filtro muestreados con $h=0.5$.

Diseño del filtro de eliminación de banda. La frecuencia asociada con la resonancia mecánica $\omega_p = 1$ está próxima a la frecuencia deseada en lazo cerrado $\omega_m = 0,5$. Por tanto, es necesario considerar la resonancia mecánica cuando se diseña el lazo de control. Un método clásico para realizar esto consiste en introducir una red de compensación que evite la excitación de los polos oscilatorios del proceso. El filtro que realiza esto se denomina *filtro de eliminación de banda* porque su diagrama de Bode tiene una hendidura para la frecuencia de los modos no deseados que los elimina. Este método asegura que los modos oscilatorios no son excitados por las señales de referencia o las acciones de control. Sin embargo, no introduce ningún amortiguamiento en los modos oscilatorios. Esto significa que el sistema responderá a la excitación de los modos oscilatorios del mismo modo que el sistema en lazo abierto. El filtro compensador de eliminación de banda puede obtenerse con el método del diseño polinómico.

Consideremos el polinomio característico del sistema en lazo cerrado. Desde luego es deseable que este polinomio tenga A_m como factor. También debe contener como factor el polinomio A_n , el cual corresponde a los modos oscilatorios, porque los modos oscilatorios se mantienen en el sistema en lazo cerrado y no están influenciados por la realimentación. El polinomio del observador A_o también debe ser un factor del polinomio característico en lazo cerrado. Además se supone que no hay cancelación de ceros del proceso. Entonces la identidad de polinomios (10.22) será

$$AR + BS = A_m A_n A_o \quad (10.55)$$

Como A_n divide a A pero no a B , se deduce que también divide a S . De este modo el regulador tendrá ceros en los polos osculatorios.

El polinomio A_m se elige a partir de las especificaciones. Se eligen dos polos para proporcionar el amortiguamiento y la frecuencia natural deseados. Los tres polos restantes se sitúan en el origen. Se supone que el observador tiene todos los polos en el origen.

Del Teorema 10.2 se deduce que el diseño requiere $\text{gra } R = 4$, $\text{gra } S = 4$ y $\text{gra } A_0 = 2$. La identidad (10.22) es de noveno orden y da los polinomios R y S del regulador. La función de transferencia en lazo cerrado es

$$\frac{BT}{AR + BS} = t_0 \frac{B}{A_m}$$

donde

$$t_0 = \frac{A_m(1)}{B(1)}$$

Por tanto, el polinomio T es

$$T = t_0 A_n A_{n_0}$$

Con esta elección de T la ganancia en estado estacionario del sistema en lazo cerrado es la unidad. Obsérvese que S y T tienen como factor A_n . La respuesta del sistema en lazo cerrado cuando se utiliza el regulador de eliminación de banda se muestra en la Figura 10.14. La señal de referencia es un salto escalón en $t=0$ y la perturbación v es un pulso en $t=25$ de magnitud 5 y duración 0,1 s. La respuesta del sistema está de acuerdo con las especificaciones.

No hay excitación de los modos débilmente amortiguados por la señal de referencia o la señal de control. Sin embargo, el pulso de perturbación excita estos modos y causa la oscilación de la respuesta. Obsérvese que la oscilación no introduce ninguna acción de control.

Amortiguamiento activo de los modos oscilatorios. Con el diseño del filtro de eliminación de banda el regulador no intenta amortiguar los modos oscilatorios. Efectuaremos un nuevo diseño para que las características del servo sean las mismas pero también se amortiguen las oscilaciones. Supongamos que el amortiguamiento de los modos oscilatorios debe cambiarse desde el amortiguamiento en lazo abierto ζ ,

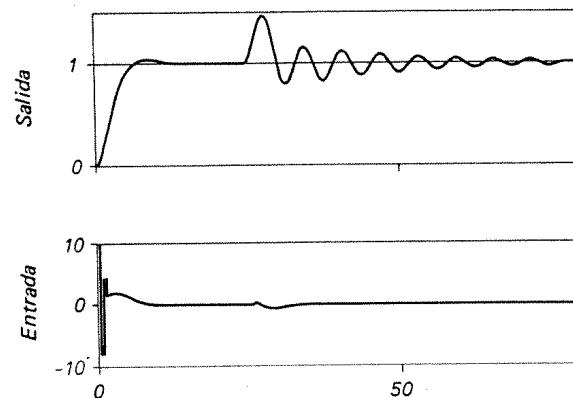


Figura 10.14 Respuesta del sistema en lazo cerrado empleando el regulador de diseño de eliminación de banda.

$=0,05$ hasta $0,707$. Además supongamos que la frecuencia de amortiguamiento debe ser la misma de antes. Esto corresponde a los polos del sistema continuo

$$p_{12} = -0.707 \pm 0.999i$$

Designemos por A_d al correspondiente polinomio del sistema discreto. El diseño se realiza del modo anterior pero con la identidad

$$AR + BS = A_m A_d A_o \quad (10.56)$$

donde A_m y A_o son los mismos de antes. La respuesta del sistema en lazo cerrado se muestra en la Figura 10.15. Compárese con la Figura 10.14. Las características del

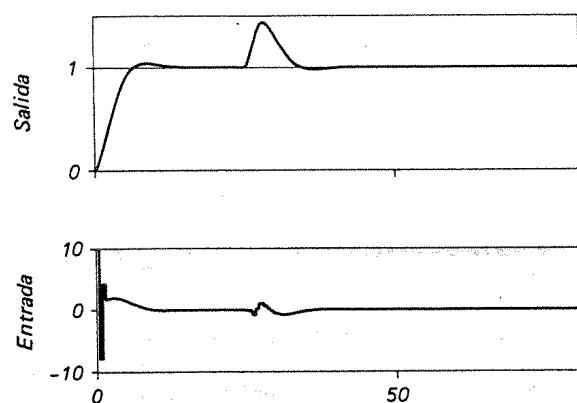


Figura 10.15 Respuesta del sistema en lazo cerrado utilizando el regulador diseñado para amortiguamiento activo.

servo son las mismas que en el caso anterior y los modos oscilatorios están ahora amortiguados por el regulador.

Comparación

El diseño del filtro de eliminación de banda y el diseño del amortiguamiento activo son dos modos de solucionar el problema del diseño utilizando el método polinómico. Los ejemplos de diseño pueden elaborarse además introduciendo un integrador y estudiando el efecto de diferentes polinomios para el observador.

La ecuación (10.51) puede utilizarse para ver la diferencia entre los dos diseños. La relación entre la variable x de la Figura 10.1, la señal de referencia y las perturbaciones está dada por

$$x(k) = \frac{BT}{AR + BS} u_c(k) + \frac{BR}{AR + BS} v(k) - \frac{BS}{AR + BS} e(k)$$

Esto demuestra que el diseño del filtro de eliminación de banda no amortigua los modos oscilatorios. El polinomio A_n es un factor del denominador de la función de transferencia de la perturbación v pero no de e . En el diseño con amortiguamiento activo, A_d es un factor del denominador y las oscilaciones están mejor amortiguadas. El amortiguamiento de las perturbaciones también puede verse estudiando la diferencia de retorno que viene dada por

$$1 + \frac{BS}{AR} = \frac{AR + BS}{AR}$$

De (10.55) y (10.56) se deduce que la diferencia de retorno es

$$\frac{A_m(z)A_n(z)A_o(z)}{A(z)R(z)}$$

para el diseño de filtro de banda eliminada y

$$\frac{A_m(z)A_d(z)A_o(z)}{A(z)R(z)}$$

para el diseño de amortiguamiento activo. La magnitud de la diferencia de retorno para $|z|=1$ se muestra en la Figura 10.16. La diferencia de retorno indica la atenuación de perturbaciones como el ruido de medida a la salida del proceso. Compárese con (10.51). Se ve que el diseño del amortiguamiento activo produce una mejor amortiguación de los modos oscilatorios..

El Teorema 10.3 puede utilizarse para evaluar la precisión necesaria del modelo para el diseño. La precisión relativa está dada por

$$\frac{1}{|H_m|} \left| \frac{H_{ff}}{H_{fb}} \right| = \left| \frac{A_m(z)T(z)}{t_0 B_m(z)S(z)} \right|$$

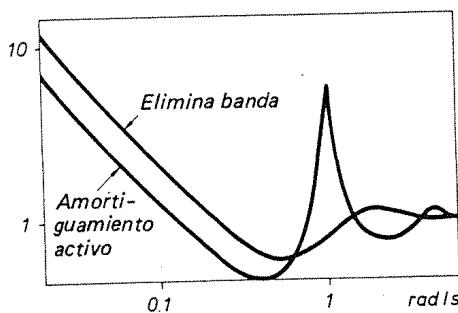


Figura 10.16 Módulo de las diferencias de retorno.

Esta función está dibujada en la Figura 10.17 para $|z|=1$. La figura muestra que el diseño de amortiguamiento activo es más insensible que el diseño del filtro de eliminación de banda a los errores de modelado. El diseño del filtro de eliminación de banda es bastante sensible alrededor de ω_p .

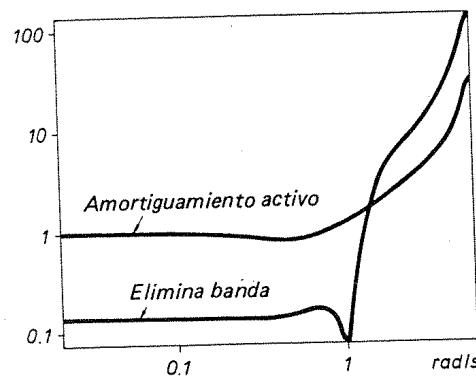


Figura 10.17 Diagrama de Bode de $A_m T / (t_0 B_m S)$ que da la precisión relativa del modelo necesaria para un diseño estable.

Resumen

El ejemplo muestra que el método polinómico de diseño es un modo flexible y útil para diseñar controladores. Sin embargo, deben tomarse precauciones para evitar problemas numéricos si el proceso es de orden elevado (cuarto orden o superior). Pueden presentarse problemas numéricos en la solución de la identidad polinómica y en la realización del regulador.

10.10. Conclusiones

Este capítulo trata un procedimiento de asignación de polos basado en manipulaciones polinómicas. El diseño comienza con un modelo del proceso, H , un

modelo deseado, H_m , y un polinomio A_o del observador. El Algoritmo 10.1 conduce a un regulador que incluye un término de realimentación desde la salida medida y un término de prealimentación desde la señal de referencia.

La elección del modelo deseado y del polinomio del observador están íntimamente relacionados con las especificaciones físicas.

10.11. Problemas

- 10.1. Utilizar el algoritmo de Euclides del Apéndice D para calcular el máximo común divisor de los polinomios

$$B(z) = z^3 - 2z^2 + 1.45z - 0.35$$

$$A(z) = z^4 - 2.6z^3 + 2.25z^2 - 0.8z + 0.1$$

- 10.2. Dado el operador de transferencia discreta

$$H(q) = \frac{1}{q + a}$$

y siendo

$$H_m(q) = \frac{1 + \alpha}{q + \alpha}$$

el sistema deseado

- a) Determinar un controlador de la forma de (10.4) utilizando el Algoritmo 10.1.
b) Determinar el polinomio característico del sistema en lazo cerrado.

- 10.3. Considerar el sistema dado por la función de transferencia

$$H(q) = \frac{z + 0.7}{z^2 - 1.8z + 0.81}$$

Utilizar el diseño polinómico para determinar un controlador tal que el sistema en lazo cerrado tenga el polinomio característico

$$z^2 - 1.5z + 0.7$$

Supongamos que el polinomio del observador tiene el orden más bajo posible y todos sus polos situados en el origen. Considerar los siguientes casos:

El cero del proceso está cancelado.

El cero del proceso no está cancelado.

Simular los dos casos y discutir las diferencias entre los dos controladores. ¿Cuál de ellos es preferible?

- 10.4. Para el sistema del Problema 10.2, suponer que la realimentación sólo puede realizarse mediante el error. Por tanto, el controlador tiene la forma

$$u(k) = \frac{S}{R}[u_c(k) - y(k)]$$

- a) Determinar S/R para obtener el sistema en lazo cerrado deseado.
b) Determinar la ecuación característica del sistema en lazo cerrado y comparar con el Problema 10.2. Considerar, por ejemplo, el caso para $|a| > 1$.



- 10.5. Considerar el sistema del Problema 10.2 y suponer que el sistema en lazo cerrado debería ser capaz de eliminar perturbaciones en escalón en la entrada del proceso. Esto significa que en la Figura 10.1 v es un escalón.
- Analizar lo que ocurre cuando se utiliza el controlador del Problema 10.2 y v es un escalón.
 - Rediseñar el controlador para que se cumplan las especificaciones.
- 10.6. Demostrar que (10.41) es correcta.
- 10.7. Considerar el sistema del Problema 10.2 y suponer que $a = -0,9$ y $\alpha = -0,5$.
- Emplear cálculo directo para determinar la influencia de los errores de modelado. Suponer que el diseño se hace para $a = -0,9$ y determinar la estabilidad del sistema en lazo cerrado si el proceso verdadero tiene un polo en a^0 .
 - Utilizar el Teorema 10.3 para determinar la influencia de los errores de modelado. ¿Qué ocurre cuando decrece α ?
- 10.8. Considerar el sistema del Problema 10.2. Utilizar (10.50) para determinar el valor máximo de la señal de control como función de a y α cuando la señal de referencia es un escalón.
- 10.9. En el Ejemplo 10.6 se hace un diseño polinómico para el motor normalizado. Simular el sistema y estudiar la sensibilidad del método de diseño respecto de la elección del intervalo de muestreo. Suponer que las especificaciones en lazo cerrado corresponden a un sistema continuo de segundo orden con amortiguamiento $\zeta = 0,7$ y frecuencia natural $\omega = 1$ rad/s. Comparar con la Figura 10.9.
- 10.10. Considerar el sistema descrito por

$$A_1(q)z(k) = B_1(q)u(k)$$

$$A_2(q)y(k) = B_2(q)z(k)$$

Suponer que la variable que ha de ser controlada es $z(k)$, pero la variable medida es $y(k)$. Además suponer que A_2 tiene sus raíces dentro del círculo unidad.

Obtener un controlador de la forma (10.4) tal que el sistema en lazo cerrado sea

$$A_m(q)z(k) = B_m(q)u_c(k)$$

¿Qué restricciones deben imponerse? ¿Cómo influyen en la función de transferencia discreta del sistema en lazo cerrado las incertidumbres en A_2 y B_2 ?

- 10.11. Considerar el sistema de dos tanques del Problema 3.10 para $h = 12$ s.
- Utilizar el método polinómico para diseñar un controlador con un integrador. Suponer que la ecuación característica en lazo cerrado deseada es

$$z^2 - 1.55z + 0.64 = 0$$

Esto corresponde a $\zeta = 0,7$ y $\omega = 0,027$ rad/s.

- Rediseñar el controlador para diferentes valores de ω y estudiar cómo varía la magnitud de la señal de control con ω .
- 1.12. Considerar el control del motor normalizado del Ejemplo A.2. Demostrar que la realimentación de velocidad puede diseñarse utilizando asignación de polos. (*Sugerencia:* Diseñar en primer lugar una ley de realimentación únicamente con realimentación de posición. Entonces demostrar que la ley de control puede reescribirse como una combinación de realimentaciones de posición y velocidad.)
- 10.13. Generalizar el resultado del Problema 10.12 a un proceso general con varias salidas.

10.12. Referencias

El método polinómico para asignación de polos se trata en

- KUČERA, V. (1979): *Discrete Linear Control*. Prague: Academia.
WOLOWICH, W. A. (1974): *Linear Multivariable Systems*. New York: Springer-Verlag.
PERNEBO, L. (1981): «An Algebraic Theory for the Design of Controllers for Multivariable Systems-Part I: Structure Matrices and Feedforward Design and Part II: Feedback realizations and feedback design. *IEEE Trans. Autom. Control*, AC-26, 171-82 and 183-94.

El método estudiado en este capítulo se ha utilizado en conexión con algoritmos adaptativos de asignación de polos en

- ÅSTRÖM, K. J. y B. WITTENMARK (1980): «Self-tuning Controllers Based on Pole-Zero Placement.» *Proc. IEE*, Part D, 127, 120-30.

El algoritmo de Dahlin-Higham se dedujo independientemente en

- DAHLIN, E. B. (1968): «Designing and Tuning Digital Controllers.» *Instruments & Control Systems*, 41, No. 6, 77-83.
HIGHAM, J. D. (1968): «Single-Term Control of First- and Second-Order Processes with Dead Time.» *Control*, February, 136-40.

El controlador de modelo algorítmico se describió en

- RICHALET, J. A., RAULT, J. L. TESTUD, y J. PAPON (1978): «Model Predictive Heuristic Control: Applications to Industrial Processes.» *Automatica*, 14, 413-28.

Métodos de diseño óptimo: formulación en el espacio de estados

Objetivo:

Desarrollar técnicas de diseño óptimo y resolver los problemas de predicción y filtrado para sistemas descritos por modelos lineales en el espacio de estados y con criterios cuadráticos.

11.1. Introducción

En los dos capítulos anteriores se ha resuelto el problema de síntesis empleando técnicas de asignación de polos. Los principales parámetros de diseño han sido la localización de los polos en lazo cerrado, y las presentaciones se han limitado a los sistemas de una sola entrada y una sola salida. En este capítulo se expondrá un problema más general. Todavía se supone que el proceso es lineal, pero puede ser variable con el tiempo y tener varias entradas y salidas. Además se introducen en los modelos ruidos de proceso y de medida. El problema de síntesis se formula como la minimización de un criterio, que es una función cuadrática de los estados y de las señales de control. El controlador óptimo resultante es lineal. El problema, que se establecerá formalmente más adelante, recibe el nombre de *problema de control Lineal Cuadrático (LQ)*, del inglés *Linear Quadratic*, o de *problema de control Lineal Cuadrático Gaussiano (LQG)* si en los modelos del proceso se permiten perturbaciones estocásticas gaussianas. La solución estacionaria al problema LQ para sistemas invariantes conduce a una ley de control de la misma estructura que el controlador de realimentación de estado del Capítulo 9. El controlador LQ puede entonces interpretarse como un controlador de asignación de polos. Los grados de libertad de la versión multivariable del controlador del Capítulo 9 se resuelven por la minimización de una función de coste en lugar de especificar sólo los polos en lazo cerrado y los vectores propios en lazo cerrado.

El control LQ es un área muy amplia tratada en muchos libros. En este capítulo sólo se dará una breve revisión de las principales ideas y resultados. En esta sección se

formula el problema y se dan algunos resultados útiles. En la Sección 11.2 se presenta la solución del problema LQ en el caso de que se disponga de todos los estados, y también se exponen las propiedades de los controladores LQ. Si todos los estados no son medibles, hay que estimarlos, lo que puede hacerse empleando un sistema dinámico como en la Sección 9.3. Para el caso de perturbaciones gaussianas, se puede determinar el estimador óptimo, que minimiza la varianza del error del estimador. Esto es lo que se llama el filtro de Kalman. El estimador tiene la misma estructura que en (9.24). Sin embargo, la matriz de ganancia, K , se determina de forma diferente y en general es variable con el tiempo. Los filtros de Kalman se exponen en la Sección 11.3. En la Sección 11.4 se resuelve el problema LQG, en el que los estados se estiman usando un filtro de Kalman. La solución se basa en el *teorema de separación* o en el principio de *equivalencia de certeza*. Esto implica que la estrategia de control óptimo se puede descomponer en dos partes: un estimador de estado, que da la mejor estimación de los estados a partir de las salidas observadas, y una ley de realimentación lineal de los estados estimados. El controlador lineal empleado es el mismo que el que se emplearía si no hubiera perturbaciones actuando sobre el sistema. En la Sección 11.5 se presentan algunos aspectos prácticos.

Formulación del problema

El problema de diseño se especifica dando el proceso, el criterio y las señales de control admisibles.

El proceso. Se supone que el proceso a controlar está descrito por el modelo continuo

$$dx = Ax dt + Bu dt + dv_c \quad (11.1)$$

donde A y B pueden ser matrices variables con el tiempo. El proceso v_c tiene valor medio cero e incrementos no correlacionados. La covarianza incremental de v_c es $R_{1c} dt$ (compárese con la Sección 6.6). El modelo (11.1) puede muestrearse como en la Sección 6.7. Es preciso hacer algunas modificaciones porque se permite que el sistema sea variable con el tiempo. La entrada $u(t)$ es constante durante el período de muestreo, y para el caso en que no haya ruido la solución de (11.1) puede escribirse como

$$x(t) = \Phi(t, kh)x(kh) + \Gamma(t, kh)u(kh) \quad (11.2)$$

donde $\Phi(t, kh)$ es la matriz fundamental de (11.1) que satisface

$$\frac{d}{dt} \Phi(t, kh) = A(t)\Phi(t, kh); \quad \Phi(kh, kh) = I \quad (11.3)$$

y

$$\Gamma(t, kh) = \int_{kh}^t \Phi(t, s)B(s) ds \quad (11.4)$$



Omitiendo los argumentos temporales de las matrices, el modelo muestreado puede escribirse como

$$\begin{aligned} x(kh + h) &= \Phi x(kh) + \Gamma u(kh) + v(kh) \\ y(kh) &= Cx(kh) + e(kh) \end{aligned} \quad (11.5)$$

donde v y e son procesos de ruido blanco discretos gaussianos con media cero y

$$Ev(kh)v^T(kh) = R_1$$

$$Ev(kh)e^T(kh) = R_{12}$$

$$Ee(kh)e^T(kh) = R_2$$

Las expresiones para las matrices de covarianza se dan en la Sección 6.7. Además se supone que el estado inicial $x(0)$ tiene una distribución gaussiana con

$$Ex(0) = m_0 \quad \text{y} \quad \text{cov}(x(0)) = R_0$$

Las matrices R_0 , R_1 y R_2 son semidefinidas positivas. Las matrices de covarianza pueden ser variables. Se supone que el modelo (11.5) es alcanzable y observable.

Como se expuso en el Capítulo 9, es posible incluir diferentes tipos de perturbaciones y efectos del entorno aumentando el vector de estado del proceso.

El criterio. El propósito del control es minimizar la función de coste

$$J = E \left\{ \int_0^{kh} [x^T(t)Q_{1c}x(t) + 2x^T(t)Q_{12c}u(t) + u^T(t)Q_{2c}u(t)] dt + x^T(Nh)Q_{0c}x(Nh) \right\} \quad (11.6)$$

donde las matrices Q_{0c} , Q_{1c} y Q_{2c} son simétricas y definidas positivas. Las matrices en la función de coste pueden depender del tiempo.

Leyes de control admisibles. Es importante especificar los datos disponibles para determinar la señal de control. La primera suposición es que se emplea un muestreo periódico y que la señal de control es constante durante los períodos de muestreo.

Si C es igual a la matriz unidad y si en (11.5) $e(kh)=0$, entonces se puede disponer de la totalidad del vector de estado. En este caso se admite que la señal de control sea una función de los estados hasta el instante kh inclusive. Esto se conoce como *información completa del estado*. En la mayoría de los casos no se conocen exactamente las variables de estado. Es lo que se denomina *información incompleta del estado*. En este caso la señal de control en el instante kh puede ser una función de las entradas y salidas hasta el instante $kh-h$.

El problema. El problema de control óptimo se define ahora del siguiente modo: encontrar la señal admisible de control que minimiza la función de coste (11.6) cuando el proceso está descrito por el modelo (11.1) o el modelo equivalente (11.5). Los parámetros de diseño son las matrices en la función de coste y el período de muestreo.

Muestreo de la función de coste

La función de coste dada en (11.6) está expresada en forma continua. Primero se transforma en una función de coste discreta. Integrando (11.6) sobre intervalos de longitud h se tiene

$$J = E \left\{ \sum_{k=0}^{N-1} J(k) + x^T(Nh)Q_{0c}x(Nh) \right\} \quad (11.7)$$

donde

$$J(k) = \int_{kh}^{kh+h} [x^T(t)Q_{1c}x(t) + 2x^T(t)Q_{12c}u(t) + u^T(t)Q_{2c}u(t)] dt \quad (11.8)$$

Usando (11.2) en (11.8) y por el hecho de que $u(t)$ es constante sobre el período de muestreo se tiene

$$J(k) = x^T(kh)Q_{1c}x(kh) + 2x^T(kh)Q_{12c}u(kh) + u^T(kh)Q_{2c}u(kh)$$

donde

$$Q_1 = \int_{kh}^{kh+h} \Phi^T(s, kh)Q_{1c}\Phi(s, kh) ds \quad (11.9)$$

$$Q_{12} = \int_{kh}^{kh+h} \Phi^T(s, kh)[Q_{1c}\Gamma(s, kh) + Q_{12c}] ds \quad (11.10)$$

$$Q_2 = \int_{kh}^{kh+h} [\Gamma^T(s, kh)Q_{1c}\Gamma(s, kh) + 2\Gamma^T(s, kh)Q_{12c} + Q_{2c}] ds \quad (11.11)$$

Entonces minimizar la función de coste (11.6) cuando $u(kh)$ es constante sobre el período de muestreo es lo mismo que minimizar la función de coste discreta

$$\begin{aligned} J = E \left\{ \sum_{k=0}^{N-1} &[x^T(kh)Q_{1c}x(kh) + 2x^T(kh)Q_{12c}u(kh) + u^T(kh)Q_{2c}u(kh)] \right. \\ &\left. + x^T(Nh)Q_{0c}x(Nh) \right\} \quad (11.12) \end{aligned}$$

Las matrices Q_1 , Q_{12} y Q_2 están dadas por (11.9)-(11.11), respectivamente, y $Q_0 = Q_{0c}$. En lo que sigue se supone que Q_1 es semidefinida positiva y que Q_2 es definida positiva. Además, se supone que $Q_{12}=0$. A continuación se muestra cómo se puede eliminar Q_{12} .

Cuando se considera el caso estocástico, en (11.12) se tiene un término adicional dependiente del ruido. Sin embargo, este término es independiente de la señal de control y puede no tenerse en cuenta cuando se realiza la minimización.

El problema del control óptimo se ha transformado ahora en el problema discreto de minimizar la función de coste (11.12) cuando el proceso está descrito por la ecuación (11.5). Para facilitar la escritura en lo que sigue se supone que se utiliza como unidad de tiempo el período de muestreo; es decir, $h=1$.

Transformación de la función de coste

Para simplificar la escritura se realiza una transformación. Se introduce una nueva señal de control

$$\tilde{u} = u + M^T x$$

donde

$$M = Q_{12}Q_2^{-1}$$

El sistema de (11.5) se transformará en

$$\begin{aligned} x(k+1) &= \tilde{\Phi}x(k) + \Gamma\tilde{u}(k) + v(k) \\ y(k) &= Cx(k) + e(k) \end{aligned}$$

donde

$$\tilde{\Phi} = \Phi - \Gamma M^T$$

Además, la función de coste (11.12) se transforma en

$$J = E \left\{ \sum_{k=0}^{N-1} [x^T(k)\tilde{Q}_1x(k) + \tilde{u}^T(k)Q_2\tilde{u}(k)] + x^T(N)Q_0x(N) \right\} \quad (11.13)$$

donde

$$\tilde{Q}_1 = Q_1 - Q_{12}Q_2^{-1}Q_{12}^T$$

Para el sistema transformado se elimina el producto cruzado entre estados y control. Con esto puede estudiarse el caso en que $Q_{12}=0$.

Valor medio de una forma cuadrática

En lo que sigue se evaluarán expresiones de la forma

$$Ex^T S x$$

donde x es una variable aleatoria gaussiana con media m y matriz de covarianza R . Tenemos

$$\begin{aligned} Ex^T S x &= E(x-m)^T S(x-m) + Em^T S x + Ex^T Sm - Em^T Sm \\ &= E(x-m)^T S(x-m) + m^T Sm \end{aligned}$$

Además,

$$\begin{aligned} E(x-m)^T S(x-m) &= Etr(x-m)^T S(x-m) = EtrS(x-m)(x-m)^T \\ &= trSE(x-m)(x-m)^T = trSR \end{aligned}$$

Entonces

$$Ex^T S x = m^T Sm + trSR \quad (11.14)$$

Formación de expresiones cuadráticas

En lo que sigue habrá que minimizar varias veces funciones cuadráticas. Una forma de encontrar el mínimo es emplear el método de completar las expresiones de los cuadrados. Considérese la función

$$F(u) = u^T Su + r^T u + u^T r$$

donde S es una matriz positiva, definida y simétrica de orden $n \times n$ y u y r son vectores de dimensión n . El mínimo de $F(u)$ puede obtenerse reescribiendo la función en la forma

$$\begin{aligned} F(u) &= u^T Su + r^T u + u^T r \\ &= u^T Su + r^T u + u^T r + r^T S^{-1}r - r^T S^{-1}r \\ &= (u + S^{-1}r)^T S(u + S^{-1}r) - r^T S^{-1}r \end{aligned}$$

El primer término es siempre no negativo; entonces el mínimo se obtiene para

$$u = -S^{-1}r \quad (11.15)$$

y el mínimo es

$$F_{\min} = -r^T S^{-1}r \quad (11.16)$$

11.2. Control lineal cuadrático

Ahora se va a resolver el problema del control LQ para el caso de información completa del estado. La solución se obtiene por medio de la programación dinámica.

El caso determinístico

En primer lugar se considera el caso determinístico, donde $v(k)=0$ y $e(k)=0$ en (11.5). Entonces el sistema se describe por

$$x(k+1) = \Phi x(k) + \Gamma u(k) \quad (11.17)$$

donde se da $x(0)$. El problema consiste ahora en determinar la secuencia $u(0), u(1), \dots, u(N-1)$ tal que minimice la función de coste (11.12). La solución se expone en el siguiente teorema.

Teorema 11.1. Considérese el sistema (11.17). Sea $u(k)$ una función de $x(k), x(k-1), \dots$. Introdúzcase

$$\begin{aligned} S(k) &= \Phi^T S(k+1)\Phi + Q_1 - L^T(k)(Q_2 + \Gamma^T S(k+1)\Gamma)L(k) \\ &= [\Phi - \Gamma L]^T S(k+1)\Phi + Q_1 \\ &= [\Phi - \Gamma L]^T S(k+1)[\Phi - \Gamma L] + Q_1 + L^T Q_2 L \end{aligned} \quad (11.18)$$

donde la matriz L está definida por

$$L(k) = (Q_2 + \Gamma^T S(k+1)\Gamma)^{-1}\Gamma^T S(k+1)\Phi \quad (11.19)$$

y con la condición final $S(N) = Q_0$. Supóngase que $S(k)$ tiene una solución que es semidefinida positiva y que $Q_2 + \Gamma^T S(k)\Gamma$ es definida positiva. Entonces existe una única estrategia de control admisible

$$u(k) = -L(k)x(k) \quad (11.20)$$

que minimiza el coste (11.12) cuando $Q_{12} = 0$. El valor mínimo del coste será

$$\min J = V_0 = x^T(0)S(0)x(0) \quad (11.21)$$

Democión. Para demostrar el teorema se emplea la programación dinámica. Introdúzcase

$$V_k = \min_{u(k), \dots, u(N-1)} E \left\{ \sum_{i=k}^{N-1} [x^T(i)Q_1x(i) + u^T(i)Q_2u(i)] + x^T(N)Q_0x(N) \right\}$$

Para $k=N$ se define

$$V_N = x^T(N)S(N)x(N)$$

donde

$$S(N) = Q_0$$

Para $k=N-1$,

$$V_{N-1} = \min_{u(N-1)} \{x^T(N-1)Q_1x(N-1) + u^T(N-1)Q_2u(N-1) + V_N\} \quad (11.22)$$

Utilizando (11.5) se tiene

$$\begin{aligned} V_{N-1} &= \min_{u(N-1)} \{x^T(N-1)Q_1x(N-1) + u^T(N-1)Q_2u(N-1) \\ &\quad + [\Phi x(N-1) + \Gamma u(N-1)]^T S(N)[\Phi x(N-1) + \Gamma u(N-1)]\} \\ &= \min_{u(N-1)} \{x^T(N-1)[Q_1 + \Phi^T S(N)\Phi]x(N-1) \\ &\quad + x^T(N-1)\Phi^T S(N)\Gamma u(N-1) + u^T(N-1)\Gamma^T S(N)\Phi x(N-1) \\ &\quad + u^T(N-1)[\Gamma^T S(N)\Gamma + Q_2]u(N-1)\} \end{aligned}$$

Con (11.15) y (11.16) la ley de control

$$u(N-1) = -L(N-1)x(N-1)$$

da el coste mínimo

$$V_{N-1} = x^T(N-1)S(N-1)x(N-1)$$

donde

$$S(N-1) = \Phi^T S(N)\Phi + Q_1 - L^T(N-1)(Q_2 + \Gamma^T S(N)\Gamma)L(N-1)$$

y

$$L(N-1) = (Q_2 + \Gamma^T S(N)\Gamma)^{-1}\Gamma^T S(N)\Phi$$

Los mismos argumentos dan

$$V_{N-2} = \min_{u(N-2)} \{x^T(N-2)Q_1x(N-2) + u^T(N-2)Q_2u(N-2) + V_{N-1}\}$$

Esta es igual que (11.22), pero con los argumentos temporales desplazados un paso. Ahora se puede repetir el proceso y V_0 , que es el mínimo de J , se obtiene iterando hacia atrás en el tiempo. \square

Nota 1. Nótese que no es necesario que Q_2 sea definida positiva; basta que $Q_2 + \Gamma^T S(k)\Gamma$ sea definida positiva.

Nota 2. Si hay un término de producto cruzado en la función de coste, como en (11.12), la matriz de ganancia se transforma en

$$L(k) = (Q_2 + \Gamma^T S(k+1)\Gamma)^{-1}(\Gamma^T S(k+1)\Phi + Q_{12}^T)$$

y la ecuación (11.18) se transforma en

$$\begin{aligned} S(k) &= \Phi^T S(k+1)\Phi + Q_1 - L^T(Q_2 + \Gamma^T S\Gamma)L \\ &= (\Phi - \Gamma L)^T S(k+1)\Phi + Q_1 - L^T Q_{12}^T \\ &= (\Phi - \Gamma L)^T S(k+1)(\Phi - \Gamma L) + Q_1 - L^T Q_{12}^T - Q_{12}L + L^T Q_2 L \end{aligned}$$

Nota 3. Los cálculos necesarios para determinar el controlador LQ sólo se pueden hacer a mano para ejemplos muy sencillos. En la práctica es necesario tener acceso a programas interactivos, que puedan calcular la ley de control y simular los sistemas.

La ecuación de Riccati

La ecuación (11.18) se conoce como la ecuación discreta de Riccati. La ecuación se puede escribir como

$$S(k) = [\Phi - \Gamma L(k)]^T S(k+1)[\Phi - \Gamma L(k)] + Q_1 + L^T(k)Q_2L(k) \quad (11.23)$$

Se supone que las matrices $S(N) = Q_0$ y Q_1 son simétricas y semidefinidas positivas. De esto se sigue que $S(k)$ es simétrica y semidefinida positiva.

Se puede utilizar la ecuación de Riccati para reescribir la función de coste (11.12) cuando $Q_{12} = 0$, lo que se presenta en el siguiente teorema.

Teorema 11.2. Supóngase que la ecuación de Riccati (11.18) tiene una solución que es definida no negativa para $0 \leq k \leq N$; entonces



$$\begin{aligned}
x^T(N)Q_0x(N) + \sum_{k=0}^{N-1} [x^T(k)Q_1x(k) + u^T(k)Q_2u(k)] &= x^T(0)S(0)x(0) \\
+ \sum_{k=0}^{N-1} [u(k) + L(k)x(k)]^T[\Gamma^T S(k+1)\Gamma + Q_2][u(k) + L(k)x(k)] \\
+ \sum_{k=0}^{N-1} \{v^T(k)S(k+1)[\Phi x(k) + \Gamma u(k)] \\
&\quad + [\Phi x(k) + \Gamma u(k)]^T S(k+1)v(k)\} \\
+ \sum_{k=0}^{N-1} v^T(k)S(k+1)v(k)
\end{aligned} \tag{11.24}$$

donde $x(k+1)$ viene dado por (11.5).

*Demuestra*ón. Tenemos la identidad

$$\begin{aligned}
x^T(N)Q_0x(N) &= x^T(N)S(N)x(N) \\
&= x^T(0)S(0)x(0) + \sum_{k=0}^{N-1} [x^T(k+1)S(k+1)x(k+1) \\
&\quad - x^T(k)S(k)x(k)]
\end{aligned} \tag{11.25}$$

Considérense los diferentes términos en el sumatorio y usemos (11.5) y (11.18). Entonces

$$\begin{aligned}
x^T(k+1)S(k+1)x(k+1) &= (\Phi x(k) + \Gamma u(k) + v(k))^T S(k+1)(\Phi x(k) \\
&\quad + \Gamma u(k) + v(k))
\end{aligned} \tag{11.26}$$

y

$$\begin{aligned}
x^T(k)S(k)x(k) &= x^T(k)[\Phi^T S(k+1)\Phi + Q_1 \\
&\quad - L^T(k)(\Gamma^T S(k+1)\Gamma + Q_2)L(k)]x(k)
\end{aligned} \tag{11.27}$$

Introduciendo (11.26) y (11.27) en (11.25) tenemos

$$\begin{aligned}
x^T(N)Q_0x(N) &= x^T(0)S(0)x(0) + \sum \{[\Phi x(k) + \Gamma u(k)]^T S(k+1)v(k) \\
&\quad + v^T(k)S(k+1)[\Phi x(k) + \Gamma u(k)] + v^T(k)S(k+1)v(k)\} \\
&\quad + \sum \{u^T(k)(\Gamma^T S(k+1)\Gamma + Q_2)u(k) \\
&\quad + u^T(k)\Gamma^T S(k+1)\Phi x(k) + x^T(k)\Phi^T S(k+1)\Gamma u(k) \\
&\quad + x^T(k)L^T(k)(\Gamma^T S(k+1)\Gamma + Q_2)L(k)x(k) \\
&\quad - x^T(k)Q_1x(k) - u^T(k)Q_2u(k)\}
\end{aligned}$$

donde se ha sumado y restado el término $u^T Q_2 u$ en la última suma. Reorganizando los términos se completa la demostración. \square

Información completa del estado

Supóngase que $v(k) \equiv 0$ en (11.5) pero que el estado inicial es incierto. El Teorema 11.2 nos da

$$\begin{aligned}
J &= E \left\{ \sum_{k=0}^{N-1} [x^T(k)Q_1x(k) + u^T(k)Q_2u(k)] + x^T(N)Q_0x(N) \right\} \\
&= E\{x^T(0)S(0)x(0)\} + E \left\{ \sum_{k=0}^{N-1} [u(k) + L(k)x(k)]^T[\Gamma^T S(k+1)\Gamma + Q_2] \right. \\
&\quad \left. [u(k) + L(k)x(k)] \right\}
\end{aligned}$$

Como $S(k)$ es semidefinida positiva, el segundo término es no negativo. Además, $S(k)$ es independiente de $u(k)$, y de ahí se deduce que

$$J \geq E x^T(0)S(0)x(0) = m_0^T S(0)m_0 + tr S(0)R_0 \tag{11.28}$$

en la que se ha usado (11.14). La igualdad se tiene para la ley de control de (11.20). El Teorema 11.2 y (11.28) dan una forma alternativa de demostrar el Teorema 11.1.

Supongamos ahora que hay perturbaciones estocásticas actuando sobre el sistema y que todavía es medible la totalidad del estado. Empleando el Teorema 11.2, y el hecho de que $v(k)$ es independiente de $u(k)$ y $x(k)$ se tiene

$$\begin{aligned}
J &= E \left\{ x^T(0)S(0)x(0) + \sum_{k=0}^{N-1} v^T(k)S(k+1)v(k) \right. \\
&\quad \left. + \sum_{k=0}^{N-1} [u(k) + L(k)x(k)]^T[\Gamma^T S(k+1)\Gamma + Q_2][u(k) + L(k)x(k)] \right\}
\end{aligned} \tag{11.29}$$

Utilizando (11.14) se tiene la relación

$$J \geq m_0^T S(0)m_0 + tr S(0)R_0 + \sum_{k=0}^{N-1} tr S(k+1)R_1 \tag{11.30}$$

La igualdad se obtiene para la ley de control de (11.20), que es una ley de control admisible. La diferencia en los costes óptimos (11.28) y (11.30) se debe a la perturbación $v(k)$. La ley de control (11.20) minimiza el coste para el caso de información completa del estado.

La solución al problema LQ conduce a un controlador variable. La matriz de realimentación no depende de x y puede precalcularse desde $k=N$ hasta $k=0$ y almacenarse en el computador. Normalmente sólo se utiliza el controlador estacionario —el controlador constante obtenido cuando aumenta el horizonte temporal. Para procesos y funciones de coste constantes, $S(k)$ —bajo hipótesis bastante generales— convergerá a una matriz constante cuando aumenta el horizonte temporal. En general existen varias soluciones. Si el sistema (11.17) es alcanzable y si (Φ, U) es un par observable, donde

$$Q_1 = U^T U \tag{11.31}$$



entonces existirá una solución única de la ecuación de Riccati que será simétrica definida no negativa. Si hay un término producto cruzado en la función de coste, $Q_{12} \neq 0$, deben considerarse en su lugar las matrices Φ y \tilde{Q}_1 .

Ejemplo 11.1—Control LQ de un integrador doble

Consideremos el integrador doble (véase el Ejemplo A.1) y empleemos el período de muestreo $h=1$. Tomemos como matrices de ponderación en (11.12)

$$Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad Q_2 = [\rho]$$

Ahora se puede estudiar la influencia de la ponderación. Se ha calculado el vector estacionario de realimentación para diferentes valores de ρ . La Figura 11.1 muestra los estados y la señal de control para diferentes valores. Cuando $\rho=0$ que significa que sólo están penalizada la salida, el controlador resultante es el mismo que el controlador de tiempo finito de la Sección 5.4. Cuando ρ aumenta, el módulo de la señal de control disminuye.

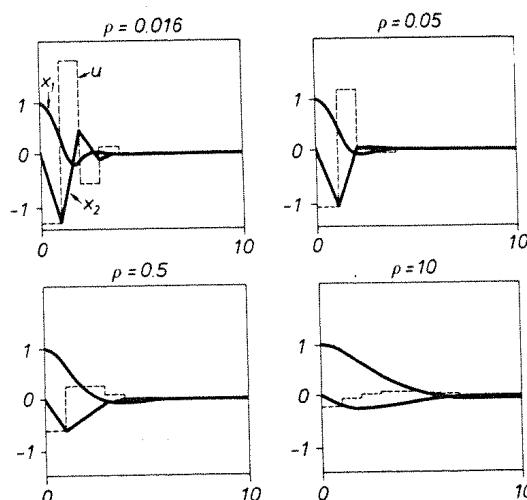


Figura 11.1 Control lineal cuadrático del integrador doble con diferentes pesos, ρ , de la señal de control. El valor inicial del estado (línea continua) es $x(0)^T = [1, 0]$. La línea discontinua es la señal de control.

Propiedades del controlador LQ

El controlador de asignación de polos de la Sección 9.2 y el controlador LQ estacionario tienen la misma estructura. Sin embargo, se han obtenido de forma diferente, y por tanto hay algunas diferencias en sus propiedades.

El controlador de realimentación lineal del estado (11.20) tiene n parámetros. En general, es difícil ajustar los parámetros directamente de forma que se consiga un buen comportamiento. Por el contrario, el procedimiento de ajuste puede ser escoger los n valores propios del sistema en lazo cerrado y utilizar el procedimiento de diseño de la

Sección 9.2. Este procedimiento es muy apropiado para sistemas de una entrada y una salida. Sin embargo, es difícil lograr un compromiso entre la velocidad del sistema y el módulo de la señal de control.

El controlador LQ tiene varias propiedades interesantes. Es aplicable a sistemas multivariados y variables en el tiempo. También, cambiando la relación de valores entre los elementos de las matrices de ponderación, es fácil llegar a un compromiso entre la velocidad de recuperación y el módulo de las señales de control. Además, con las hipótesis hechas sobre el sistema (alcanzable) y la función de coste (simétrica y definida positiva) el controlador LQ resultante siempre dará un sistema en lazo cerrado estable.

Teorema 11.3—Estabilidad del sistema en lazo cerrado. Sea el sistema (11.5) invariante y sea la función de coste (11.12) tal que Q_1 y Q_2 son definidas positivas y $Q_{12} = 0$. Supongamos que (11.18) tiene una solución, \bar{S} , estacionaria y definida positiva. Entonces la estrategia estacionaria de control óptimo

$$u(k) = -Lx(k) = -(Q_2 + \Gamma^T \bar{S} \Gamma)^{-1} \Gamma^T \bar{S} \Phi x(k)$$

da un sistema en lazo cerrado asintóticamente estable

$$x(k+1) = (\Phi - \Gamma L)x(k)$$

Demostración. Se puede emplear el Teorema 5.6 para demostrar que el sistema en lazo cerrado es asintóticamente estable. Hay que demostrar que la función

$$V(x(k)) = x^T(k) \bar{S} x(k)$$

es una función de Lyapunov. V es definida positiva y

$$\begin{aligned} \Delta V(x(k)) &= x^T(k+1) \bar{S} x(k+1) - x^T(k) \bar{S} x(k) \\ &= x^T(k)[\Phi - \Gamma L]^T \bar{S} [\Phi - \Gamma L] x(k) - x^T(k) \bar{S} x(k) \\ &= -x^T(k)[Q_1 + L^T Q_2 L] x(k) \end{aligned}$$

donde se ha hecho uso de (11.23). Como $Q_1 + L^T Q_2 L$ es definida positiva, ΔV es definida negativa. Con lo que el sistema en lazo cerrado es asintóticamente estable. \square

Los polos del sistema en lazo cerrado se pueden obtener de varias formas. Cuando se ha terminado el diseño, los polos se obtienen a partir de

$$\det(\lambda I - \Phi + \Gamma L) = 0$$

Se puede demostrar que los polos son los n valores propios estables del problema de los valores propios generalizados

$$\det \left\{ \begin{bmatrix} I & 0 \\ Q_1 & \Phi^T \end{bmatrix} \lambda - \begin{bmatrix} \Phi & -\Gamma Q_2^{-1} \Gamma \\ 0 & I \end{bmatrix} \right\} = 0 \quad (11.32)$$

La ecuación (11.32) se conoce como *ecuación de Euler* del problema LQ.

A continuación se enuncia sin prueba el Teorema 11.4 para el caso de una entrada y una salida. En la Sección 12.6 se da una demostración.

Teorema 11.4 – Los polos en lazo cerrado de un sistema SISO. Sea un sistema con entrada y salida escalares y supongamos que se utiliza una realimentación estacionaria óptima para un sistema invariante. Supongamos además que en la función de costo se penalizan la salida y la señal de control; es decir, $Q_1 = C^T C$ y $Q_2 = \rho$. Los polos del sistema en lazo cerrado son las n raíces dentro del círculo unidad de la ecuación de orden $2n$

$$\rho + H(z^{-1})H(z) = 0 \quad (11.33)$$

donde

$$H(z) = C(zI - \Phi)^{-1}\Gamma \quad \square$$

Ejemplo 11.2 – Control LQ del integrador doble

Para ilustrar la dependencia de las matrices de ponderación de los polos en lazo cerrado, reconsideraremos el Ejemplo 11.1. La Figura 11.2 muestra los polos del sistema en lazo cerrado para diferentes valores de ρ . Para $\rho=0$ el lugar de las raíces comienza en $z=-1$ y $z=0$. Cuando ρ aumenta, las raíces se desplazan hacia los polos de $H(z)$, $z=1$.

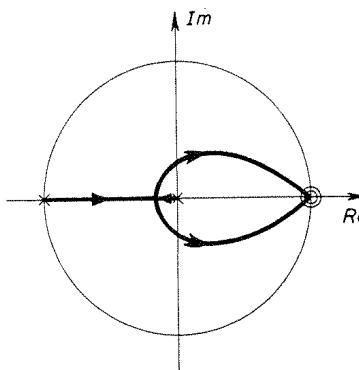


Figura 11.2 Polos en lazo cerrado obtenidos de (11.33) cuando se controla el integrador doble con el controlador óptimo con ρ variando entre 0 e ∞ .

Margen de ganancia del controlador LQ

El Teorema 11.3 muestra que el sistema en lazo cerrado es estable cuando se emplea un controlador LQ . También es posible determinar el margen de ganancia del sistema en lazo cerrado. Considérese el sistema (11.5) con $v(k)=e(k)=0$. La función de transferencia discreta del sistema en lazo abierto es

$$H(z) = C(zI - \Phi)^{-1}\Gamma$$

Supongamos que en la función de coste (11.12) sólo se penalizan las entradas y las salidas; es decir,

$$Q_1 = C^T C$$

ya que $Q_{12}=0$. Controlemos el sistema por un controlador LQ estacionario de realimentación de estado. Entonces el controlador se define por las ecuaciones

$$\begin{aligned} S &= \Phi^T S \Phi + Q_1 - L^T R L \\ L &= R^{-1} \Gamma^T S \Phi \\ R &= \Gamma^T S \Gamma + Q_2 \end{aligned} \quad (11.34)$$

La ecuación algebraica de Riccati, (11.34), puede escribirse

$$Q_1 = (z^{-1}I - \Phi)^T S (zI - \Phi) + (z^{-1}I - \Phi)^T S \Phi + \Phi^T S (zI - \Phi) + L^T R L$$

Ahora se puede usar la ecuación de Riccati para reescribir una ecuación que corresponda a (11.33). Esto da una expresión para los polos en lazo cerrado.

$$\begin{aligned} Q_2 + H^T(z^{-1})H(z) &= Q_2 + \Gamma^T(z^{-1}I - \Phi)^{-T}C^T C(zI - \Phi)^{-1}\Gamma \\ &= Q_2 + \Gamma^T\{S + S\Phi(zI - \Phi)^{-1} + (z^{-1}I - \Phi)^{-T}\Phi^T S \\ &\quad + (z^{-1}I - \Phi)^{-T}L^T R L(zI - \Phi)^{-1}\}\Gamma \\ &= R + RL(zI - \Phi)^{-1}\Gamma + \Gamma^T(z^{-1}I - \Phi)^{-T}L^T R \\ &\quad + \Gamma^T(z^{-1}I - \Phi)^{-T}L^T R L(zI - \Phi)^{-1}\Gamma \\ &= [I + L(z^{-1}I - \Phi)^{-1}\Gamma]^T R [I + L(zI - \Phi)^{-1}\Gamma] \\ &= (I + H_1(z^{-1}))^T R (I + H_1(z)) \end{aligned} \quad (11.35)$$

donde

$$H_1(z) = L(zI - \Phi)^{-1}\Gamma$$

La ecuación (11.35) da una factorización espectral de

$$Q_2 + H^T(z^{-1})H(z)$$

Consideremos el caso SISO. Entonces

$$H(z) = C(zI - \Phi)^{-1}\Gamma = \frac{B(z)}{A(z)}$$

y el sistema en lazo cerrado se define por

$$H_2(z) = C[zI - (\Phi - \Gamma L)]^{-1}\Gamma = \frac{B(z)}{P(z)}$$

Además, la diferencia de retorno del sistema con el controlador LQ es

$$1 + L(zI - \Phi)^{-1}\Gamma = \frac{P(z)}{A(z)}$$

Por tanto,

$$H_1(z) = \frac{P(z) - A(z)}{A(z)}$$

Supongamos ahora que el controlador de (11.20) se sustituye por

$$u(k) = -\beta Lx(k) \quad (11.36)$$

donde β es un escalar positivo. Cuando se emplea (11.36) la diferencia de retorno es

$$1 + \beta H_1(z)$$

Entonces, cuando se utiliza (11.36), la estabilidad del sistema en lazo cerrado se determina partir de

$$A(z) + \beta(P(z) - A(z)) = 0 \quad (11.37)$$

Ahora el margen de ganancia se puede determinar a partir de (11.37) empleando el lugar de las raíces o dibujando la curva de Nyquist para $(P - A)/A$. Como P y A son mónicos y $\text{gra}(P) = \text{gra}(A)$, se tiene que $\text{gra}(P - A) \leq n - 1$. Esto implica que el lugar de las raíces de (11.37) con respecto a β tiende a infinito con al menos una asíntota. Por tanto, el controlador LQ discreto tiene un margen de ganancia finito, lo contrario que el controlador LQ continuo, que tiene un margen de ganancia infinito.

En el caso escalar, (11.35) se puede escribir en la forma

$$\rho A(z^{-1})A(z) + B(z^{-1})B(z) = rP(z^{-1})P(z) \quad (11.38)$$

donde $r = \Gamma^T S \Gamma + \rho$. Empleando (11.35) se tiene la estimación para $|z| = 1$

$$\left| 1 + \frac{P(z) - A(z)}{A(z)} \right|^2 = |1 + H_1(z)|^2 \geq \rho/r$$

Entonces

$$\begin{aligned} |1 + \beta H_1(z)| &= \beta \left| \frac{1}{\beta} - 1 + 1 + H_1(z) \right| \\ &\geq \beta \left| \frac{1}{\beta} - 1 \right| - |1 + H_1(z)| > 0 \end{aligned}$$

La última desigualdad se mantiene si

$$\frac{1}{1 + \sqrt{\rho/r}} < \beta < \frac{1}{1 - \sqrt{\rho/r}} \quad (11.39)$$

La ecuación (11.39) da una estimación de los valores de β para los que (11.37) es estable.

Cómo encontrar las matrices de ponderación

Cuando se usa la teoría de optimización, la función de coste debería construirse según razones físicas. En estos casos, la teoría de control LQG puede verse como una aproximación cuando las ecuaciones de estado se obtienen mediante la linealización de las ecuaciones de movimiento y la función de coste se saca de una función de coste no lineal. Desgraciadamente, estas formulaciones sólo se pueden lograr en unos pocos casos. Uno de ellos es el Ejemplo 11.3.

Ejemplo 11.3 – Pilotaje de un barco

La dinámica linealizada que describe el pilotaje de los barcos se puede describir por la ecuación

$$\frac{d}{dt} \begin{bmatrix} v \\ r \\ \Psi \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ r \\ \Psi \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} \delta \quad (11.40)$$

donde δ es el ángulo del timón, Ψ es el ángulo de dirección, r es la velocidad de giro y v es la velocidad de balanceo. El aumento relativo en la resistencia lateral debido al pilotaje se puede aproximar mediante la expresión

$$\frac{\Delta R}{R} = \frac{\alpha}{T} \int_0^T [vr + \rho \delta^2] dt \quad (11.41)$$

El primer término representa la fuerza de Coriolis debida al acoplamiento de la velocidad de balanceo y a la variación de giro. El segundo término representa la resistencia lateral producida por las deformaciones del timón.

En muchos casos es difícil encontrar funciones de coste cuadráticas naturales. La teoría de control LQ ha encontrado un empleo considerable incluso cuando no se pueden obtener aquéllas. En estos casos, el diseñador de control escoge una función de coste. La ley de realimentación se obtiene directamente resolviendo la ecuación de Riccati. Entonces se analiza el sistema en lazo cerrado respecto a la respuesta transitoria, la respuesta en frecuencia, robustez, etc. Con esta información se modifican los elementos de la función de coste hasta conseguir los resultados deseados. El empleo de un procedimiento de este tipo puede parecer una forma un tanto rara de usar la teoría de optimización. Se puede uno preguntar por qué no se utilizan en su lugar otros métodos, como la búsqueda directa sobre la ganancia de realimentación o la asignación de polos. Se ha visto empíricamente que la teoría LQ es muy fácil de utilizar de esta forma. La búsqueda garantiza automáticamente la obtención de sistemas estables en lazo cerrado con márgenes razonables. A menudo es bastante fácil ver cómo se pueden escoger las matrices de ponderación para actuar sobre las propiedades del sistema en lazo cerrado. En primer lugar se escogen las variables z_i , que corresponden a variables físicas significativas. A continuación se elige la función de coste como una suma ponderada de z_i^2 . Pesos grandes corresponden a respuestas pequeñas. Entonces se evalúan las respuestas del sistema en lazo cerrado a las perturbaciones típicas. Un aspecto particularmente difícil es encontrar la relación de

los pesos entre las variables de estado y las de control, lo que puede hacerse por aproximaciones sucesivas. Algunas veces se dan las especificaciones en términos de las desviaciones máximas permitidas en los estados y en las señales de control para una perturbación dada. Una regla empírica para decidir los pesos en (11.6) es escoger los elementos diagonales como los valores inversos de las desviaciones permitidas. Otra forma es considerar sólo penalizaciones sobre las variables de estado y restricciones sobre las desviaciones del control. Si las restricciones son cuadráticas, los métodos que emplean los multiplicadores de Lagrange dan criterios del tipo de (11.12).

11.3. Teoría de predicción y filtrado

Cuando se emplea el controlador LQ debe ser medible la totalidad del estado. En esta sección se estudia el problema de la estimación de los estados de (11.5) a partir de las medidas de la salida. Se postula un estimador de la misma estructura que el de la Sección 9.3, pero ahora el vector de ganancia se determina de forma diferente. El problema se resuelve como un problema de optimización paramétrica, donde se minimiza la varianza del error de estimación.

Predicción, filtrado y suavizado

Se pueden obtener diferentes estimadores para los estados de (11.5) dependiendo de las medidas de que se disponga. Supongamos que se conocen los datos

$$Y_k = \{y(i), u(i) | i \leq k\}$$

Utilizando Y_k queremos estimar $x(k+m)$. Tenemos tres casos:

- Suavizado ($m < 0$).
- Filtrado ($m = 0$).
- Predicción ($m > 0$).

La Figura 11.3 ilustra los diferentes casos. En esta sección se estudian los problemas de predicción y filtrado. El sistema dinámico resultante recibe el nombre de filtro, independientemente del problema resuelto.

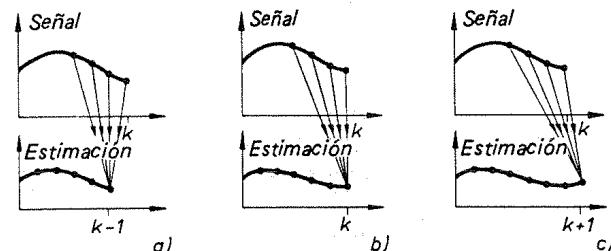


Figura 11.3 Suavizando, filtrado y predicción.

El filtro de Kalman

En el problema de predicción con anticipación de un paso supongamos el proceso descrito por (11.5) con $h=1$. Por simplicidad, supongamos en primer lugar que $R_{12}=0$. Tomemos un estimador de la forma

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k-1) + \Gamma u(k) + K(k)[y(k) - C\hat{x}(k|k-1)] \quad (11.42)$$

El error de reconstrucción $\tilde{x} = x - \hat{x}$ se comporta según

$$\begin{aligned} \tilde{x}(k+1) &= \Phi\tilde{x}(k) + v(k) - K(k)[y(k) - C\hat{x}(k|k-1)] \\ &= (\Phi - K(k)C)\tilde{x}(k) + v(k) - K(k)e(k) \end{aligned} \quad (11.43)$$

En la Sección 9.3 K se utilizaba para dar al sistema del tipo del (11.43) los valores propios deseados. Aquí el problema se enfoca de manera diferente: se tienen en cuenta las propiedades del ruido y el criterio es minimizar la varianza del error de estimación, que se designa por $P(k)$.

$$P(k) = E(\tilde{x}(k) - E\tilde{x}(k))(\tilde{x}(k) - E\tilde{x}(k))^T$$

El valor medio de \tilde{x} se obtiene a partir de (11.43)

$$E\tilde{x}(k+1) = (\Phi - K(k)C)E\tilde{x}(k)$$

Como $E\tilde{x}(0) = m_0$, el valor medio del error de reconstrucción es cero para todos los instantes $k \geq 0$ independiente de K si $E\tilde{x}(0) = m_0$. Ahora la ecuación (11.43) da

$$\begin{aligned} P(k+1) &= E\tilde{x}(k+1)\tilde{x}(k+1)^T \\ &= (\Phi - K(k)C)P(k)(\Phi - K(k)C)^T + R_1 + K(k)R_2K(k)^T \end{aligned} \quad (11.44)$$

ya que $\tilde{x}(k)$, $v(k)$ y $e(k)$ son independientes. Además, $P(0) = R_0$. De (11.44) se deduce que si $P(k)$ es semidefinida positiva, entonces $P(k+1)$ también es semidefinida positiva. Se supone que el criterio es minimizar el escalar $\alpha^T P(k+1) \alpha$, donde α es un vector arbitrario. También se supone que hasta el instante $k-1$ se ha utilizado el vector de ganancia óptima, K . De (11.44),

$$\begin{aligned} \alpha^T P(k+1) \alpha &= \alpha^T [\Phi P(k) \Phi^T + R_1 - K(k)C P(k) \Phi^T - \Phi P(k) C^T K^T(k) \\ &\quad + K(k)(R_2 + C P(k) C^T) K^T(k)] \alpha \end{aligned} \quad (11.45)$$

La ganancia $K(k)$ se puede determinar a partir de (11.45) completando las expresiones de los cuadrados.

$$\begin{aligned} \alpha^T P(k+1) \alpha &= \alpha^T [\Phi P(k) \Phi^T + R_1 - \Phi P(k) C^T (R_2 + C P(k) C^T)^{-1} C P(k) \Phi^T] \alpha \\ &\quad + \alpha^T [K(k) - \Phi P(k) C^T (R_2 + C P(k) C^T)^{-1}] [R_2 + C P(k) C^T] \\ &\quad [K(k) - \Phi P(k) C^T (R_2 + C P(k) C^T)^{-1}]^T \alpha \end{aligned} \quad (11.46)$$



El criterio de (11.46) tiene dos términos. La primera parte es independiente de K y la segunda es no negativa, ya que la matriz $R_2 + CP(k)C^T$ es definida positiva. Entonces el mínimo se obtiene si se elige K de forma que la segunda parte de (11.46) sea cero. Entonces

$$K(k) = \Phi P(k) C^T (R_2 + CP(k)C^T)^{-1} \quad (11.47)$$

$$\begin{aligned} P(k+1) &= \Phi P(k) \Phi^T + R_1 \\ &\quad - \Phi P(k) C^T (R_2 + CP(k)C^T)^{-1} CP(k) \Phi^T \end{aligned} \quad (11.48)$$

Nótese que la ganancia K es independiente de α y que (11.44) también se verifica para el óptimo $K(k)$. La reconstrucción definida mediante las expresiones (11.42), (11.47) y (11.48) recibe el nombre de *filtro de Kalman*. Esto se sintetiza en el siguiente teorema.

Teorema 11.5—El filtro de Kalman. Considerérese el proceso dado por (11.5). La reconstrucción de los estados usando el modelo (11.42) es óptima en el sentido de que se minimiza la varianza del error de reconstrucción si la matriz $R_2 + CP(k)C^T$ es definida positiva y si la matriz de ganancia se escoge de acuerdo con (11.47) y (11.48). La varianza del error de reconstrucción viene dada por (11.48). \square

Nota 1. El problema de reconstrucción se ha resuelto como un problema de optimización paramétrica suponiendo la estructura del estimador de (11.42). De hecho es cierto que la estructura es óptima para perturbaciones gaussianas. También se puede interpretar el criterio como una minimización del error de estimación de una combinación lineal de los estados, $\alpha^T x(k)$.

Nota 2. Mejor que la notación tradicional, $P(k)$, para la varianza, es $P(k|k-1)$. Esta última notación indica que se han empleado las medidas hasta el instante $k-1$ inclusive. Los distintos términos de la ecuación de la varianza (11.48) pueden interpretarse de la siguiente manera: el término $\Phi P \Phi^T$ muestra cómo cambia la varianza debido a la dinámica del sistema, y R_1 representa el incremento de la varianza debido al ruido v [compárese con (6.20)]. El último término muestra cómo disminuye la varianza por la información obtenida a través de las medidas. Nótese que $P(k)$ no depende de las observaciones. Entonces puede precalcularse la ganancia en los instantes futuros y almacenarse en el computador.

Nota 3. El filtro de Kalman se puede interpretar también como la media condicional del estado en el instante $k+1$ dado Y_k ; es decir,

$$\hat{x}(k+1|k) = E\{x(k+1)|Y_k\}$$

$$P(k+1) = E\{[x(k+1) - \hat{x}(k+1|k)][x(k+1) - \hat{x}(k+1|k)]^T | Y_k\}$$

Nota 4. Si $R_{12} \neq 0$, la ganancia y las ecuaciones de la varianza dados en (11.47) y (11.48) resultan

$$K(k) = (\Phi P(k) C^T + R_{12})(CP(k)C^T + R_2)^{-1}$$

$$P(k+1) = \Phi P(k) \Phi^T + R_1 - K(k)(CP(k)C^T + R_2)K^T(k)$$

Las otras ecuaciones se mantienen sin cambios.

Nota 5. El predictor de (11.42) tiene la propiedad de que el estado en el instante k se reconstruye a partir de $y(k-1)$, $y(k-2)$, ... Del mismo modo se puede obtener un filtro que también utilice $y(k)$ para estimar $x(k)$. El problema del filtro queda resuelto por

$$\begin{aligned} \hat{x}(k+1|k+1) &= \Phi \hat{x}(k|k) + \Gamma u(k) + K(k+1)y(k+1) \\ &\quad - C(\Phi \hat{x}(k|k) + \Gamma u(k)) \end{aligned} \quad (11.49)$$

donde

$$K(k) = P(k|k-1)C^T[R_2 + CP(k|k-1)C^T]^{-1}$$

$$P(k|k-1) = \Phi P(k-1|k-1) \Phi^T + R_1$$

$$P(k|k) = P(k|k-1) - K(k)CP(k|k-1)$$

$$P(0|0) = R_0$$

Aquí se utiliza la notación $P(k|k-1)$ en lugar de $P(k)$ para especificar los datos disponibles; $P(k|k)$ puede interpretarse como la varianza del error de estimación en el instante k dado Y_k .

Ejemplo 11.4

Consideremos el sistema escalar

$$x(k+1) = x(k)$$

$$y(k) = x(k) + e(k)$$

donde e tiene una desviación estándar σ y $x(0)$ tiene de valor medio -2 y de varianza 0.5 . El estado es constante y hay que reconstruirlo a partir de medidas con ruido. El filtro de Kalman viene dado por

$$\hat{x}(k+1|k) = \hat{x}(k|k-1) + K(k)(y(k) - \hat{x}(k|k-1)) \quad (11.51)$$

$$K(k) = \frac{P(k)}{\sigma^2 + P(k)} \quad (11.52)$$

$$P(k+1) = \frac{\sigma^2 P(k)}{\sigma^2 + P(k)}$$

La varianza y la ganancia disminuyen con el tiempo. La Figura 11.4 muestra unas realizaciones del error de estimación cuando se usa el filtro de Kalman y cuando se emplea (11.51) con ganancia constante. Una gran ganancia fija da una disminución inicial rápida del error, mientras que la varianza estacionaria es grande. Una ganancia fija pequeña da una disminución del error lenta, pero un mejor comportamiento en estado estacionario.

Ejemplo 11.5

Considérese el sistema de primer orden

$$y(k) + ay(k-1) = e(k) + ce(k-1) \quad (11.53)$$

donde e tiene una desviación estándar σ . Supongamos además que $|c| < 1$. Una representación de (11.53) en el espacio de estados es la dada por

$$x(k+1) = -ax(k) + e(k)$$

$$y(k) = (c-a)x(k) + e(k)$$

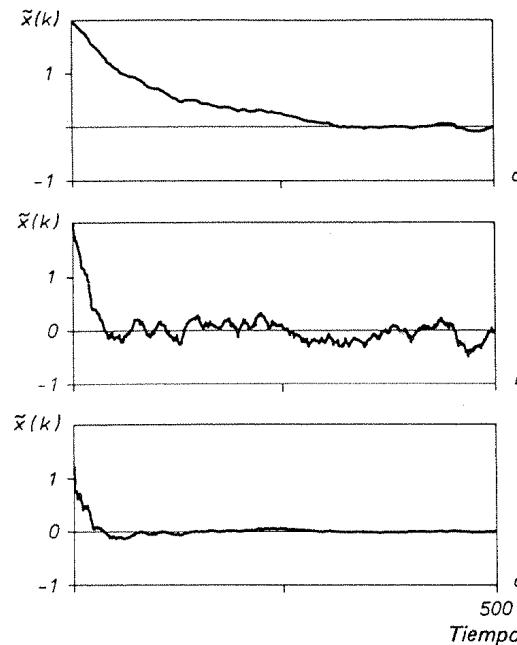


Figura 11.4 Error de estimación del sistema del Ejemplo 11.4 cuando $\sigma = 1$ y cuando se emplean:
a) $K = 0,01$,
b) $K = 0,05$,
c) la ganancia óptima de (11.52).

Es fácil comprobar que el filtro de Kalman en estado estacionario se caracteriza por $P = 0$ y $K = 1$. El predictor de x con anticipación de un paso viene dado por

$$\begin{aligned}\hat{x}(k+1|k) &= -a\hat{x}(k|k-1) + y(k) - (c-a)\hat{x}(k|k-1) \\ &= -c\hat{x}(k|k-1) + y(k)\end{aligned}$$

Además, en estado estacionario la predicción de la salida con anticipación de un paso es

$$\begin{aligned}\hat{y}(k+1|k) &= (c-a)\hat{x}(k+1|k) \\ &= \frac{c-a}{1+cq^{-1}}y(k)\end{aligned}$$

Propiedades de los filtros de Kalman en el dominio de la frecuencia

Cuando se resuelven problemas de diseño empleando técnicas de optimización es muy importante el modelado porque el regulador óptimo, o el filtro óptimo, es sólo una transformación del modelo. Por tanto, es muy útil comprender las propiedades de esta transformación. En esta sección se dará una idea del diseño de los filtros de Kalman analizando las características en el dominio de la frecuencia de un filtro de Kalman estacionario.

Consideremos el problema de estimar el estado del sistema

$$x(k+1) = \Phi_1 x(k) + v(k)$$

a partir de observaciones con ruido

$$y(k) = C_1 x(k) + n(k)$$

donde el ruido n viene dado por

$$n(k) = C_2 z(k) + e(k)$$

$$z(k+1) = \Phi_2 z(k) + w(k)$$

En estos modelos, $\{v(k)\}$, $\{e(k)\}$ y $\{w(k)\}$ son secuencias de variables aleatorias no correlacionadas. El filtro de Kalman estacionario para predicción de x de un paso es

$$\begin{bmatrix} \hat{x}(k+1) \\ \hat{z}(k+1) \end{bmatrix} = \begin{bmatrix} \Phi_1 & 0 \\ 0 & \Phi_2 \end{bmatrix} \begin{bmatrix} \hat{x}(k) \\ \hat{z}(k) \end{bmatrix} + \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} [y(k) - C_1 \hat{x}(k) - C_2 \hat{z}(k)]$$

o

$$\begin{bmatrix} \hat{x}(k+1) \\ \hat{z}(k+1) \end{bmatrix} = \begin{bmatrix} \Phi_1 - K_1 C_1 & -K_1 C_2 \\ -K_2 C_1 & \Phi_2 - K_2 C_2 \end{bmatrix} \begin{bmatrix} \hat{x}(k) \\ \hat{z}(k) \end{bmatrix} + \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} y(k) \quad (11.54)$$

Entonces el filtro de Kalman se caracteriza por la función de transferencia discreta

$$H(z) = [I \quad 0] \begin{bmatrix} zI - \Phi_1 + K_1 C_1 & K_1 C_2 \\ K_2 C_1 & zI - \Phi_2 + K_2 C_2 \end{bmatrix}^{-1} \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \quad (11.55)$$

Un dibujo de la respuesta en frecuencia de la función de transferencia nos muestra cómo el filtro atenúa las diferentes frecuencias. Es muy útil determinar las respuestas en frecuencia del filtro cuando se diseñan filtros de Kalman. En general, las propiedades de la respuesta en frecuencia dependerán del modelo de una forma compleja. Sin embargo, hay algunas propiedades generales que se pueden comprender fácilmente sin necesidad de cálculos detallados.

Lema 11.1. Los ceros de transmisión de la función de transferencia discreta, (11.55), del filtro de Kalman estacionario se obtienen de

$$\det[zI - \Phi_2] = 0$$

Demostración. Un cero de transmisión es un número complejo z tal que una señal de entrada de la forma $z^k y_0$ da salida cero. Para el sistema (11.54) esto implica que $\hat{z}(k) = \hat{z}_0 z^k$ donde $-\infty < k < \infty$ e $y_0 \neq 0$ tal que

$$\begin{aligned}K_1 C_2 \hat{z}_0 - K_1 y_0 &= 0 \\ (zI - \Phi_2 + K_2 C_2) \hat{z}_0 - K_2 y_0 &= 0\end{aligned}$$

$$\begin{bmatrix} K_1 C_2 & -K_1 \\ zI - \Phi_2 + K_2 C_2 & -K_2 \end{bmatrix} \begin{bmatrix} \hat{z}_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 0 & -K_1 \\ zI - \Phi_2 & -K_2 \end{bmatrix} \begin{bmatrix} I & 0 \\ -C_2 & I \end{bmatrix} \begin{bmatrix} \hat{z}_0 \\ y_0 \end{bmatrix} = 0$$

Obviamente esta condición se puede satisfacer cuando $y_0 \neq 0$ y $\hat{z} \neq 0$ para aquellos z que son valores propios de la matriz Φ_2 .

Entonces el filtro de Kalman tendrá ceros en los polos del modelo del ruido. Obtener un filtro de Kalman que bloquee ciertas frecuencias (un filtro elimina banda) se reduce a escoger un modelo de ruido con polos a dichas frecuencias. La atenuación de ciertas frecuencias por el filtro de Kalman se refuerza si la energía del ruido aumenta a dichas frecuencias en el modelo del ruido. \square

11.4. Control lineal cuadrático Gaussiano

En el problema del control LQG se supone que el sistema está gobernado por (11.5) y que la función de coste está dada por (11.12) con $Q_{12}=0$. Se supone que los controles admisibles son tales que $u(k)$ es una función de Y_{k-1} . Esto significa que hay un retardo de cálculo de un período de muestreo.

Todavía son válidos el Teorema 11.2 y (11.29) para el caso de información incompleta del estado. Como (11.20) no es una estrategia admisible de control, el tercer término de (11.29) no se puede hacer igual a cero. El siguiente teorema da la solución, y se enuncia sin demostración.

Teorema 11.6 – El teorema de separación. Considérese el sistema en (11.5). Sean las estrategias admisibles de control tales que $u(k)$ es una función de Y_{k-1} . Supóngase que (11.18), con condiciones iniciales $S(N)=Q_0$, tiene una solución $S(k)$ que es definida no negativa y que $\Gamma^T S(k) \Gamma + Q_2$ es definida positiva. Entonces existe una única estrategia admisible de control

$$u(k) = -L(k)\hat{x}(k|k-1) \quad (11.56)$$

que minimiza el coste esperado (11.12) para $Q_{12}=0$. El valor mínimo de la función de coste es el dado por la expresión

$$J = m_0^T S(0) m_0 + \text{tr}S(0)R_0 + \sum_{k=0}^{N-1} \text{tr}S(k+1)R_1 + \sum_{k=0}^{N-1} \text{tr}P(k)L^T(k)[\Gamma^T S(k+1)\Gamma + Q_2]L(k) \quad (11.57) \quad \square$$

Nota 1. La diferencia en los costes mínimos entre (11.30) y (11.57) se debe a la estimación de las variables de estado.

Nota 2. Es posible modificar el Teorema 11.6 para otras estrategias admisibles de control; por ejemplo, el caso en que $u(k)$ pueda ser una función de Y_{k-1} y de $y(k)$.

Una consecuencia del teorema de separación es que se puede dividir el problema de síntesis en dos partes, que se pueden resolver separadamente. En primer lugar se

resuelve el problema de control determinístico, obteniendo $L(k)$. En segundo lugar, se estima el estado empleando un filtro de Kalman. En la Figura 11.5 se muestra un diagrama de bloques del sistema con una ley de control óptima.

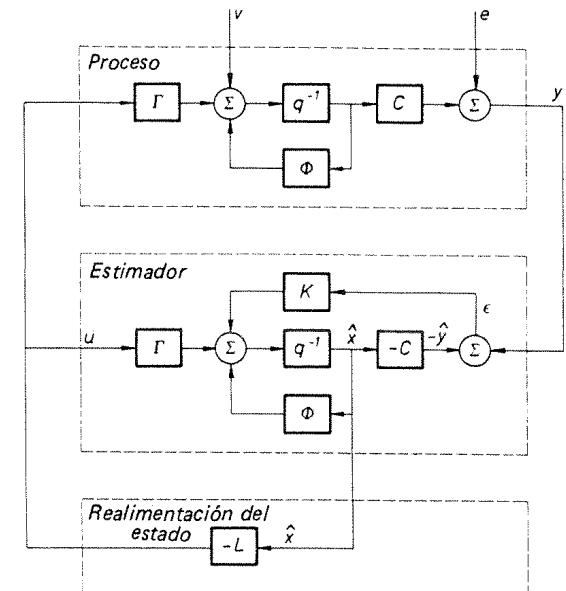


Figura 11.5 Teorema de separación.

Dualidad

Las soluciones del problema de control LQ y del de estimación del estado son muy similares. Se puede demostrar que el problema de estimación del estado es equivalente a un problema LQ . La equivalencia se ilustra mediante la tabla siguiente, que muestra las sustituciones necesarias para convertir el problema de control óptimo en un problema de estimación del estado.

Problema de control óptimo	Problema de estimación del estado
k	$N-k$
Φ	Φ^T
Γ	C^T
Q_0	R_0
Q_1	R_1
Q_{12}	R_{12}
S	P
L	K^T



Propiedades del sistema en lazo cerrado

El sistema en lazo cerrado con control LQG se describe por

$$x(k+1) = \Phi x(k) + \Gamma u(k) + v(k)$$

$$y(k) = Cx(k) + e(k)$$

$$u(k) = -L\hat{x}(k|k-1)$$

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k-1) + \Gamma u(k) + K[y(k) - C\hat{x}(k|k-1)]$$

Introduciendo x y $\tilde{x} = x - \hat{x}$, las ecuaciones se pueden escribir en la forma

$$\begin{bmatrix} x(k+1) \\ \tilde{x}(k+1) \end{bmatrix} = \begin{bmatrix} \Phi - \Gamma L & \Gamma L \\ 0 & \Phi - KC \end{bmatrix} \begin{bmatrix} x(k) \\ \tilde{x}(k) \end{bmatrix} + \begin{bmatrix} I \\ I \end{bmatrix} v(k) + \begin{bmatrix} 0 \\ -K \end{bmatrix} e(k)$$

La dinámica del sistema en lazo cerrado queda determinada por $\Phi - \Gamma L$ y $\Phi - KC$; es decir, por la dinámica del problema de control LQ determinístico correspondiente y por la del filtro óptimo (compárese con la Sección 9.4).

El problema del servo

En la Sección 9.5 se ha estudiado el problema del servo para un controlador de realimentación de estado. Para el problema LQG se puede introducir la señal de referencia de la misma forma que en la Figura 9.7. La única diferencia es que la matriz de realimentación L se obtiene minimizando la función de coste cuadrática.

11.5. Aspectos prácticos

Las secciones anteriores han mostrado cómo se pueden resolver los problemas de control LQ y LQG. Cuando se aplica control LQG se presentan varios problemas de tipo práctico. Uno es el de escoger los parámetros de diseño —es decir, los pesos en la función de coste— que se analizó en la Sección 11.2 y el período de muestreo. Otro problema es la dificultad para obtener buenos modelos para el proceso y para las perturbaciones. Otro problema más es la realización de los cálculos numéricos necesarios para lograr el controlador resultante. En esta sección analizaremos algunos de estos problemas.

Complejidad del modelo

Una de las críticas que se hace al control LQ es que se debe disponer de un modelo exacto completo del proceso. La mayor parte de los procesos físicos son de orden elevado. Sin embargo, para el control, a menudo es suficiente con emplear una aproximación de bajo orden. En el Capítulo 13 se exponen formas de obtener modelos matemáticos.

Una forma de disminuir la sensibilidad a los errores de modelado es reducir el ancho de banda de sistema en lazo cerrado cambiando los pesos en la función de coste. Compárese esto con los resultados de robustez de la Sección 5.2. Otra forma de disminuir la sensibilidad a los errores de modelado es introducir un ruido artificial, lo que significa que las covarianzas del ruido usadas en el diseño del filtro de Kalman son mayores que los valores reales.

Solución de la ecuación de Riccati

En muchos casos, sólo se instala el controlador óptimo estacionario, lo que significa que hay que calcular los valores estacionarios de las ecuaciones de Riccati, (11.18) y (11.48). Hay varias formas de hacerlo numéricamente. Una forma es suponer un S o P constantes y resolver las ecuaciones algebraicas. Una forma inmediata para lograr la solución es iterar las ecuaciones hasta que se obtiene una condición estacionaria. Sin embargo, es importante realizar los cálculos de forma que se garantice que la solución es simétrica y definida positiva. Se han realizado métodos especiales para resolver la ecuación de Riccati, como los algoritmos de la raíz cuadrada y del doblado. Cuando se utiliza el método de la raíz cuadrada, se calcula la raíz cuadrada de S o de P . Esto proporciona mejores propiedades numéricas. El algoritmo del doblado o algoritmo rápido, acelera el cálculo del valor estacionario calculando la solución en el instante $2k$ cuando se conocer la solución en el instante k . Existen muchos libros y trabajos sobre diferentes métodos.

Elección del período de muestreo

La elección del período de muestreo está influenciada por el modo en que se dan las especificaciones del problema de control. Se pueden considerar dos casos.

En el primer caso se supone que se dan las especificaciones en forma del amortiguamiento y la respuesta deseados del sistema de control en lazo cerrado sin utilizar señales de control excesivamente grandes. Entonces es natural determinar el control mediante iteraciones en los pesos de la función de coste muestreada (11.12). Para ponerlo en marcha hay que hacer una primera elección del período de muestreo basándose en las especificaciones. Lo normal es escoger el período de muestreo en relación con la dinámica del sistema en lazo cerrado, como se expuso en la Sección 9.2. Esto significa que puede hacer falta realizar una o dos iteraciones en el período de muestreo. La dinámica en lazo cerrado es una función complicada de la función de coste.

En el segundo caso se supone que las especificaciones se dan en términos de la función de coste continua (11.6). Entonces el coste lo minimiza el controlador LQ continuo. El coste aumenta cuadráticamente con el período de muestreo h , para h pequeños, lo que significa que existe un límite para el que, en la práctica, no merece la pena reducir más el período de muestreo. Es posible obtener una aproximación del aumento del coste debido a un aumento del período de muestreo (véanse las referencias). Cuando se dispone de buenos programas de diseño interactivo, es fácil comprobar el coste y el comportamiento para algunos períodos de muestreo.

11.6. Conclusiones

En este capítulo se ha expuesto el diseño óptimo basado en modelos en el espacio de estados. Los controladores LQ y los filtros de Kalman tienen muchas propiedades muy buenas. Sólo es posible una breve visión de conjunto. El principal problema del control LQ es trasladar las especificaciones del sistema a la función de coste. Esto se suele hacer por un procedimiento iterativo, por lo que es necesario disponer de buenos programas interactivos.

11.7. Problemas

- 11.1. Considérese el sistema de primer orden

$$\dot{x} = -ax + bu$$

Supóngase que hay que minimizar la función de coste (11.6) con $Q_{1c}=1$ y $Q_{2c}=\rho$. Determíñese la correspondiente función de coste discreta (11.12).

- 11.2. Considérese el integrador doble continuo del Ejemplo A.1. Supóngase que hay que minimizar la función de coste (11.6) con

$$Q_{1c} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad Q_{2c} = 1$$

Determíñense Q_1 , Q_{12} y Q_2 en la función de coste discreta (11.12).

- 11.3. Dado el sistema

$$x(k+1) = ax(k) + bu(k)$$

con la función de coste

$$J = \sum_{k=0}^N x^2(k)$$

Sea la estrategia admisible de control tal que $u(k)$ es una función de $x(k)$. Determíñese la estrategia que minimiza el coste.

- 11.4. Considérese el sistema del Problema 11.3. Determíñese la estrategia de control que minimiza el coste cuando las estrategias admisibles de control son tales que $u(k)$ es una función de $x(k-1)$.

- 11.5. El modelo de inventario del Ejemplo A.5 puede describirse por

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u(k) \\ y(k) &= [1 \ 0]x(k) \end{aligned}$$

- a) Determíñese el controlador LQ estacionario cuando $Q_1=C^TC$ y $Q_2=\rho$.
 - b) Determíñense los polos del sistema en lazo cerrado e investigúese cómo dependen del peso dado a la señal de control, ρ .
 - c) Simúlese el sistema empleando el controlador de a). Supóngase que $x(0)^T=[1 \ 1]$ y considérense la salida y la señal de control para diferentes valores de ρ .
- 11.6. Considérese el sistema de dos tanques con el operador de transferencia discreto dado en el Problema 3.10c). Usese (11.33) y dibújese el lugar de las raíces que muestra los polos

en lazo cerrado cuando el sistema se controla mediante un controlador LQ estacionario para la función de coste

$$J = \sum_{k=0}^{\infty} [y(k)^2 + \rho u(k)^2]$$

- 11.7. Compruébese que se puede obtener una ley de control de tiempo mínimo, tal que la matriz $\Phi - \Gamma L$ tenga todos sus valores propios en el origen, mediante la optimización discreta con $Q_2=0$, $Q_1=0$ y $Q_0=I$.
- 11.8. Considérese el problema del pilotaje del barco caracterizado por el modelo de (11.40) y la función de coste de (11.41). Usense los valores $a_{11}=-0,454$, $a_{12}=-0,433$, $a_{21}=-4,005$, $a_{22}=-0,807$, $b_1=0,097$, $b_2=-0,807$, $\alpha=0,014$ y $\rho=0,08$. Determíñese la realimentación de estado óptimo cuando $h=5$ s.
- 11.9. El problema del pilotaje del barco se aproxima a veces empleando el modelo de segundo orden

$$\frac{d}{dt} \begin{bmatrix} \Psi \\ r \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -\alpha \end{bmatrix} \begin{bmatrix} \Psi \\ r \end{bmatrix} + \begin{bmatrix} 0 \\ k \end{bmatrix} \delta$$

y la siguiente aproximación de la función de coste:

$$J = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [\Psi^2 + \rho \delta^2] dt$$

Determíñese la realimentación óptima para un regulador muestreado. Usense los parámetros $\alpha=0,001$, $k=0,0005$ y $\rho=0,08$ y el período de muestreo $h=5$ s.

- 11.10. Considérese el controlador LQ determinado en el Problema 11.5 para el modelo de inventario.
- Usese (11.39) para obtener una estimación del margen de ganancia para el sistema en lazo cerrado.
 - Usese (11.37) para determinar el margen de ganancia y compárese con el estimado mediante a).
- 11.11. Se genera un proceso estocástico mediante

$$\begin{aligned} x(k+1) &= 0.5x(k) + v(k) \\ y(k) &= x(k) + e(k) \end{aligned}$$

donde v y e son procesos de ruido blanco no correlacionados con covarianzas r_1 y r_2 , respectivamente. Además, $x(0)$ tiene una distribución normal con media cero y desviación estándar σ . ¿Cuál es la ganancia en estado estacionario? Calcúlese el polo del filtro estacionario y compárese con el polo del sistema.

- 11.12. El integrador doble con ruido de proceso puede describirse mediante las ecuaciones

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}x(k) + \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}u(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}v(k) \\ y(k) &= [1 \ 0]x(k) \end{aligned}$$

donde $v(k)$ es una secuencia de variables aleatorias normales de media cero e independientes. Supóngase que $x(0)$ es normal con media $Ex(0)=[1 \ 1]^T$ y matriz de covarianza unidad.

- a) Determinense las ecuaciones de la matriz de covarianza del error de reconstrucción y del vector de ganancia en el filtro de Kalman.
 b) Simúlense las ecuaciones de la covarianza y de la ganancia y determinense la velocidad de la convergencia y los valores estacionarios.

11.13. Considérese el integrador doble del Problema 11.12, pero sea la salida

$$y(k) = [1 \ 0]x(k) + v(k)$$

- a) Determinense las ecuaciones para la matriz de covarianza del error de reconstrucción y el vector de ganancia en el filtro de Kalman.
 b) Simúlense las ecuaciones de la covarianza y de la ganancia y determinense la velocidad de la convergencia y los valores estacionarios.

11.14. Dado el sistema

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}v(k) + \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \\ y(k) &= [1 \ 0]x(k) \end{aligned}$$

donde $v(k)$ es un ruido blanco de media cero con desviación estándar 0,1. Supóngase que se conoce $x(0)$ exactamente. Determinese la estimación de $x(k+3)$, dado $y(k)$ que minimiza el error de predicción. Usese para determinar la mejor estimación de $y(3)$ y su varianza.

11.15. Se define la señal $x(k)$ como

$$\begin{aligned} x(k+1) &= ax(k) + v(k) \\ y(k) &= x(k) + e(k) \end{aligned}$$

donde v y e son procesos independientes de ruido blanco con media cero. Las varianzas son, respectivamente, 1 y σ^2 . Se estima la señal x utilizando un suavizado exponencial del tipo

$$\hat{x}(k|k) = \alpha\hat{x}(k-1|k-1) + (1-\alpha)y(k)$$

Determinese una expresión para la dependencia del error de estimación de los parámetros α y σ^2 . Compárese con el filtro de Kalman óptimo estacionario.

- 11.16. Muéstrese que el Teorema 11.5 se puede generalizar a las situaciones en las que las perturbaciones $e(k)$ y $v(k)$ tienen valores medios constantes aunque desconocidos. (Compárese con la Sección 9.4.)
 11.17. Se mide una variable constante x mediante dos sensores diferentes. Sin embargo, las medidas están contaminadas por ruido y tienen diferente exactitud. Supóngase que el sistema está descrito por

$$\begin{aligned} x(k+1) &= x(k) \\ \hat{y}(k) &= Cx(k) + e(k) \end{aligned}$$

donde $C^T = [1 \ 1]$ y $e(k)$ es un vector de ruido blanco con media cero y con matriz de covarianza

$$R_2 = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}$$

Estímese x mediante

$$\hat{x}(k) = a_1y_1(k) + a_2y_2(k)$$

Determinense las constantes a_1 y a_2 tales que el valor medio del error de predicción sea cero y tal que la varianza del error de predicción sea lo más pequeña posible. Compárese la varianza mínima con los casos en los que sólo se emplea una medida. Compárese la solución con el filtro de Kalman.

- 11.18. Probar que el filtro de estimación dado en (11.49) y (11.50) es el filtro óptimo en el sentido de que minimiza la varianza del error de estimación.

- 11.19. Considérese el diseño de un filtro de Kalman para estimar la velocidad en el mando de un motor basado en medidas angulares. La dinámica básica del motor, que relaciona la corriente con el ángulo, viene dada por

$$G(s) = \frac{1}{s(s+1)}$$

Supóngase que hay perturbaciones de baja frecuencia (fricción) que se modelan por la expresión

$$z_1(k+h) = z_1(k) + w_1(k)$$

Supóngase también que se desea filtrar las perturbaciones debido a la resonancia mecánica a la frecuencia ω . Esta señal se modela como la señal obtenida gobernando con un ruido blanco un sistema con función de transferencia

$$G(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$

Determinense los diagramas de Bode para el filtro de Kalman para $\zeta=0.05$, $\omega=0.1$ y $\omega=2$. Considérese un período de muestreo de 0,05 s. Estúdiese también la influencia de diferentes intensidades relativas de la perturbación de baja frecuencia y de la de banda limitada.

11.8. Referencias

- El control LQG y los filtros óptimos se tratan en muchos libros de texto; por ejemplo, KWAKernaak, H. y R. SIVAN (1972): *Linear Optimal Control Systems*. New York: Wiley. Interscience.
 ANDERSON, B. D. O. y J. B. MOORE (1971): *Linear Optimal Control*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.
 ÅSTRÖM, K. J. (1970): *Introduction to Stochastic Control Theory*. New York: Academic Press.
 ANDERSON, B. D. O. y J. B. MOORE (1979): *Optimal Filtering*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.

Kalman y Bucy fueron quienes hicieron las aportaciones más importantes al desarrollo de los filtros óptimos recursivos expuestos en la Sección 11.3. Véanse

- KALMAN, R. E. (1960): «A New Approach to Linear Filtering and Prediction Problems,» *Journal Basic Eng.*, 82, March, 34-45.



- KALMAN, R. E. y R. S. BUCY (1961): «New Results in Linear Filtering and Prediction Theory,» *Trans. ASME, Ser. D, Journal Basic Eng.*, 83, December, 95-107.
BUCY, R. S. (1959): «Optimum Finite Time Filters for a Special Nonstationary Class of Inputs,» Internal memorandum BBD-600, Johns Hopkins University, Applied Physics Lab.

Los algoritmos numéricos para resolver la ecuación de Riccati se exponen en

- BIERMAN, G. (1977): *Factorization Methods for Discrete Estimation*. New York: Academic Press.
PAPPAS, T., A. J. LAUB, and N. R. SANDELL, JR. (1980): «On the Numerical Solution of the Discrete Time Riccati Equation,» *IEEE Trans. Autom. Contr.*, AC-25, August, 631-41.
VAN DOOREN, P. (1981): «A Generalized Eigenvalue Approach for Solving Riccati Equations,» *SIAM Journal Sci. Stat. Comp.*, 2, 121-35.

La elección del período de muestreo para los controladores LQ se expone en

- ÅSTRÖM, K. J. (1963): «On the Choice of Sampling Rates in Optimal Linear Systems,» Internal report IBM San Jose Research Laboratory.
MELZER, S. M. y B. C. KUO (1971): «Sampling Period Sensitivity of the Optimal Sampled Data Linear Regulator,» *Automatica*, 7, 367-70.

El teorema de separación apareció por primera vez en una revista de economía:

- SIMON, H. A. (1956): «Dynamic Programming under Uncertainty with a Quadratic Criterion Function,» *Econometrica*, 24, 74.

Versiónes discretas del teorema de separación se pueden encontrar en

- GUNKEL, III, T. L., G. F. FRANKLIN (1963): «A General Solution for Linear Sampled Data Control,» *Trans. ASME Journal Basic Eng.* 85-D, 197-201.

El margen de ganancia para los controladores LQG discretos se expone en

- SAFONOV, M. G. (1980): *Stability and Robustness of Multivariable Feedback Systems*. Cambridge, Mass.: MIT Press.

La robustez de los controladores LQG se expone en

- DOYLE J. C., G. STEIN (1981): «Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis,» *IEEE Trans. Autom. Contr.*, AC-26, Feb., 4-16.

Métodos de diseño óptimo: formulación entrada-salida

Objetivo:

Mostrar cómo se pueden formular y resolver los problemas de regulación óptima utilizando los modelos de entrada-salida. Deducir fórmulas para predicción óptima, control de varianza mínima y control LQG.

12.1. Introducción

En este capítulo se consideran los métodos de diseño óptimo basados en modelos de entrada-salida. Se exponen los diseños de reguladores basados en modelos lineales y criterios cuadráticos. Esta es una clase de problemas que admite soluciones perfectamente conocidas. Los problemas están resueltos en el Capítulo 11 por otros métodos. Basar el método en el modelo de entrada-salida proporciona una mejor comprensión y también permite obtener diferentes algoritmos numéricos.

La formulación del problema se da en la Sección 12.2. Esto incluye el análisis del modelo en cuanto a dinámica, perturbaciones y criterios, así como las especificaciones de los controles admisibles. El modelo se da en función de tres polinomios. También se resuelve un problema muy simple utilizando los principios fundamentales. Este ejemplo muestra claramente que los problemas de control óptimo y filtrado óptimo están estrechamente relacionados. El problema de predicción se resuelve en la Sección 12.3. La solución se obtiene fácilmente por división de polinomios. Se da una fórmula explícita simple para la función de transferencia del predictor óptimo. La ley de control de varianza mínima se deduce en la Sección 12.4. Para sistemas con inversos estables, la ley de control se obtiene en función de los polinomios que caracterizan al predictor óptimo. Para sistemas con inversos inestables, la solución se obtiene resolviendo una ecuación polinómica diofántica del tipo discutido en el Capítulo 10. El problema de control de varianza mínima puede interpretarse, de este modo, como un problema de asignación de polos. Esto ayuda a una elección adecuada de los polos en lazo cerrado y de los polos del observador en el problema de asignación de polos.



El problema del control LQG se resuelve en la Sección 12.5. Se demuestra que la solución puede expresarse en función de la factorización espectral y la solución de una ecuación diofántica. Aspectos prácticos, como la elección del período de muestreo, se dan en la Sección 12.6.

12.2. Formulación del problema

Se supone que el proceso que se va a controlar es lineal e invariante en el tiempo y que tiene una entrada u y una salida y . La dinámica del proceso está caracterizada por una combinación de un retardo de tiempo y una función de transferencia racional. También se supone que las perturbaciones pueden describirse como ruido blanco filtrado. Se considera un problema de regulación en estado estacionario. El criterio se basa en la desviación cuadrática media de las señales de control y de salida. En la definición formal del problema que se da a continuación, se supone que el modelo y el criterio están muestreados (compárese con las Secciones 3.2 y 11.1).

Dinámica del proceso

Supongamos que la dinámica del proceso está caracterizada por

$$x(k) = \frac{B_1(q)}{A_1(q)}u(k) \quad (12.1)$$

donde $A_1(q)$ y $B_1(q)$ son polinomios en el operador desplazamiento hacia adelante.

Perturbaciones

Supongamos que la influencia del entorno en el proceso puede caracterizarse por perturbaciones que son procesos estocásticos. Como el sistema es lineal, puede utilizarse el principio de superposición para reducir todas las perturbaciones a una perturbación equivalente v en la salida del sistema. Por tanto, la salida del sistema está definida por

$$y(k) = x(k) + v(k) \quad (12.2)$$

Además supongamos que la perturbación v puede representarse como la salida de un sistema lineal alimentado con ruido blanco; es decir,

$$v(k) = \frac{C_1(q)}{A_2(q)}e(k) \quad (12.3)$$

donde $C_1(q)$ y $A_2(q)$ son polinomios en el operador desplazamiento hacia adelante y $\{e(k)\}$ es una secuencia de variables aleatorias independientes o no correlacionadas de media cero y desviación estándar σ . La perturbación v puede ser un proceso aleatorio

estacionario. Sin embargo, éste puede desplazarse, porque el polinomio $A_2(q)$ puede ser inestable. El modelo del proceso y su entorno pueden reducirse a una forma típica. Eliminando v y x entre (12.1), (12.2) y (12.3) e introduciendo

$$\begin{aligned} A &= A_1 A_2 \\ B &= B_1 A_2 \\ C &= C_1 A_1 \end{aligned} \quad (12.4)$$

se obtiene el siguiente modelo

$$A(q)y(k) = B(q)u(k) + C(q)e(k) \quad (12.5)$$

Este es el modelo canónico, el cual será la base del diseño del control. En el caso especial en el que no hay perturbaciones, el modelo es simplemente una función de transferencia discreta racional (véase la Sección 3.6). Cuando no hay señal de control, el modelo es un proceso estocástico con densidad espectral racional o proceso ARMA (véase la Sección 6.5).

La ecuación (12.5) puede normalizarse para que los coeficientes principales de $A(q)$ y $C(q)$ sean la unidad. Tales polinomios se denominan *mónicos*. El polinomio C también puede multiplicarse por una potencia arbitraria de q , puesto que no cambia la estructura de correlación de $C(q)e(t)$. Esto puede utilizarse para normalizar C de modo que $\text{gra } C = \text{gra } A$ o $\text{grad } C = \text{gra } A - 1$. Los polinomios $A(q)$ y $B(q)$ pueden tener ceros dentro o fuera del círculo unidad. Se supone que todos los ceros del polinomio $C(q)$ están dentro del círculo unidad. Por factorización espectral (Teorema 6.3), el polinomio $C(q)$ puede cambiarse para que todos sus ceros estén dentro del círculo unidad o sobre la frontera. Se emplea un ejemplo para ilustrar este importante punto.

Ejemplo 12.1 – Modificaciones del polinomio C

Considérese el polinomio

$$C(z) = z + 2$$

el cual tiene el cero $z = -2$ fuera del círculo unidad. Consideremos la señal

$$n(k) = C(q)e(k)$$

donde $\{e(k)\}$ es una secuencia de variables aleatorias no correlacionadas de media nula y varianza unidad. La densidad espectral de n está dada por

$$\phi(e^{i\omega h}) = \frac{1}{2\pi} C(e^{i\omega h})C(e^{-i\omega h})$$

Puesto que

$$\begin{aligned} C(z)C(z^{-1}) &= (z + 2)(z^{-1} + 2) = (1 + 2z^{-1})(1 + 2z) \\ &= (2z + 1)(2z^{-1} + 1) \\ &= 4(z + 0.5)(z^{-1} + 0.5) \end{aligned}$$

la señal n también puede representarse como

$$n(k) = C^*(q)e(k)$$

donde

$$C^*(z) = 2z + 1$$

es la recíproca del polinomio $C(z)$ (véase la Sección 3.3).

Si los cálculos de (12.4) dan un polinomio $C(q)$ que tiene todos sus polos fuera del círculo unidad, el polinomio C se factoriza como

$$C = C^+ C^-$$

donde C^- contiene todos los factores con ceros fuera del círculo unidad. Entonces, el polinomio C se sustituye por $C^+ C^{-*}$.

El nombre *Gaussiano* en LQG es en realidad ligeramente engañoso. Las demostraciones muestran que la distribución de probabilidad no tiene importancia mientras las variables aleatorias $\{e(k)\}$ sean independientes.

Criterios

En la regulación en estado estacionario tiene sentido expresar los criterios en términos de la *varianza en estado estacionario* de las variables de control y de la salida del proceso. Para la regulación de sistemas con una salida, el criterio puede ser minimizar la varianza de la salida. Esto se expuso en la Sección 7.6. Compárese también con la Figura 7.7. Esto conduce al criterio

$$J_{mv} = E[y^2(k)] \quad (12.6)$$

donde se elige la escala de modo que $y=0$ corresponda al punto deseado. Una ley de control que minimice el criterio (12.6) se denomina *control de varianza mínima*. También se puede expresar el criterio como

$$J_\infty = \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_1^N y^2(k)$$

Obsérvese que este criterio es una aproximación de la función de coste continua

$$J_c = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T y^2(t) dt \quad (12.7)$$

En la Sección 11.1 se presenta una aproximación más exacta, que tiene en cuenta el comportamiento de las señales entre los instantes de muestreo. Algunas consecuencias de esta aproximación se tratan en la Sección 12.6.

Las propiedades de las señales de control bajo control de varianza mínima dependen críticamente de los períodos de muestreo. Un período de muestreo corto produce una gran varianza de las señales de control y períodos de muestreo grandes producen una varianza pequeña. En algunos casos se desea llegar a un compromiso

entre las varianzas de las señales de control y de salida. Esto puede hacerse introduciendo la función de coste

$$J_{lq} = E[y^2(k) + \rho u^2(k)] \quad (12.8)$$

La ley de control que minimiza este criterio se denomina *ley de control lineal cuadrática*.

Controles admisibles

Supongamos que la ley de control es $u(k)$; es decir, el valor de la señal de control en el instante k es una función de $y(k), y(k-1), \dots, y(k-1), u(k-1), u(k-2), \dots$. De este modo el retardo del cálculo es insignificante en comparación con el período de muestreo. Es muy fácil modificar el resultado para incluir el efecto del retardo computacional.

Existen dos versiones de la teoría. Puede postularse una ley de control lineal. Entonces es suficiente suponer que las perturbaciones $e(i)$ y $e(j)$ no están correlacionadas para $i \neq j$. Si se supone que $e(i)$ y $e(j)$ son independientes, puede demostrarse que la ley de control óptima es lineal. La fórmula para la ley de control óptima es la misma en ambos casos.

Un ejemplo

El problema de control óptimo definido por el modelo de (12.5) y el criterio de (12.6) se soluciona de un modo especial. La solución, que se obtiene fácilmente de los principios fundamentales, aclara las suposiciones realizadas. También indica cómo puede resolverse el problema general.

Consideremos el sistema de primer orden

$$y(k+1) + ay(k) = bu(k) + e(k+1) + ce(k) \quad (12.9)$$

donde $|c| < 1$ y $\{e(k)\}$ es una secuencia de variables aleatorias independientes con varianza unidad.

Consideremos la situación en el instante k . Se han observado las salidas $y(k), y(k-1), \dots$. El control $u(k)$ debería determinarse para que y esté tan próximo a cero como sea posible. Se deduce de (12.9) que $y(k+1)$ puede cambiarse arbitrariamente mediante una elección adecuada de $u(k)$. Como $e(k+1)$ es independiente de $y(k)$ y de los términos de la parte derecha de (12.9), se deduce que

$$\text{Var } y(k+1) \geq \text{Var } e(k+1) = 1 \quad (12.10)$$

El término $e(k)$ puede calcularse en función de los datos conocidos $y(k), y(k-1), \dots$ y $u(k-1), u(k-2), \dots$. Cuando se conocen las variables $y(k)$ y $e(k)$, la ley de control

$$u(k) = [ay(k) - ce(k)]/b \quad (12.11)$$

produce

$$y(k+1) = e(k+1) \quad (12.12)$$

lo cual corresponde al límite inferior de (12.10). Si se utiliza en cada paso la ley de control (12.11), la ecuación (12.12) se mantiene para todo k . Entonces, el cálculo de $e(k)$ a partir de los datos disponibles en el instante k es trivial y la ley de control de (12.11) puede escribirse como

$$u(k) = \frac{a - c}{b} y(k) \quad (12.13)$$

De este modo, el control óptimo es una simple realimentación proporcional con ganancia $(a - c)/b$.

Para analizar las propiedades del sistema en lazo cerrado bajo control óptimo, eliminemos u entre (12.9) y (12.13). Esto produce

$$y(k+1) + cy(k) = e(k+1) + ce(k)$$

Obsérvese que el sistema en lazo cerrado tiene el polinomio característico

$$C(z) = z + c$$

Esto demuestra la importancia de suponer que el polinomio $C(z)$ es estable. Esta ecuación en diferencias tiene la solución

$$y(k) = e(k) + (-c)^{k-k_0}[y(k_0) - e(k_0)]$$

Como el módulo de c es menor que la unidad, el último término tiende a cero cuando $k - k_0$ tiende a infinito. De este modo, la ley de control en (12.13) proporciona la varianza mínima en estado estacionario.

Con este resultado pueden hacerse algunas observaciones. La cantidad $-ay(k) + bu(k) + ce(k)$ puede interpretarse como la mejor estimación de $y(k+1)$, con los datos disponibles en el instante k . La cantidad $e(k+1)$ es el error de predicción. La ley de control de (12.13) implica que la señal de control se elige de manera que el valor de la predicción sea igual al valor de referencia, el cual es cero en este caso. Entonces, el error de control es igual al error de predicción. De este modo, la solución al problema del control de varianza mínima está estrechamente relacionada con la solución del problema de predicción. Además, el problema de predicción se resuelve antes de buscar la solución general del problema de control de varianza mínima.

12.3. Predicción óptima

La teoría de predicción puede establecerse de diferentes modos, los cuales se diferencian en las suposiciones que se hagan sobre el proceso, los criterios y los predictores admisibles. En la Sección 11.3 se dio una formulación. En esta sección haremos las siguientes suposiciones:

El proceso que va a predecirse se genera filtrando ruido blanco gaussiano.

El mejor predictor es el que minimiza el valor cuadrático medio del error de predicción.

Un predictor admisible de m pasos para $y(k+m)$ es una función arbitraria de $y(k)$, $y(k-1)$, ...

En primer lugar se da una deducción intuitiva de un predictor. Después se formalizará el resultado.

Heurística

Consideremos la señal y generada por el modelo

$$y(k) = \frac{C(q)}{A(q)} e(k) = \frac{C^*(q^{-1})}{A^*(q^{-1})} e(k) \quad (12.14)$$

donde A^* y C^* son las recíprocas de A y C , y q^{-1} es el operador de desplazamiento hacia atrás. Es conveniente introducir este operador porque la exposición se basa en la causalidad. Se supone que A y C son de orden n .

Consideremos la situación en el instante k . Se han observado las variables $y(k)$, $y(k-1)$, ... y se desea predecir $y(k+m)$. Un desarrollo formal en serie de C^*/A^* en q^{-1} produce

$$\begin{aligned} y(k+m) &= \frac{C^*(q^{-1})}{A^*(q^{-1})} e(k+m) \\ &= \{e(k+m) + f_1 e(k+m-1) + \cdots + f_{m-1} e(k+1)\} \\ &\quad + [f_m e(k) + f_{m+1} e(k-1) + \cdots] \end{aligned} \quad (12.15)$$

Los términos de la parte derecha son todos independientes porque $\{e(k)\}$ es una secuencia de variables aleatorias independientes. Se deduce del modelo (12.14) que si el polinomio C es estable, entonces $e(i)$ puede calcularse exactamente a partir de $y(i)$, $y(i-1)$, ... Los términos dentro de los corchetes son funciones conocidas de los datos disponibles en el instante k . Los términos entre llaves son independientes de los datos en el instante k . De este modo se deduce que el predictor óptimo está dado por

$$\hat{y}(k+m|k) = f_m e(k) + f_{m+1} e(k-1) + f_{m+2} e(k-2) + \cdots$$

y que el error de predicción es

$$\tilde{y}(k+m|k) = e(k+m) + f_1 e(k+m-1) + \cdots + f_{m-1} e(k+1)$$

Para proporcionar una demostración formal, falta mostrar cómo pueden calcularse los números f_i a partir de A y C y cómo puede expresarse $e(k)$ en función de los datos anteriores.

Resultado fundamental

El resultado fundamental puede establecerse del siguiente modo.

Teorema 12.1 – Predicción óptima. Sea $\{y(k)\}$ un proceso aleatorio generado por el modelo de (12.14), donde todos los ceros del polinomio $C(z)$ están dentro del círculo unidad y $\{e(k)\}$ es una secuencia de variables aleatorias independientes. El predictor de varianza mínima sobre m pasos es

$$\hat{y} = \hat{y}(k+m|k) = \frac{qG(q)}{C(q)}y(k) = \frac{G^*(q^{-1})}{C^*(q^{-1})}y(k) \quad (12.16)$$

donde los polinomios F y G son el cociente y el resto obtenidos al dividir $q^{m-1}C$ entre A ; es decir,

$$q^{m-1}C(q) = A(q)F(q) + G(q) \quad (12.17)$$

El error de predicción es de media móvil

$$\tilde{y}(k+m|k) = y(k+m) - \hat{y}(k+m|k) = F(q)e(k+1) \quad (12.18)$$

Tiene media cero y varianza

$$E\tilde{y}(k+m|k)^2 = [1 + f_1^2 + \dots + f_{m-1}^2]\sigma^2 \quad (12.19)$$

Demostración. El polinomio F es de grado $m-1$ y G es de menor grado que n . Por tanto,

$$F(q) = q^{m-1} + f_1q^{m-2} + \dots + f_{m-1}$$

y

$$G(q) = g_0q^{n-1} + g_1q^{n-2} + \dots + g_{n-1}$$

Introduzcamos

$$F^*(q^{-1}) = 1 + f_1q^{-1} + \dots + f_{m-1}q^{-m+1}$$

y

$$G^*(q^{-1}) = g_0 + g_1q^{-1} + \dots + g_{n-1}q^{-n+1}$$

Se deduce de (12.17) que

$$C^*(q^{-1}) = A^*(q^{-1})F^*(q^{-1}) + q^{-m}G^*(q^{-1}) \quad (12.20)$$

Entonces la ecuación (12.15) puede escribirse como

$$y(k+m) = \frac{C^*(q^{-1})}{A^*(q^{-1})}e(k+m) = F^*(q^{-1})e(k+m) + \frac{G^*(q^{-1})}{A^*(q^{-1})}e(k)$$

Utilizando la ecuación (12.14) la señal e del último término puede expresarse en función de los datos disponibles en el instante k . Por tanto,

$$y(k+m) = F^*(q^{-1})e(k+m) + \frac{G^*(q^{-1})}{C^*(q^{-1})}y(k) \quad (12.21)$$

El primer término de la parte derecha es una función lineal de $e(k+1)$, $e(k+2)$, ..., $e(k+m)$, los cuales son independientes de los datos $y(k)$, $y(k-1)$, $y(k-2)$, ... disponibles en los instantes k . El último término es una función lineal de los datos. Sea \hat{y} una función arbitraria de $y(k)$, $y(k+1)$, ... Entonces

$$\begin{aligned} E[y(k+m) - \hat{y}]^2 &= E[F^*(q^{-1})e(k+m)]^2 + E\left[\frac{G^*(q^{-1})}{C^*(q^{-1})}y(k) - \hat{y}\right]^2 \\ &\quad + 2E[F^*(q^{-1})e(k+m)]\left[\frac{G^*(q^{-1})}{C^*(q^{-1})}y(k) - \hat{y}\right] \end{aligned} \quad (12.22)$$

El último término es cero porque $e(k+m)$, $e(k+m-1)$, ..., $e(k+1)$ tienen valores medios nulos y son independientes de $y(k)$, $y(k-1)$, ... El predictor que minimiza el valor cuadrático medio del error de predicción está definido por (12.16) y el error de predicción por (12.18). La demostración se completa tomando el valor medio de los cuadrados de los errores de predicción (12.18). Con esto se obtiene (12.19). \square

Nota 1. Obsérvese que el mejor predictor es lineal. La linealidad no depende críticamente del criterio de varianza mínima. Si la densidad de probabilidad de $\{y(k)\}$ es simétrica, el predictor (12.16) es óptimo para todos los criterios de la forma $Eg[(y(k+m) - \hat{y})^2]$ para g simétrica.

Nota 2. La suposición de que $e(i)$ y $e(j)$ son independientes para $i \neq j$ es esencial para que el último término de (12.22) sea nulo. Si las variables no están correlacionadas, el término será nulo si se exige que el predictor \hat{y} sea lineal.

Nota 3. Se deduce de (12.18) que

$$\tilde{y}(k+1|k) = y(k+1) - \hat{y}(k+1|k) = e(k+1)$$

Las variables aleatorias $\{e(k)\}$ pueden interpretarse como las innovaciones del proceso $\{y(k)\}$ (compárese con la Sección 6.5).

Nota 4. Obsérvese que la función

$$J(m) = \sigma^2[1 + f_1^2 + \dots + f_{m-1}^2]$$

es la varianza del error de predicción sobre el intervalo de tiempo mh . La función $J(m)$ se acerca a la varianza de y cuando $m \rightarrow \infty$. Una gráfica de la función J muestra cómo puede predecirse el proceso sobre diferentes horizontes.

Nota 5. El predictor expuesto en esta sección es equivalente al predictor en estado estacionario obtenido utilizando el filtro de Kalman en la Sección 11.3 (véase el Ejemplo 11.5).



Cálculo del predictor óptimo

Los polinomios F y G se pueden obtener por división de polinomios. También se puede dar una fórmula explícita para los coeficientes de los polinomios. Igualando en (12.17) los coeficientes de potencias iguales de q se obtienen las siguientes ecuaciones:

$$\begin{aligned} c_1 &= a_1 + f_1 \\ c_2 &= a_2 + a_1 f_1 + f_2 \\ &\vdots \\ c_{m-1} &= a_{m-1} + a_{m-2} f_1 + \cdots + a_1 f_{m-2} + f_{m-1} \\ c_m &= a_m + a_{m-1} f_1 + \cdots + a_1 f_{m-1} + g_0 \\ c_{m+1} &= a_{m+1} + a_m f_1 + \cdots + a_2 f_{m-1} + g_1 \\ &\vdots \\ c_n &= a_n + a_{n-1} f_1 + \cdots + a_{n-m+1} f_{m-1} + g_{n-m} \\ 0 &= a_n f_1 + a_{n-1} f_2 + \cdots + a_{n-m+2} f_{m-1} + g_{n-m+1} \\ &\vdots \\ 0 &= a_n f_{m-1} + g_{n-1} \end{aligned} \tag{12.23}$$

Estas ecuaciones pueden resolverse fácilmente por recursión.

Caso en el que C tiene los ceros sobre la frontera del círculo unidad

El predictor de (12.16) es un sistema dinámico con el polinomio característico $C(z)$. Así, la suposición de que C tiene todos sus ceros dentro del círculo unidad garantiza que el predictor es estable en estado estacionario. Las condiciones iniciales son irrelevantes porque su influencia decrece exponencialmente. De la factorizaciónpectral se deduce que C puede elegirse con sus ceros dentro del círculo unidad o sobre la frontera. Por tanto, falta discutir el caso en que C tiene sus ceros sobre la frontera del círculo unidad.

Ejemplo 12.2

Consideremos el proceso

$$y(k) = e(k) - e(k-1) \tag{12.24}$$

En este caso el polinomio $C(z)=z-1$ tiene un cero sobre la frontera del círculo unidad. Aplicando el método anterior se obtiene el predictor de un paso

$$\hat{y}(k+1|k) = -e(k)$$

Calculando $e(k)$ a partir de $y(k), y(k-1), \dots, y(k_0)$ como se hizo anteriormente se obtiene

$$e(k) = e(k_0 - 1) + \sum_{i=k_0}^k y(i) = e(k_0 - 1) + z(k) \tag{12.25}$$

La presencia del término $e(k_0 - 1)$, el cual no tiende a cero cuando $k_0 \rightarrow -\infty$, muestra las consecuencias de que C sea inestable. Sin embargo, puede utilizarse la teoría del filtro de Kalman para determinar el predictor óptimo. La señal proporcionada por (12.24) puede escribirse como

$$\begin{aligned} x(k+1) &= e(k) \\ y(k) &= -x(k) + e(k) \end{aligned}$$

donde $R_1 = R_2 = R_{12} = \sigma^2$ con la notación utilizada en la Sección 11.3. El filtro de Kalman es

$$\hat{x}(k+1|k) = K(k)[y(k) + \hat{x}(k|k-1)]$$

$$P(k+1) = \frac{\sigma^2 P(k)}{P(k) + \sigma^2}$$

$$K(k) = \frac{\sigma^2}{P(k) + \sigma^2}$$

con las condiciones iniciales

$$\hat{x}(k_0|k_0-1) = 0$$

$$P(k_0) = \sigma^2$$

El predictor para la salida es

$$\hat{y}(k+1|k) = -\hat{x}(k+1|k) = -K(k)[y(k) - \hat{y}(k|k-1)]$$

Con cálculos simples se obtiene

$$\hat{y}(k+1|k) = -\frac{1}{k-k_0+2} \sum_{n=0}^{k-k_0} (n+1)y(k_0+n)$$

Por tanto, el predictor óptimo es un sistema variable con el tiempo. Obsérvese que la influencia de la condición inicial $y(k_0)$ tiende a cero según $1/(k+2-k_0)$. Es mucho más lento que el caso en que C es estable.

Del ejemplo se deduce que el predictor óptimo es un *sistema variable con el tiempo* si el polinomio C tiene los ceros sobre la frontera del círculo unidad. Estos modelos deben evitarse si se desean predictores invariantes. Desafortunadamente, este hecho no siempre se tiene en cuenta como ilustra el Ejemplo 12.3.

Ejemplo 12.3 – Cómo modelar errores permanentes

El modelo

$$A(q)y(k) = C(q)e(k) + b$$

donde b es una constante desconocida, representa una señal con una desviación. La constante b puede eliminarse tomando diferencias. Por tanto,

$$(q-1)A(q)y(k) = (q-1)C(q)e(k)$$



El factor $q - 1$ puede eliminarse considerando $\Delta y(k) = (q - 1)y(k)$ como la salida. Entonces se obtiene el modelo

$$A(q)\Delta y(k) = (q - 1)C(q)e(k) = \tilde{C}(q)e(k)$$

En este modelo el polinomio \tilde{C} aparentemente tiene un cero en el círculo unidad. Sin embargo, este modelo *no* es muy deseable porque el predictor óptimo es un sistema variable en el tiempo. Es mucho mejor modelar un error permanente como un proceso de Wiener. Esto conduce a un modelo del proceso con $A(1)=0$ que es inestable con un predictor estacionario.

En la Sección 12.6 se dan otras razones para evitar los modelos donde el polinomio $C(z)$ tiene ceros próximos al círculo unidad.

12.4. Control de varianza mínima

Para determinar la ley de control de varianza mínima se estudiará en primer lugar el caso especial en que el polinomio B en (12.5) es estable. Esto significa que la dinámica del proceso tiene una inversa estable. Por abuso del lenguaje, este caso también se denomina caso de fase mínima porque la función de transferencia discreta tiene todos sus ceros dentro del círculo unidad. La solución al problema del control es muy simple en este caso especial. La solución también aclara las propiedades del problema de control.

Sistemas con inversos estables

Introduciendo el operador desplazamiento hacia atrás q^{-1} , el modelo (12.25) puede escribirse como

$$\begin{aligned} y(k) &= \frac{B(q)}{A(q)}u(k) + \frac{C(q)}{A(q)}e(k) \\ &= \frac{B^*(q^{-1})}{A^*(q^{-1})}q^{-d}u(k) + \frac{C^*(q^{-1})}{A^*(q^{-1})}e(k) \end{aligned} \quad (12.26)$$

donde

$$d = \text{gra } A - \text{gra } B$$

es el *exceso de polos* del sistema (véase Sección 3.4). Además, $\text{gra } A = \text{gra } C = n$. Se introducen los polinomios recíprocos para hacer más transparente la discusión basada en argumentos de causalidad.

Se deduce de (12.26) que

$$\begin{aligned} y(k+d) &= \frac{C^*(q^{-1})}{A^*(q^{-1})}e(k+d) + \frac{B^*(q^{-1})}{A^*(q^{-1})}u(k) \\ &= F^*(q^{-1})e(k+d) + \frac{G^*(q^{-1})}{A^*(q^{-1})}e(k) + \frac{B^*(q^{-1})}{A^*(q^{-1})}u(k) \end{aligned} \quad (12.27)$$

donde se ha utilizado la ecuación (12.17) con $m=d$ para obtener la última igualdad. El primer término del lado derecho es independiente de los datos disponibles en el instante k y por tanto también del segundo y tercer término. El segundo término puede calcularse exactamente a partir de los datos disponibles en el instante k . Para realizar esto, se calcula la variable e a partir de (12.26); es decir,

$$e(k) = \frac{A^*}{C^*}y(k) - q^{-d}\frac{B^*}{C^*}u(k)$$

donde se ha quitado el argumento para simplificar la escritura. Utilizando esta expresión para e , la ecuación (12.27) puede escribirse como

$$\begin{aligned} y(k+d) &= F^*e(k+d) + \frac{G^*}{C^*}y(k) - q^{-d}\frac{B^*G^*}{A^*C^*}u(k) + \frac{B^*}{A^*}u(k) \\ &= F^*e(k+d) + \frac{G^*}{C^*}y(k) + \frac{B^*F^*}{C^*}u(k) \end{aligned} \quad (12.28)$$

Sea ahora $u(k)$ una función arbitraria de $y(k), y(k-1), \dots$ y $u(k-1), u(k-2), \dots$ Entonces

$$Ey^2(k+d) = E[F^*e(k+d)]^2 + E\left[\frac{G^*}{C^*}y(k) + \frac{B^*F^*}{C^*}u(k)\right]^2 \quad (12.29)$$

Los términos mezclados se anulan porque $e(k+d), \dots, e(k+1)$ son independientes de $y(k), y(k-1), \dots$ y $u(k), u(k-1), \dots$ Puesto que el último término de (12.29) no es negativo, se deduce que

$$Ey^2(k+d) \geq [1 + f_1^2 + \dots + f_{d-1}^2]\sigma^2 \quad (12.30)$$

donde la igualdad se obtiene para

$$u(k) = -\frac{G^*(q^{-1})}{B^*(q^{-1})F^*(q^{-1})}y(k) = -\frac{G(q)}{B(q)F(q)}y(k) \quad (12.31)$$

la cual es la ley de control de varianza mínima deseada. El resultado puede resumirse del siguiente modo.

Teorema 12.2. Considérese un proceso descrito por (12.5), donde $\{e(k)\}$ es una secuencia de variables aleatorias independientes de valor medio cero y desviación estándar σ . Consideremos los polinomios B y C con todos sus ceros dentro del círculo unidad. Entonces, la ley de control de varianza mínima está dada por (12.31), donde los polinomios F^* y G^* están definidos por (12.17). Esta ley de control produce la salida

$$\begin{aligned} y(k) &= F^*(q^{-1})e(k) \\ &= e(k) + f_1e(k-1) + \dots + f_{d-1}e(k-d+1) \end{aligned} \quad (12.32)$$

en estado estacionario. \square



Nota 1. El teorema se mantiene cuando $e(i)$ y $e(j)$ no están correlacionadas para $i \neq j$ si se postula una ley de control lineal.

Nota 2. El resultado está estrechamente relacionado con la solución del problema de predicción (Teorema 12.1). La identidad (12.17) se utilizó en ambos casos. Los dos últimos términos de (12.28) pueden interpretarse como la predicción de la salida con d pasos de antelación. De este modo, la estrategia de varianza mínima se obtiene prediciendo la salida d pasos más adelante y eligiendo un control que haga la predicción igual a la salida deseada. Por tanto, el problema del control estocástico puede separarse en dos problemas, un problema de predicción estocástica y un problema de control determinístico. El Teorema 12.2 se denomina, por tanto, *teorema de separación*.

Nota 3. El error bajo control de varianza mínima es de media móvil de orden $d-1$. De este modo la función de covarianza del error de regulación será nula para argumentos superiores a $d-1$. Este hecho puede utilizarse para diagnosticar si debe utilizarse una estrategia de varianza mínima.

Nota 4. Cuando se utiliza la ley de control (12.31) se cancelan todos los ceros del proceso. La consecuencia de esta propiedad se discutirá más adelante.

Es muy fácil calcular la ley de control de varianza mínima para un modelo dado (12.5), como se ilustra en el Ejemplo 12.4.

Ejemplo 12.4

Consideremos un sistema definido por (12.5), donde

$$A(q) = q^3 - 1.7q^2 + 0.7q$$

$$B(q) = q + 0.5$$

$$C(q) = q^3 - 0.9q^2$$

El exceso de polos es $d=2$. Dividiendo $q^{d-1}C(q)$ entre $A(q)$ se obtiene el cociente

$$F(q) = q + 0.8$$

y el resto

$$G(q) = 0.66q^2 - 0.56q$$

De este modo, la ley de control de varianza mínima es

$$u(k) = -\frac{q(0.66q - 0.56)}{(q + 0.5)(q + 0.8)}y(k)$$

Interpretación como diseño por asignación de polos

La ley de control de varianza mínima puede interpretarse mediante el diseño por asignación de polos expuesto en el Capítulo 10. Para ver la relación, analizaremos el

sistema en lazo cerrado que se obtiene cuando se aplica al sistema de (12.5) la ley de control (12.31). Las ecuaciones (12.5) y (12.31) pueden escribirse como

$$\begin{bmatrix} A(q) & -B(q) \\ G(q) & F(q)B(q) \end{bmatrix} \begin{bmatrix} y(k) \\ u(k) \end{bmatrix} = \begin{bmatrix} C(q) \\ 0 \end{bmatrix} e(k) \quad (12.33)$$

La ecuación característica del sistema en lazo cerrado se obtiene igualando a cero el determinante de la matriz del lado izquierdo de (12.33). Por tanto,

$$\begin{aligned} A(q)F(q)B(q) + G(q)B(q) &= B(q)[A(q)F(q) + G(q)] \\ &= q^{d-1}B(q)C(q) = 0 \end{aligned} \quad (12.34)$$

donde la ecuación (12.17), con $m=d$, se utiliza para obtener la segunda igualdad. Por tanto, el sistema en lazo cerrado tiene $2n-d$ polos en los ceros de B y C , y $d-1$ polos adicionales en el origen.

La estrategia de control de varianza mínima puede interpretarse como un diseño por asignación de polos, donde los polos se sitúan en los ceros de (12.34). La similitud con la asignación de polos puede verse con mayor claridad si se escribe la ley de control (12.31) como

$$u(k) = -\frac{S(q)}{R(q)}y(k)$$

donde $S=G$ y $R=FB$ [compárese con la ecuación (10.4)]. Multiplicando (12.17) por B se obtiene

$$\begin{aligned} q^{d-1}C(q)B(q) &= A(q)F(q)B(q) + G(q)B(q) \\ &= A(q)R(q) + B(q)S(q) \end{aligned}$$

Esta ecuación es un caso especial de la ecuación diofántica (10.22) cuando $B^+=B$ y con $A_m=q^{d-1}$ y $A_o=C$.

Sistemas con inversos inestables

La nota 4 del Teorema 12.2 menciona que en la ley de control dada por el Teorema 12.2 se cancelan todos los ceros del proceso. Si hay ceros del proceso fuera del círculo unidad, el sistema en lazo cerrado tendrá modos inestables que son inobservables desde la salida. En primer lugar discutiremos la implicaciones de este fenómeno. A continuación se presentarán otras leyes de control que no requieren que todos los ceros de $B(z)$ estén dentro del círculo unidad.

Resolviendo la ecuación (12.33) para y y u se obtiene

$$y(k) = \frac{F(q)}{q^{d-1}}e(k)$$

$$u(k) = -\frac{G(q)}{q^{d-1}B(q)}e(k)$$

La necesidad de suponer que B sea estable se ve claramente en estas ecuaciones. Si el polinomio B es inestable, el sistema tiene modos inestables, los cuales se excitan con las perturbaciones. Estos modos inestables están acoplados a la señal del control con lo que ésta crece exponencialmente. Sin embargo, la señal de salida permanece acotada porque los modos inestables no están acoplados con la salida. Un ejemplo ilustra lo que ocurre.

Ejemplo 12.5

Consideremos un sistema descrito por los polinomios

$$A(z) = (z - 1)(z - 0.7)$$

$$B(z) = 0.9z + 1$$

$$C(z) = z(z - 0.7)$$

El polinomio $B(z)$ tiene un cero en $z = -10/9$, el cual está fuera del círculo unidad. En la Figura 12.1 se muestra una simulación del controlador de varianza mínima. La presencia de los modos inestables se ve claramente en la señal de control, aunque no es importante en la salida del sistema. Esto se denomina rizado. Si se continúa la simulación, la señal de control finalmente será tan grande que se producirán desbordamientos o errores numéricos. En un problema práctico la señal se hará rápidamente tan grande que la aproximación lineal ya no será válida. Después de un corto período de tiempo los modos inestables se observarán en la salida.

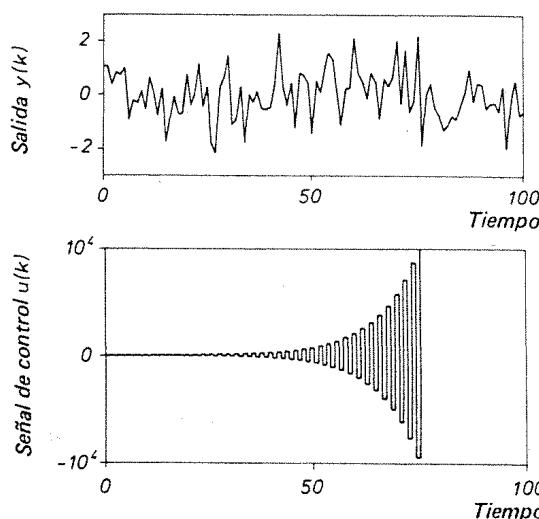


Figura 12.1 Simulación del sistema del Ejemplo 12.5 con la ley de control dada por el Teorema 12.2 que cancela un cero inestable del proceso.

En el Teorema 12.3 se extiende la ley de control de varianza mínima al caso en que el polinomio B tenga ceros fuera del círculo unidad.

Teorema 12.3. Consideremos un sistema descrito por (12.5). Factoricemos el polinomio $B(z)$ como

$$B(z) = B^+(z)B^-(z) \quad (12.35)$$

Todos los ceros del polinomio $B^+(z)$ están dentro del círculo unidad y todos los ceros de $B^-(z)$ están fuera del círculo unidad o sobre la frontera. Supongamos que todos los ceros del polinomio C están dentro del círculo unidad y que los polinomios $A(z)$ y $B^-(z)$ no tienen factores comunes. Entonces, la ley de control de varianza mínima está dada por

$$u(k) = -\frac{G(q)}{B^+(q)F(q)}y(k) \quad (12.36)$$

donde $F(z)$ y $G(z)$ son polinomios que satisfacen la ecuación diofántica

$$q^{d-1}C(q)B^{-*}(q) = A(q)F(q) + B^-(q)G(q) \quad (12.37)$$

en la cual el polinomio $F(z)$ tiene el grado $d + \text{gra } B^- - 1$.

*Demuestra*ón. En la demostración se utiliza un truco ideado por Wiener. Consideremos el operador

$$\frac{1}{q+a}$$

donde $|a| > 1$. Este operador se interpreta normalmente como un operador causal inestable (no acotado). Como $|a| > 1$, y el operador desplazamiento tiene norma $\|q\| = 1$, el desarrollo en serie

$$\frac{1}{q+a} = \frac{1}{a} \frac{1}{1+q/a} = \frac{1}{a} \left[1 - \frac{q}{a} + \frac{q^2}{a^2} - \dots \right]$$

converge. Así, el operador $(q+a)^{-1}$ puede interpretarse como un operador estable no causal; es decir,

$$\frac{1}{q+a}y(k) = \frac{1}{a} \left[y(k) - \frac{1}{a}y(k+1) + \frac{1}{a^2}y(k+2) - \dots \right]$$

Con esta interpretación, se deduce que

$$(q+a) \left[\frac{1}{q+a}y(k) \right] = y(k)$$

Los cálculos requeridos en la demostración se realizan convenientemente utilizando el operador desplazamiento hacia atrás. Se deduce del modelo del proceso de (12.5) que

$$y(k+d) = \frac{B^*(q^{-1})}{A^*(q^{-1})}u(k) + \frac{C^*(q^{-1})}{A^*(q^{-1})}e(k+d)$$



Introduzcamos

$$w(k) = \frac{B^-(q^{-1})}{B^{-*}(q^{-1})} y(k)$$

donde el operador $1/B^{-*}(q^{-1})$ se interpreta como un operador estable no causal. Las señales y y w tienen la misma varianza en estado estacionario porque B^- y B^{-*} son polinomios reciprocos y

$$\left| \frac{B^-(e^{i\omega})}{B^{-*}(e^{i\omega})} \right| = 1$$

Una ley de control admisible que minimiza la varianza de w también minimiza la varianza de y . Con esto se deduce que

$$w(k+d) = \frac{B^{+*}(q^{-1})B^-(q^{-1})}{A^*(q^{-1})} u(k) + \frac{C^*(q^{-1})B^-(q^{-1})}{A^*(q^{-1})B^{-*}(q^{-1})} e(k+d) \quad (12.38)$$

La suposición de que $A(z)$ y $B^-(z)$ no tienen factores comunes garantiza que (12.37) tiene solución.

La ecuación (12.37) implica que

$$C^*(q^{-1})B^-(q^{-1}) = A^*(q^{-1})F^*(q^{-1}) + q^{-d}B^{-*}(q^{-1})G^*(q^{-1})$$

Dividiendo por A^*B^{-*} se obtiene

$$\frac{C^*(q^{-1})B^-(q^{-1})}{A^*(q^{-1})B^{-*}(q^{-1})} = \frac{F^*(q^{-1})}{B^{-*}(q^{-1})} + q^{-d} \frac{G^*(q^{-1})}{A^*(q^{-1})}$$

Utilizando esta ecuación, el último término de (12.38) puede escribirse como

$$w(k+d) = \frac{F^*(q^{-1})}{B^{-*}(q^{-1})} e(k+d) + \frac{B^{+*}(q^{-1})B^-(q^{-1})}{A^*(q^{-1})} u(k) + \frac{G^*(q^{-1})}{A^*(q^{-1})} e(k) \quad (12.39)$$

Puesto que el operador $1/B^{-*}(q^{-1})$ se interpreta como un operador acotado no causal y puesto que $\text{gra } F^* = d + \text{gra } B^- - 1$, se deduce que

$$\frac{F^*(q^{-1})}{B^{-*}(q^{-1})} e(k+d) = \alpha_1 e(k+1) + \alpha_2 e(k+2) + \dots$$

Todos estos términos son independientes de los dos últimos términos de (12.39). Utilizando los argumentos que se dieron con detalle en la demostración del Teorema 12.1, encontramos que la ley de control óptimo se obtiene anulando la suma de los dos últimos términos de (12.39). Esto da

$$u(k) = -\frac{G^*(q^{-1})}{B^{+*}(q^{-1})B^-(q^{-1})} e(k) \quad (12.40)$$

y

$$y(k) = \frac{B^-(q^{-1})}{B^-(q^{-1})} w(k) = \frac{F^*(q^{-1})}{B^-(q^{-1})} e(k) = \frac{F(q)}{q^{d-1}B^{-*}(q)} e(k) \quad (12.41)$$

Eliminando $e(k)$ entre (12.40) y (12.41) se obtiene

$$u(k) = -\frac{G^*(q^{-1})}{B^{+*}(q^{-1})F^*(q^{-1})} y(k) \quad (12.42)$$

El numerador y denominador tienen el mismo grado, y la ley de control puede reescribirse como (12.36). \square

Nota 1. Los ceros estables del proceso se cancelan con la ley de control óptima.

Nota 2. De las demostraciones de los Teoremas 12.2 y 12.3 se deduce que la varianza de la salida de un sistema como el (12.5) puede tener varios mínimos locales si el polinomio $B(z)$ tiene ceros fuera del círculo unidad. Hay un mínimo absoluto definido por el Teorema 12.2. Sin embargo, este mínimo dará señales de control infinitamente grandes. El mínimo local dado por el Teorema 12.3 es el más grande de los mínimos locales. En este caso la señal de control está acotada.

Nota 3. La factorización de (12.35) es arbitraria porque B^+ podría multiplicarse por un mínimo y B^- podría dividirse por el mismo número. Es conveniente seleccionar los factores para que el polinomio $B^{-*}(q)$ sea mónico.

Ejemplo 12.6

Consideremos el Ejemplo 12.5 donde $d=1$ y

$$B^+(z) = 1$$

$$B^-(z) = B(z)$$

$$B^{-*}(z) = z + 0.9$$

La ecuación (12.37) se convierte en

$$z(z - 0.7)(z + 0.9) = (z - 1)(z - 0.7)(z + f_1) + (0.9z + 1)(g_0 z + g_1)$$

Sea $z=0.7$, $z=1$ y $z=-\frac{10}{9}$. Esto da

$$0.7g_0 + g_1 = 0$$

$$g_0 + g_1 = 0.3$$

$$f_1 = 1$$

Entonces, la ley de control es

$$u(k) = -\frac{G(q)}{B^+(q)F(q)} y(k) = -\frac{q - 0.7}{q + 1} y(k)$$

La salida es

$$\begin{aligned} y(k) &= \frac{F(q)}{B^{-*}(q)} e(k+d-1) = \frac{q+1}{q+0.9} e(k) \\ &= e(k) + \frac{0.1}{q+0.9} e(k) \end{aligned}$$

La varianza de la salida es

$$Ey^2 = \left[1 + \frac{0.1^2}{1 - 0.9^2} \right] \sigma^2 = \frac{20}{19} \sigma^2$$

la cual es aproximadamente un 5% mayor que cuando se utiliza el controlador del Ejemplo 12.4. La Figura 12.2 muestra una simulación de la ley de control. Se observa en la figura que el regulador funciona correctamente. Compárese también con la Figura 12.1, la cual muestra el efecto de la cancelación de los ceros inestables.

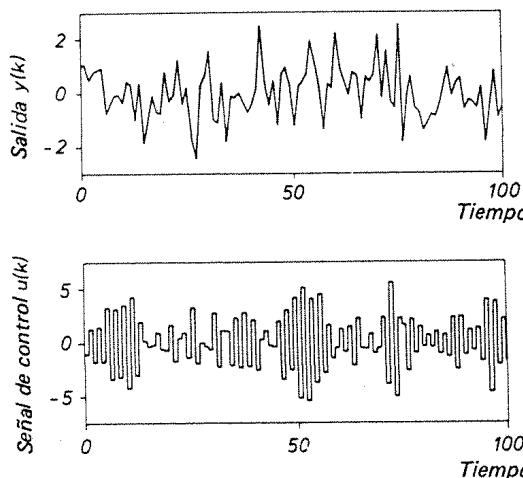


Figura 12.2 Simulación del sistema del Ejemplo 12.6.

Una interpretación como asignación de polos

Cálculos simples muestran que la ecuación característica del sistema en lazo cerrado obtenido con (12.5) y (12.36) es

$$z^{d-1} B^+(z) B^{-*}(z) C(z) = 0 \quad (12.43)$$

De este modo, la ley de control de (12.36) puede interpretarse como un controlador de asignación de polos, el cual satisface esta ecuación característica.

Multiplicando (12.37) por B^+ se obtiene la ecuación

$$A(z)R(z) + B(z)S(z) = z^{d-1} B^+(z) B^{-*}(z) C(z) \quad (12.44)$$

donde $R(z) = B^+(z)F(z)$ y $S(z) = G(z)$. Esta ecuación es la misma ecuación diofántica que se utilizó en el diseño por asignación de polos [compárese con la ecuación (10.22)]. El sistema en lazo cerrado tiene los polos correspondientes al observador dinámico, a los ceros estables del proceso y a las reflexiones en el círculo unidad de los

ceros inestables del proceso. Obsérvese que la función de transferencia $B(z)/A(z)$ puede interpretarse como si se tuvieran $d = \text{gra } A - \text{gra } B$ ceros en el infinito. Las reflexiones de estos ceros en el círculo unidad también aparecen como polos en lazo cerrado.

12.5. Control LQG

Vamos a resolver el problema del control óptimo para el sistema de (12.5) con el criterio de (12.8). La ley de control de varianza mínima expuesta en la Sección 12.4 puede expresarse en función de la solución de una ecuación polinómica. La solución del problema LQG puede obtenerse de un modo similar. Sin embargo, se necesitan dos ecuaciones polinómicas. Estas ecuaciones se tratarán antes de presentar el resultado fundamental.

Factorización espectral

El problema LQ se resolvió en la Sección 11.2 utilizando la representación del espacio de estados, el cual conduce a una ecuación de Riccati en estado estacionario. Se deduce de la ecuación de Riccati que

$$rP(z)P(z^{-1}) = \rho A(z)A(z^{-1}) + B(z)B(z^{-1}) \quad (12.45)$$

donde el polinomio mónico $P(z)$ es el polinomio característico del sistema en lazo cerrado. Este resultado se dedujo en la Sección 11.2 [véase la ecuación (11.38)]. El polinomio característico en lazo cerrado puede obtenerse resolviendo una ecuación de Riccati en estado estacionario. Una alternativa consiste en encontrar un polinomio $P(z)$ que satisfaga (12.45) directamente. Una realimentación que proporcione los polos deseados en lazo cerrado puede determinarse por asignación de polos. Al problema de encontrar un polinomio $P(z)$ que satisfaga (12.45) se le denomina *factorización espectral*.

Consideremos en primer lugar un polinomio de la forma

$$F(z) = f_0 z^{2n} + f_1 z^{2n-1} + \cdots + f_{n-1} z^{n+1} + f_n z^n + f_{n-1} z^{n-1} + \cdots + f_1 z + f_0$$

Tal polinomio es auto-recíproco porque

$$F^*(z) = z^{2n} F(z^{-1}) = F(z)$$

De este modo se deduce que si $z = a$ es un cero de $F(z)$, entonces $z = 1/a$ también es un cero. Además, si los coeficientes f_i son reales, entonces $z = \bar{a}$ y $z = 1/\bar{a}$ también son ceros. Ahora puede establecerse el siguiente resultado.

Lema 12.1. Sean los polinomios $A(z)$ y $B(z)$ primos entre sí con $\text{gra } A(z) > \text{gra } B(z)$. Entonces existe un polinomio único $P(z)$ con $\text{gra } P(z) = \text{gra } A(z) = n$ y todos sus ceros dentro o sobre la frontera del círculo unidad tal que cumple (12.45). Si $\rho > 0$, entonces $P(z)$ no tiene ceros sobre la frontera del círculo unidad.



Demostración. Si el lado derecho de (12.45) se multiplica por z^n se obtiene un polinomio auto-recíproco. De este modo, los ceros del lado derecho son imágenes especulares respecto al círculo unidad. Como los coeficientes son reales, los ceros también son simétricos respecto al eje real. El lado derecho de (12.45) no puede tener ceros sobre el círculo unidad porque si $z = e^{i\omega}$ es cero, entonces

$$\rho A(e^{i\omega})A(e^{-i\omega}) + B(e^{i\omega})B(e^{-i\omega}) = \rho |A(e^{i\omega})|^2 + |B(e^{i\omega})|^2 = 0$$

Como $\rho > 0$, esto implica que $z = \exp(i\omega)$ es un cero de $A(z)$ y de $B(z)$, lo cual contradice la suposición de que $A(z)$ y $B(z)$ son primos entre sí. La condición $\text{gra } P(z) = n$ es necesaria para asegurar que $P(z)$ sea único porque si $P(z)$ satisface (12.45), entonces $z^k P(z)$ también satisface (12.45) para un $k \geq 0$ arbitrario. \square

Nota. Introduciendo polinomios recíprocos, la ecuación (12.45) puede escribirse como

$$rP^*(z)P(z) = \rho A(z)A^*(z) + z^d B(z)B^*(z) \quad (12.46)$$

Como $\text{gra } P(z) = \text{gra } A(z) = d + \text{gra } B(z)$ se deduce que

$$\text{gra } P^*(z) = \max [\text{gra } A^*(z), \text{gra } B^*(z)] \quad (12.47)$$

Conceptualmente el problema de factorización espectral puede resolverse encontrando los ceros del lado derecho de (12.45) y ordenándolos. También existen algoritmos recursivos eficaces para resolver el problema.

La ecuación diofántica

El polinomio característico $P(z)$ para el problema LQ se obtiene por factorización espectral. También se necesita el polinomio del observador para obtener el polinomio característico del problema LQG. En la Sección 12.3 se muestra que $C(z)$ es el polinomio del observador para el modelo (12.5). De la solución del problema de asignación de polos, puede suponerse que la ley de control óptimo está dada por

$$u(k) = -\frac{S(q)}{R(q)}y(k) \quad (12.48)$$

donde $R(z)$ y $S(z)$ son soluciones de la ecuación diofántica

$$P(z)C(z) = A(z)R(z) + B(z)S(z) \quad (12.49)$$

Esta suposición es cierta. Sin embargo, como (12.49) tiene infinitas soluciones, es necesario elegir la solución correcta. Para hacerlo resultan útiles algunas propiedades algebraicas de las ecuaciones (12.45) y (12.49).

Si pérdida de generalidad, puede suponerse que $\text{gra } A(z) = \text{gra } C(z) = \text{gra } P(z) = n$. Además, supongamos que $A(z)$ y $B(z)$ son primos entre sí con $\text{gra } A(z) > \text{gra } B(z)$. Entonces se deduce del corolario del Teorema 10.1 que hay una solución

única de (12.49) con $\text{gra } S(z) < n$ y $\text{gra } R(z) = n$. Esta solución particular corresponde a la ley de control, (12.48), con un retardo de al menos un periodo de muestreo.

Si se desea una ley de control sin retardo de cálculo, los polinomios $R(z)$ y $S(z)$ deben tener el mismo grado. Tales soluciones no son únicas. Una solución única se obtiene resolviendo el problema de control óptimo. Como preparación se desarrolla un resultado algebraico. Si $R(z)$ y $S(z)$ tienen grado n , se deduce de (12.49) que

$$A^*(z)R^*(z) + z^d B^*(z)S^*(z) = P^*(z)C^*(z)$$

donde

$$d = \text{gra } A(z) - \text{gra } B(z)$$

Se produce el siguiente resultado.

Lema 12.2. Sean los polinomios $A(z)$, $C(z)$, $P(z)$, $R(z)$ y $S(z)$ de grado n . Supongamos que los polinomios $R(z)$, $S(z)$ y $P(z)$ satisfacen (12.45) y (12.49); entonces existe un polinomio $X(z)$ tal que

$$A^*(z)X(z) + rP(z)S^*(z) = B(z)C^*(z) \quad (12.50)$$

y

$$P^*(z)X(z) + \rho A(z)S^*(z) = R^*(z)B(z) \quad (12.51)$$

Demostración. Consideremos

$$\begin{aligned} A^*(z)[R^*(z)B(z) - \rho A(z)S^*(z)] \\ &= A^*(z)R^*(z)B(z) - S^*(z)[rP(z)P^*(z) - z^d B(z)B^*(z)] \\ &= B(z)[A^*(z)R^*(z) + z^d B^*(z)S^*(z)] - rS^*(z)P(z)P^*(z) \\ &= B(z)P^*(z)C^*(z) - rS^*(z)P(z)P^*(z) \\ &= P^*(z)[B(z)C^*(z) - rS^*(z)P(z)] \end{aligned} \quad (12.52)$$

donde la primera igualdad se deduce de (12.46) y la tercera de (12.49).

En primer lugar supongamos que $A(z)$ y $B(z)$ son primos entre sí, como lo son $P(z)$ y $A(z)$. Se deduce entonces que $A^*(z)$ divide a $B(z)C^*(z) - rS^*(z)P(z)$ y que $P^*(z)$ divide a $R^*(z)B(z) - \rho A(z)S^*(z)$. Entonces existe un polinomio $X(z)$ tal que

$$B(z)C^*(z) - rS^*(z)P(z) = A^*(z)X(z)$$

y

$$R^*(z)B(z) - \rho A(z)S^*(z) = P^*(z)X(z)$$

lo cual implica (12.50) y (12.51).

Ahora considérese el caso cuando $A(z)$ y $B(z)$ tienen factores comunes. Sea $A_2(z)$ el máximo común divisor de $A(z)$ y $B(z)$. Se deduce de (12.45) que $A_2(z)$ también divide a $P(z)$. Dividiendo (12.52) por $A_2(z)$ y repitiendo el argumento anterior se obtiene el resultado. Obsérvese que en este caso $A_2(z)$ también divide a $X(z)$. \square

Resultado fundamental

La solución del problema LQG viene dada por el Teorema 12.4.

Teorema 12.4. Considérese el sistema (12.5) con $\text{gra } A(z) = \text{gra } C(z) = n$. Supongamos que todos los ceros del polinomio $C(z)$ están dentro del círculo unidad, que no hay factores comunes entre los polinomios $A(z)$, $B(z)$ y $C(z)$ y que todos los ceros de los posibles factores comunes de $A(z)$ y $B(z)$ están dentro del círculo unidad. Sea el polinomio mónico $P(z)$, el cual tiene todos sus ceros dentro del círculo unidad, la solución de (12.45), con $\text{gra } P(z) = n$. La ley de control admisible que minimiza el criterio de (12.8) está dada por

$$u(k) = -\frac{S(q)}{R(q)}y(k) \quad (12.53)$$

donde los polinomios $R(z)$ y $S(z)$ están dados por

$$\begin{aligned} S(z) &= z^{n-\text{gra } S^*} S^*(z^{-1}) \\ R(z) &= z^{n-\text{gra } R^*} R^*(z^{-1}) \end{aligned} \quad (12.54)$$

y el polinomio $S^*(z)$ es la solución única a (12.50) con $\text{gra } X(z) < n$. El polinomio $R^*(z)$ satisface (12.51) y los polinomios $R(z)$ y $S(z)$ también satisfacen (12.49).

Con la ley de control de (12.53), la salida es

$$y(k) = \frac{R(q)}{P(q)}e(k) \quad (12.55)$$

y la señal de control es

$$u(k) = -\frac{S(q)}{P(q)}e(k) \quad (12.56)$$

El valor mínimo de la función de coste es

$$\min E[y^2 + \rho u^2] = \frac{\sigma^2}{2\pi i} \oint \frac{R(z)R(z^{-1}) + \rho S(z)S(z^{-1})}{P(z)P(z^{-1})} \frac{dz}{z} \quad (12.57)$$

Demostración. Introduzcamos

$$u = v - \frac{S}{R}y \quad (12.58)$$

donde v puede considerarse como una variable de control transformada. De las ecuaciones (12.5) y (12.49) se obtiene

$$y = \frac{BRv + CRe}{AR + BS} = \frac{BRv + CRe}{PC} = \frac{BR}{PC}v + \frac{R}{P}e \quad (12.59)$$

Entonces se deduce de (12.58) que

$$u = v - \frac{SBv + SCe}{PC} = \frac{PC - BS}{PC}v - \frac{S}{P}e = \frac{AR}{PC}v - \frac{S}{P}e \quad (12.60)$$

La función de coste de (12.8) puede escribirse como

$$\begin{aligned} J &= E[y^2 + \rho u^2] = E\left[\frac{BR}{PC}v + \frac{R}{P}e\right]^2 + \rho E\left[\frac{AR}{PC}v - \frac{S}{P}e\right]^2 \\ &= J_1 + 2J_2 + J_3 \end{aligned} \quad (12.61)$$

donde

$$\begin{aligned} J_1 &= E\left[\left(\frac{BR}{PC}v\right)^2 + \rho\left(\frac{AR}{PC}v\right)^2\right] \\ J_2 &= E\left[\left(\frac{BR}{PC}v\right)\left(\frac{R}{P}e\right) - \rho\left(\frac{AR}{PC}v\right)\left(\frac{S}{P}e\right)\right] \\ J_3 &= E\left[\left(\frac{R}{P}e\right)^2 + \rho\left(\frac{S}{P}e\right)^2\right] \end{aligned}$$

Evaluemos separadamente estos términos. De la nota 3 del Teorema 6.2 y (12.45) se deduce que

$$\begin{aligned} J_1 &= E\left[\left(\frac{B(q)R(q)B(q^{-1})R(q^{-1}) + \rho A(q)R(q)A(q^{-1})R(q^{-1})}{P(q)P(q^{-1})C(q)C(q^{-1})}v\right)v\right] \\ &= E\left[\left(\frac{R(q)R(q^{-1})rP(q)P(q^{-1})}{P(q)P(q^{-1})C(q)C(q^{-1})}v\right)v\right] = E\left[\left(\frac{rR(q)R(q^{-1})}{C(q)C(q^{-1})}v\right)v\right] \\ &= E\left[r\left(\frac{R(q)}{C(q)}v\right)^2\right] \end{aligned}$$

Además,

$$\begin{aligned} J_2 &= E\left[\left(\frac{B(q)R(q)R(q^{-1}) - \rho A(q)R(q)S(q^{-1})}{P(q)C(q)P(q^{-1})}v\right)e\right] \\ &= E\left[\left(\frac{R(q)[B(q)R(q^{-1}) - \rho A(q)S(q^{-1})]}{P(q)C(q)P(q^{-1})}v\right)e\right] \end{aligned}$$

Pero del Lema 12.2 se tiene que

$$R(q)[B(q)R(q^{-1}) - \rho A(q)S(q^{-1})] = R(q)P(q^{-1})X(q)$$

Por tanto,

$$\begin{aligned} J_2 &= E\left[\left(\frac{R(q)P(q^{-1})X(q)}{P(q)C(q)P(q^{-1})}v\right)e\right] \\ &= E\left[\left(\frac{R(q)X(q)}{P(q)C(q)}v\right)\left(\frac{P(q)}{P(q)}e\right)\right] = E\left[\left(\frac{R(q)X(q)}{P(q)C(q)}v\right)e\right] \end{aligned} \quad (12.62)$$



De las suposiciones se deduce que $P(z)$ y $C(z)$ son estables. Como $X(z)$ es la solución de (12.50) con $\text{gra } X(z) < \text{gra } P(z) = n$ y $\text{gra } R(z) = \text{gra } C(z) = n$, se deduce que $R(z)X(z)$ es de grado menor que $P(z)C(z)$. De este modo el término

$$\frac{R(q)X(q)}{P(q)C(q)}v(k)$$

es una función de $v(k-1), v(k-2), \dots$ Todos estos términos son independientes de $e(k)$ y J_2 es nulo.

Por tanto, la función de coste puede escribirse como

$$J = E\left(\frac{rR}{C}v\right)^2 + E\left(\frac{R}{P}e\right)^2 + \rho E\left(\frac{S}{P}e\right)^2$$

donde P y C son polinomios estables. Se deduce que la función de coste alcanza su mínimo (12.57) para $v=0$, lo cual por (12.58) corresponde a la ley de control de (12.53). Las ecuaciones (12.55) y (12.56) con (12.59) y (12.60) y los Teoremas 6.2 y 6.4 proporcionan la expresión (12.57). \square

Nota 1. La ley de control de varianza mínima es un caso especial del Teorema 12.4, con $\rho=0$.

Nota 2. Puesto que $\text{gra } X(z) < \text{gra } P(z)$ se deduce de (12.51) que $\text{gra } R^*(z) < n$ si $\rho=0$ y $\text{gra } S^*(z) < n$ si $\rho=0$.

Nota 3. Si la ley admisible de control tiene un retardo de tiempo de un período de muestreo, la ley de control está definida por (12.53), donde $R(z)$ y $S(z)$ son la solución única de (12.49) con $\text{gra } S(z) < n$. En este caso

$$A^*(z)R^*(z) + z^{d+1}B^*(z)S^*(z) = P^*(z)C^*(z) \quad (12.63)$$

El Teorema 12.4 proporciona un método adecuado para calcular la ley de control LQ para sistemas de una entrada y una salida (SISO). El polinomio P se determina en primer lugar resolviendo el problema de factorización espectral (12.45). Si A y B tienen un factor común estable A_2 , resulta conveniente factorizar A y B como $A=A_1A_2$ y $B=B_1A_2$ y resolver el problema de factorización espectral

$$rP_1(z)P_1(z^{-1}) = \rho A_1(z)A_1(z^{-1}) + B_1(z)B_1(z^{-1}) \quad (12.64)$$

Entonces, el polinomio P es P_1A_2 . Hay varias posibilidades para calcular $R(z)$ y $S(z)$. El polinomio $S(z)$ puede determinarse de (12.50) y (12.54). Entonces el polinomio R puede obtenerse de (12.51) y (12.54) o bien de (12.49). Otro método consiste en utilizar (12.50) para determinar el grado de $S^*(z)$. Entonces, $R(z)$ y $S(z)$ pueden obtenerse de (12.49).

El teorema muestra la conexión entre los métodos de diseño LQG y de asignación de polos. El problema LQG puede interpretarse como un método de asignación de polos en el cual los polos en lazo cerrado están situados en los ceros de $P(z)$ y $C(z)$. Obsérvese que el método LQG da el polinomio del observador y también

una solución particular de la ecuación diofántica de (12.49). La elección del peso ρ se trata en el Capítulo 11.

El Teorema 12.4 se ilustra con un ejemplo.

Ejemplo 12.7

Consideremos un sistema caracterizado por

$$A(z) = z + a$$

$$B(z) = b$$

$$C(z) = z + c$$

Para encontrar la ley de control que minimiza el criterio (12.8) se resuelve en primer lugar el problema de factorización espectral. La ecuación (12.45) puede escribirse como

$$r(z + p_1)(z^{-1} + p_1) = \rho(z + a)(z^{-1} + a) + b^2$$

Igualando los coeficientes de potencias iguales de z se obtiene

$$rp_1 = \rho a$$

$$r(1 + p_1^2) = \rho(1 + a^2) + b^2$$

Con cálculos sencillos se obtiene

$$r = \frac{\rho a}{p_1}$$

y

$$p_1 = \frac{\rho(1 + a^2) + b^2 - \sqrt{\rho^2(1 - a^2)^2 + b^4 + 2\rho b^2(1 + a^2)}}{2\rho a}$$

Se elige la raíz negativa para asegurar que la raíz de $P(z)$ está dentro del círculo unidad. La ecuación (12.50) se transforma en

$$(1 + az)x_0 + r(z + p_1)s_0 = b(1 + cz)$$

Por tanto,

$$S^*(z) = s_0 = \frac{b(c - a)}{r(1 - ap_1)}$$

$$X(z) = x_0 = \frac{b(1 - cp_1)}{1 - ap_1}$$

De este modo, los polinomios $X(z)$ y $S(z)$ son de grado cero porque $P(z)=1$. El polinomio R^* está definido por (12.51). Supongamos que

$$R(z) = r_0z + r_1$$

utilizando $rp_1 = \rho a$ se obtiene

$$r_0 = 1$$

$$r_1 = \frac{cp_1}{a}$$



Otro método de solucionar el problema consiste en empezar del mismo modo y observar que $\text{gra } S^*(z)=0$. La ecuación (12.49) se transforma en

$$(z + p_1)(z + c) = (z + a)(r_0 z + r_1) + b s_0 z$$

Haciendo $z = -a$ se obtiene

$$s_0 = \frac{(p_1 - a)(a - c)}{ab}$$

utilizando (12.45) para este ejemplo con $z = -p_1$ se obtiene

$$p_1 - a = \frac{b^2 p_1}{\rho(a p_1 - 1)}$$

Y se obtiene el mismo s_0 que con el método anterior. Igualando los términos constantes de (12.49) se obtiene

$$p_1 c = a r_1$$

o

$$r_1 = \frac{p_1 c}{a}$$

y

$$r_0 = 1$$

La ley de control es

$$u(k) = -\frac{S(q)}{R(q)}y(k) = -\frac{(p_1 - a)(a - c)}{ab} \frac{q}{q + p_1 c/a} y(k)$$

Modos incontrolables e inestables

Los modelos con la propiedad de que los polinomios $A(z)$ y $B(z)$ tienen un factor común que no lo es de $C(z)$ son importantes en la práctica. Aparecen cuando hay modos que se excitan con perturbaciones incontrolables en la entrada. Como los modos no son controlables, no influye en ellos la realimentación.

El Teorema 12.4 comprende el caso de factores comunes estables, pero no actúa para factores comunes inestables. Los factores comunes inestables son importantes en la práctica porque proporcionan un modo de obtener reguladores con acción integral.

Para ver lo que ocurre cuando hay factores comunes inestables, designemos por A_2 al máximo común divisor de A y B y por A_2^- al factor de A_2 con ceros fuera del círculo unidad o sobre su frontera. Consideremos la realimentación

$$u(k) = -\frac{S(q)}{R(q)}y(k)$$

donde $R(z)$ y $S(z)$ son primos entre sí. Se deduce de (12.5) que

$$y(k) = \frac{R(q)C(q)}{A(q)R(q) + B(q)S(q)}e(k) \quad (12.65)$$

$$u(k) = -\frac{S(q)C(q)}{A(q)R(q) + B(q)S(q)}e(k) \quad (12.66)$$

El factor inestable A_2^- divide al denominador del lado derecho de (12.65) y (12.66). Tanto y como u no estarán acotadas a menos que $R(z)$ y $S(z)$ se elijan de un modo especial. La señal y estará acotada si $R(z)$ es divisible por A_2^- , y u estará acotada si A_2^- divide a $S(z)$. Como $R(z)$ y $S(z)$ son primos entre sí, no es posible mantener acotadas a y y u . Esto es normal porque son necesarias acciones de control infinitamente grandes para compensar perturbaciones infinitas.

Para describir un problema de este tipo como un problema de optimización, debe modificarse el criterio (12.8). Una posibilidad consiste en introducir la variable

$$w(k) = q^{-m}A_2^-(q)u(k) \quad (12.67)$$

donde $m = \text{gra } A_2^-(z)$, e introducir el criterio

$$J'_{1q} = E[y^2(k) + \rho w^2(k)] \quad (12.68)$$

Entonces puede establecerse el siguiente resultado.

Teorema 12.5. Consideremos el sistema de (12.5). Sean $A(z)$ y $C(z)$ de grado n . Supongamos que todos los ceros de $C(z)$ están dentro del círculo unidad y supongamos que no hay polinomios que dividen a $A(z)$, $B(z)$ y $C(z)$. Sea $A_2^-(z)$ el máximo común divisor de $A(z)$ y $B(z)$ y sea A_2^- el factor de $A_2(z)$ que tiene todos sus ceros fuera del círculo unidad o sobre la frontera. Entonces, la ley de control admisible que minimiza (12.68) está dada por (12.53), donde

$$P(z) = P_1(z)A_2(z)z^m \quad (12.69)$$

$P_1(z)$ está definido por (12.64), y $m = \text{gra } A_2^-(z)$. La solución de (12.49) deberán elegirse para que se verifiquen las siguientes condiciones:

1. $\text{gra } R(s) = \text{gra } S(z) = n + m$.
2. $A_2^-(z)$ divide a $R(z)$.
3. $\text{gra } S^*(z) < n$ si $\rho > 0$ y $\text{gra } R^*(z) < n + m$ si $\rho = 0$.

Demostración. Introduzcamos la señal w definida por (12.67). Como

$$B(z) = B_1(z)A_2^-(z)A_2^+(z) = \tilde{B}(z)z^{-m}A_2^-(z)$$

donde

$$\tilde{B}(z) = B_1(z)A_2^+(z)z^m \quad (12.70)$$

el modelo (12.5) puede escribirse como

$$A(q)y(k) = \tilde{B}(q)w(k) + C(q)e(k) \quad (12.71)$$

Los polinomios $A(q)$ y $\tilde{B}(q)$ no tienen factores comunes fuera del círculo unidad o sobre la frontera, por eso se aplica el Teorema 12.4 al sistema (12.71) con el criterio de (12.68). Entonces, la ley de control está dada por

$$w(k) = -\frac{\tilde{S}(q)}{\tilde{R}(q)}y(k) \quad (12.72)$$

donde

$$\tilde{P}(z)C(z) = A(z)\tilde{R}(z) + \tilde{B}(z)\tilde{S}(z) \quad (12.73)$$

y

$$\tilde{P}(z) = P_1(z)A_2^*(z)$$

en la cual $\text{gra } \tilde{S}^*(z) < n$ si $\rho > 0$ y $\text{gra } \tilde{R}^*(z) < n$ si $\rho = 0$. Las ecuaciones (12.67) y (12.72) dan

$$u(k) = -\frac{S(q)}{R(q)}y(k)$$

donde

$$\begin{aligned} R(z) &= A_2^*(z)\tilde{R}(z) \\ S(z) &= z^m\tilde{S}(z) \end{aligned} \quad (12.74)$$

Por tanto,

$$\begin{aligned} R^*(z) &= A_2^*(z)\tilde{R}^*(z) \\ S^*(z) &= \tilde{S}^* \end{aligned}$$

Obsérvese que $A_2^*(z)$ tiene grado m porque todos sus ceros están fuera del círculo unidad.

Multiplicando (12.73) por $A_2^*(z)$ se obtiene (12.49) porque de (12.71) y (12.47) se deduce que

$$\tilde{B}(z)A_2^*(z) = B_1(z)A_2^*(z)A_2^*(z)z^m = B(z)z^m \quad \square$$

Nota. Un modo muy apropiado de calcular la ley de control consiste en determinar en primer lugar el máximo común divisor (mcd) A_2 de A y B y calcular $A_1 = A/\text{mcd}(A, B)$. Entonces se resuelve el problema de factorización espectral de (12.64). También se resuelve la ecuación diofántica

$$P_1C = A_1A_2^*R_1 + B_1S_1 \quad (12.75)$$

para $\text{gra } R_1 = \text{gra } S_1 = n$. El polinomio $S^*(z)$ es de grado inferior a n si $\rho > 0$ y el polinomio $R^*(z)$ tiene grado menor que n si $\rho = 0$. Entonces los polinomios $R(z)$ y $S(z)$ están definidos por

$$\begin{aligned} R(z) &= A_2^*(z)R_1(z) \\ S(z) &= z^mS_1(z) \end{aligned} \quad (12.76)$$

Comparación con los métodos en el espacio de estados

Los problemas discutidos en este capítulo también pueden resolverse con los métodos en el espacio de estados del Capítulo 11. Para comparar los diferentes métodos, se da en primer lugar una representación en el espacio de estados del modelo (12.5). Para

este propósito se supone que el modelo está normalizado, esto es, que $\text{gra } C(z) = \text{gra } A(z)$. Entonces, el modelo de (12.5) puede representarse como

$$\begin{aligned} x(k+1) &= \Phi x(k) + \Gamma u(k) + K e(k) \\ y(k) &= Cx(k) + e(k) \end{aligned} \quad (12.77)$$

donde

$$\Phi = \begin{bmatrix} -a_1 & 1 & 0 & \dots & 0 \\ -a_2 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ -a_{n-1} & 0 & 0 & \dots & 1 \\ -a_n & 0 & 0 & \dots & 0 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}, \quad K = \begin{bmatrix} c_1 - a_1 \\ c_2 - a_2 \\ \vdots \\ c_n - a_n \end{bmatrix} \quad (12.78)$$

$$C = [1 \ 0 \ \dots \ 0]$$

Como ésta es una representación de innovación, el filtro de Kalman se obtiene por inspección:

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k-1) + \Gamma u(k) + K[y(k) - C\hat{x}(k|k-1)] \quad (12.79)$$

El filtro de Kalman tiene el polinomio característico

$$\det[zI - (\Phi - KC)] = C(z) \quad (12.80)$$

Las leyes de control deducidas en el Capítulo 11 suponen un retardo de cálculo de un período de muestreo. Entonces, la ley de control óptima es

$$u(k) = -L\hat{x}(k|k-1)$$

La función de transferencia del regulador óptimo es

$$H_r(z) = -L(zI - \Phi + KC + \Gamma L)^{-1}K = -\frac{S(z)}{R(z)}$$

donde $S(z)$ y $R(z)$ forman la solución a (12.49) con $\text{gra } S(z) < n$ y $\text{gra } R(z) = n$. Se deduce de esta discusión que

$$R(z) = \det(zI - \Phi + KC + \Gamma L) \quad (12.81)$$

y

$$P(z) = \det(zI - \Phi + \Gamma L) \quad (12.82)$$

Señales de mando

La exposición de este capítulo se ha limitado al problema del regulador. Para introducir las señales de mando, nos referiremos al tratamiento del Capítulo 10. La

clave del problema consiste en introducir una señal de mando de tal modo que no genere errores de reconstrucción innecesarios. Esto se consigue con la ley de control

$$R(q)u(k) = C(q)u_c(k) - S(q)y(k)$$

Entonces se deduce de (12.5) que la salida del sistema es

$$y(k) = \frac{B(q)}{P(q)}u_c(k) + \frac{R(q)}{P(q)}e(k)$$

donde $\text{gra } R = n$.

La función de transferencia discreta desde la señal de mando es $B(z)/P(z)$. Esta respuesta se puede conformar situando en cascada un precompensador que tenga una función de transferencia arbitraria estable $H_p(z)$. Con esto la ley de control será

$$u(k) = \frac{C(q)}{R(q)}H_p(q)u_c(k) - \frac{S(q)}{R(q)}y(k) \quad (12.83)$$

lo cual produce

$$y(k) = \frac{B(q)}{P(q)}H_p(q)u_c(k) + \frac{R(q)}{P(q)}e(k) \quad (12.84)$$

Como el polinomio P es estable, éste puede cancelarse con el precompensador. De esto se deduce que la respuesta a perturbaciones y a señales de mando se pueden conformar separadamente.

En primer lugar se diseña la realimentación S/R para asegurar una buena respuesta frente a las perturbaciones. Después, se elige el precompensador H_p para obtener la respuesta deseada a las señales de mando.

12.6. Aspectos prácticos

Muchas de las arbitrariedades del diseño parecen desaparecer cuando los problemas de diseño se formulan como problemas de optimización. Se establecen los modelos y los criterios, y las leyes de control se obtienen simplemente como la solución de un problema de optimización. Esta simplicidad es engañosa porque las arbitrariedades se transfieren al modelado y a la formulación de los criterios. Una aplicación afortunada de la teoría de optimización requiere conocimientos de cómo se reflejan las propiedades del modelo y los criterios en las leyes de control. Cuestiones típicas son: ¿Qué debe tenerse en cuenta en el modelo para obtener reguladores con acción integral? ¿Qué definiciones del problema producen reguladores con estructura PID? Algunos de estos problemas se discuten en esta sección, los cuales también aclaran las propiedades de las leyes de control óptimo. Esto muestra que algunos resultados pueden formularse como reglas de diseño. El método polinómico, que opera directamente con la función de transferencia, es adecuado para hacer esto.

También se discuten otros aspectos de relevancia práctica, tales como sensibilidad y robustez. Se da un breve tratamiento del rizado intermuestreo de la función de coste junto a algunos aspectos de la elección del período de muestreo.

Propiedades del regulador óptimo

Algunas propiedades del modelo influyen en las leyes de control óptimo. El modelo básico utilizado es el definido por (12.5); es decir:

$$A(q)y(k) = B(q)u(k) + C(q)e(k) \quad (12.85)$$

La razón B/A representa la función de transferencia discreta del proceso y la razón C/A representa la función de transferencia discreta que genera la perturbación de la salida del proceso. Los polinomios A , B y C pueden tener factores comunes que reflejan el modo en que la señal de control y las perturbaciones se acoplan en el sistema. Sin embargo, no hay factores comunes a los tres polinomios. Compárese con la exposición de la Sección 12.2, donde se dedujo el modelo. La presencia de factores comunes que influyen directamente en las propiedades de los reguladores se estudian a continuación.

El principio del modelo interno. Los factores comunes a los polinomios A y B corresponden a modos de perturbación que no son controlables desde u . Tales modos aparecerán como factores de P . Sea

$$A_2 = \text{mcd}(A, B) \quad (12.86)$$

el máximo común divisor de los polinomios A y B . Si A_2 es estable, se deduce del Teorema 12.4 que A_2 también divide a P . Si A_2 tiene un factor A_2^+ con todos sus ceros fuera del círculo unidad, el resultado correspondiente se deduce del Teorema 12.5. En este caso también se deduce del Teorema 12.5 que A_2^+ divide a R . Esta observación se denomina el *principio del modelo interno*; y dice que para regular un sistema con perturbaciones inestables, la dinámica de las perturbaciones tiene que aparecer en la dinámica del regulador. Unos pocos ejemplos ilustran esta idea.

Ejemplo 12.8 – Acción integral

Un regulador tiene acción integral si $z - 1$ divide a $R(z)$. Se deduce del Teorema 12.5 y del principio del modelo interno, que esto ocurrirá si $z - 1$ divide a A y B , lo cual significa que el modelo es de la forma

$$A_1(q)(q - 1)y(k) = B_1(q)(q - 1)u(k) + C(q)e(k) \quad (12.87)$$

Esto significa que hay una perturbación de desplazamiento.

Ejemplo 12.9 – Eliminación de una perturbación sinusoidal

Una perturbación sinusoidal de banda estrecha con frecuencias centradas en ω puede representarse como un ruido blanco alimentando a un sistema con numerador

$$D(q) = q^2 - 2q \cos \omega h + 1$$



Si los polos de la dinámica del sistema no corresponden a D , el modelo es

$$A_1(q)D(q)y(k) = B_1(q)D(q)u(k) + C(q)e(k) \quad (12.88)$$

Entonces, el regulador óptimo es tal que $D(z)$ divide a $R(z)$.

CANCELACIÓN DE POLOS DEL PROCESO. Un factor común de A y C corresponde a modos controlables que no se excitan con las perturbaciones. Sea A_2 el máximo común divisor de A y C . El polinomio A_2 es estable porque C es estable, y no divide a B porque no hay factores que dividan a A , B y C . De (12.49) se deduce que A_2 también divide al polinomio S , el cual es el numerador de la función de transferencia del regulador. De este modo, *los polos estables del proceso que no se excitan por las perturbaciones pueden ser cancelados*.

CANCELACIÓN DE CEROS DEL PROCESO. Los factores comunes de B y C corresponden a ceros de proceso que bloquean la transmisión para la señal de control u y para la perturbación e . Sea B_2 el máximo común divisor de B y C . El polinomio B_2 es estable y no divide a A . Entonces, se deduce de (12.49) que B_2 divide a R . Esto significa que los ceros correspondientes a $B_2=0$ son cancelados por el regulador. Por tanto, *los ceros del proceso que también son ceros de transmisión para la perturbación C son cancelados por el regulador*.

Para el control de varianza mínima, de (12.46) con $\rho=0$ se deduce que

$$P = q^d B^+ B^- *$$

y de (12.49) que B^+ divide a R . De este modo, todos los ceros estables se cancelan con la ley de control de varianza mínima.

Un análisis de las propiedades de la ley de control óptimo da, de este modo, respuestas parciales al problema clásico de cancelación.

Sensibilidad y robustez

Es importante que un sistema de control sea insensible o robusto frente a los errores de medida, perturbaciones en la planta y errores de modelado. Esto puede analizarse como se hizo en la Sección 10.6 para el problema de la asignación de polos. Las propiedades de la robustez se expresan adecuadamente en términos de la ganancia del lazo:

$$H_{lg} = \frac{BS}{AR} \quad (12.89)$$

o de la diferencia de retorno

$$H_{rd} = 1 + \frac{BS}{AR} = \frac{AR + BS}{AR} = \frac{PC}{AR} \quad (12.90)$$

La ganancia del lazo $H_{lg}(\exp i\omega h)$ es normalmente grande para las bajas frecuencias y pequeña para las altas frecuencias. La frecuencia de cruce ω_c es la frecuencia mínima para la que

$$|H_{lg}(e^{j\omega h})| = 1$$

El sistema en lazo cerrado es insensible a las perturbaciones de la planta para aquellas frecuencias en las que la ganancia es alta. Para tener baja sensibilidad frente a los malos modelados de la dinámica de alta frecuencia de la planta, interesa que la ganancia del lazo decrezca rápidamente para las frecuencias superiores a la de cruce. Es posible asegurar que la ganancia del lazo es alta para ciertas frecuencias eligiendo modelos con una estructura especial, como se hizo en los Ejemplos 12.8 y 12.9. También son útiles para evaluar la sensibilidad representaciones similares a las de la Figura 10.5. En un sistema muestreado diseñado adecuadamente, se dispondrá de filtros anti-enmascaramiento de frecuencias, que eliminan la transmisión de señal de las frecuencias superiores a las de Nyquist. La selección de una velocidad baja de muestreo asegura que la ganancia del lazo es baja a partir de una frecuencia dada. Esto significa también que los errores de modelado en altas frecuencias tienen poca influencia. Obsérvese, sin embargo, que las representaciones de la ganancia del lazo y la diferencia de retorno no darán la representación completa porque puede haber cancelación de polos y ceros que no aparece en este tipo de representaciones.

En este caso, es útil un análisis de la ecuación característica. Para efectuar tal análisis, supongamos que el sistema está gobernado por

$$A^0(q)y(k) = B^0(q)u(k) + C^0(q)e(k) \quad (12.91)$$

pero que se diseña un regulador basándose en un modelo diferente, como en (12.85). Los reguladores definidos por los Teoremas 12.4 y 12.5 proporcionan un sistema en lazo cerrado con el polinomio característico

$$\begin{aligned} A^0R + B^0S &= A^0R - AR + B^0S - BS + AR + BS \\ &= PC + (A^0 - A)R + (B^0 - B)S \end{aligned} \quad (12.92)$$

Cuando el modelo de (12.85) es igual al sistema (12.91) el polinomio característico es $PC = P_1 A_2 C$, como se esperaba. También se deduce por continuidad que pequeños cambios en el sistema producen pequeños cambios en los polos de lazo cerrado. El sistema es sensible a los cambios en los parámetros si los polinomios P_1 o C tienen ceros cerca del círculo unidad.

Para garantizar sistemas con baja sensibilidad, es necesario además imponer restricciones. Recuérdese que C y P_1 se obtuvieron como solución de un problema de factorización espectral.

Sistemas en lazo cerrado con estabilidad exponencial garantizada

Las leyes de control proporcionadas por los Teoremas 12.2, 12.3, 12.4 y 12.5 dan sistemas en lazo cerrado con polos dentro del círculo unidad. Algunas veces se desea disponer de leyes de control tales que el sistema en lazo cerrado tenga sus polos dentro de un círculo de radio r . Es inmediato formular problemas de optimización que proporcionen estas leyes de control.

Introducamos el criterio

$$J = E \left[\left(\frac{1}{r} \right)^{2k} (\eta^2(k) + \mu u^2(k)) \right] \quad (12.93)$$

Si se puede encontrar una ley de control que minimice este criterio, las variables $y(k)$ y $u(k)$ tienen que converger a cero al menos tan rápido como lo hace r^k cuando k se incrementa. Para obtener este resultado, debe suponerse que el modelo de (12.5) es tal que la covarianza de $e(k)$ también tiende a cero como r^k .

Introducamos las variables de escala η , μ y ϵ definidas por

$$\begin{aligned} y(k) &= r^k \eta(k) \\ u(k) &= r^k \mu(k) \\ e(k) &= r^k \epsilon(k) \end{aligned} \quad (12.94)$$

Puesto que

$$q'y(k) = q'[r^k \eta(k)] = r^{k+l} \eta(k+l) = r^k (rq)' \eta(k)$$

se deduce que

$$A(q)y(k) = A(q)[r^k \eta(k)] = r^k A(rq)\eta(k)$$

Introduciendo los polinomios transformados

$$\tilde{A}(z) = A(rz)$$

$$\tilde{B}(z) = B(rz)$$

$$\tilde{C}(z) = C(rz)$$

el modelo de (12.5) puede escribirse como

$$\tilde{A}(q)\eta(k) = \tilde{B}(q)\mu(k) + \tilde{C}(q)\epsilon(k) \quad (12.95)$$

y el criterio de (12.93) será

$$J = E[\eta^2(k) + \rho \mu^2(k)] \quad (12.96)$$

Entonces, la ley de control que minimiza (12.96) para el sistema de (12.95) viene expresada por el Teorema 12.4. Esta ley de control da un sistema en lazo cerrado en el que todos los ceros de la ecuación característica

$$\tilde{P}(z)\tilde{C}(z) = 0$$

están dentro del círculo unidad. Retornando a las variables originales resulta la ecuación característica

$$P(z)C(z) = \tilde{P}\left(\frac{z}{r}\right)\tilde{C}\left(\frac{z}{r}\right) = 0$$

Todos los ceros de esta ecuación están dentro del círculo $|z|=r$.

De este modo se ha obtenido un procedimiento simple para obtener leyes de realimentación que producen sistemas en lazo cerrado con todos los polos dentro del círculo $|z|=r$.

Reducción de perturbaciones

La diferencia de retorno es

$$H_{rd}(z) = 1 + H_{ls}(z) = 1 + \frac{BS}{AR} = \frac{AR + BS}{AR}$$

La inversa de la diferencia de retorno es una medida de la eficiencia del sistema en lazo cerrado para eliminar perturbaciones.

Consideremos el modelo de (12.85). Sin control la salida es

$$y_{ol} = \frac{C}{A} e$$

Con la ley de control LQ, la salida será

$$y_{lsg} = \frac{R}{P} e$$

Eliminando e entre estas ecuaciones se obtiene

$$y_{lsg} = \frac{AR}{PC} y_{ol} = \frac{1}{PC} y_{ol} = \frac{1}{1 + \frac{BS}{AR}} y_{ol} = \frac{1}{H_{rd}} y_{ol} \quad (12.97)$$

De este modo la función de transferencia discreta

$$H(z) = \frac{A(z)R(z)}{P(z)C(z)} = \frac{1}{H_{rd}(z)} \quad (12.98)$$

que es la inversa de la diferencia de retorno, muestra lo que se atenúan las perturbaciones de las diferentes frecuencias.

Selección del período de muestreo

Hay una diferencia sustancial entre la ley de control de varianza mínima estudiada en la Sección 12.4 y la ley de control LQG estudiada en la Sección 12.5 en función de la influencia del período de muestreo. La elección del período de muestreo es crítica para el control de varianza mínima. Un período de muestreo corto produce un sistema con un gran ancho de banda, que se asienta rápidamente. Las acciones de control también serán grandes cuando el período de muestreo es corto. A este respecto, la ley de control de varianza mínima es similar a la ley de control de tiempo



finito estudiada en la Sección 9.2. El periodo de muestreo es menos crítico en el control LQG. Del análisis de la Sección 11.5 se deduce que la ley de control se aproxima al control continuo cuando el periodo de muestreo h tiende a cero. Para períodos de muestreo pequeños, la diferencia entre las funciones de coste continua y muestreada también es cuadrática en h . Por tanto, el siguiente estudio se centra en el control de varianza mínima.

Variación de la varianza de la salida entre instantes de muestreo

La ley de control de varianza mínima minimiza la varianza de la salida *en los instantes de muestreo*. Sin embargo, el principal objetivo puede ser minimizar la función de coste continua de (12.7). Esto puede realizarse muestreando, en primer lugar, la función de coste continua y posteriormente minimizando la correspondiente función de coste discreta. Esto constituye un procedimiento de diseño más complicado. En algunos casos, las leyes de control de varianza mínima son una aproximación suficientemente buena. Es útil estudiar la variación de la función de coste entre los instantes de muestreo. Este análisis es similar al análisis del rizado entre instantes de muestreo para los sistemas determinísticos de la Sección 10.5. Un ejemplo ilustra esta idea.

Ejemplo 12.10 – Variación de la función de coste entre intervalos de muestreo

Consideremos el sistema continuo

$$dx = u dt + dv \quad (12.99)$$

donde $\{v(t)\}$ es un proceso de Wiener con covarianza incremental $\sigma_v^2 dt$. Supongamos que la salida se observa sin filtro anti-enmascaramiento en los instantes $t_k = k \cdot h$, donde h es el periodo de muestreo. Por tanto,

$$y(t_k) = x(kh) + \epsilon(t_k)$$

donde $\{\epsilon(t_k)\}$ es una secuencia de variables aleatorias independientes de media cero y covarianza σ_ϵ^2 . Muestreando los sistemas se obtiene

$$x(kh + h) = x(kh) + hu(kh) + v(kh + h) - v(kh)$$

$$y(kh) = x(kh) + \epsilon(kh)$$

Por tanto,

$$y(kh + h) = y(kh) + hu(kh) + v(kh + h) + \epsilon(kh + h) - \epsilon(kh) + v(kh + h) - v(kh)$$

La perturbación del lado derecho puede representarse como

$$w(kh + h) = e(kh + h) + ce(kh)$$

donde $\{e(kh)\}$ es una secuencia de variables aleatorias independientes, de media cero con desviación estándar σ_e .

Se obtiene fácilmente que

$$c = -1 - \frac{h\sigma_v^2}{2\sigma_e^2} + \sqrt{\frac{h\sigma_v^2}{\sigma_e^2} + \frac{h^2\sigma_v^4}{4\sigma_e^4}} \quad (12.100)$$

$$\sigma^2 = -\frac{\sigma_e^2}{c} \quad (12.101)$$

La ley de control de varianza mínima para el sistema es

$$u(kh) = -\frac{1 + c}{h} y(kh)$$

La desviación estándar de la salida bajo control de varianza mínima es

$$Ey^2(t) = \sigma^2 \quad t = h, 2h, \dots$$

La desviación estándar de la variable de estado x es

$$Ex^2(t) = \sigma^2 - \sigma_e^2 \quad t = h, 2h, \dots$$

Integramos la ecuación (12.99) para determinar la varianza de la variable de estado entre los instantes de muestreo. Esto da

$$\begin{aligned} x(kh + s) &= x(kh) + su(kh) + v(kh + s) - v(kh) \\ &= (1 - \alpha s)x(kh) - \alpha s \epsilon(kh) + v(kh + s) - v(kh) \end{aligned}$$

donde

$$\alpha = (1 + c)/h$$

Introduzcamos

$$P(s) = Ex^2(kh + s)$$

Entonces, se deduce que

$$P(s) = (1 - \alpha s)^2(\sigma^2 - \sigma_e^2) + (\alpha s)^2\sigma_e^2 + s\sigma_v^2$$

La función $P(s)$ se muestra en la Figura 12.3. Obsérvese que

$$\max_s [P(0) - P(s)] = h^2\sigma_v^2/2$$

Por tanto, la variación de P sobre un intervalo de muestreo decrece cuando h decrece.

El análisis es similar en el caso general. La única diferencia está en que debe utilizarse el Teorema 6.5 para calcular la covarianza del estado.

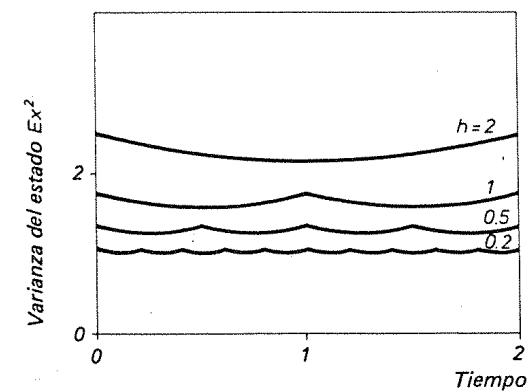


Figura 12.3 Variaciones de la varianza de la salida con el tiempo en el Ejemplo 12.10 para reguladores que tienen diferentes períodos de muestreo.

Efectos del período de muestreo en el control de varianza mínima

Para sistemas estables con inversos estables, donde el retardo de tiempo es un múltiplo entero del período de muestreo, la varianza mínima es el error en la predicción de la salida sobre un horizonte de tiempo que corresponde a la suma de un retardo de tiempo y el período de muestreo. De este modo, una representación del error en la predicción en función del horizonte de tiempo dará una estimación de cómo depende la varianza del período de muestreo (véase la Figura 12.4). La varianza

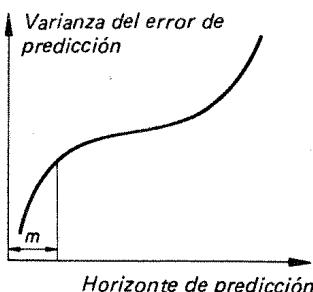


Figura 12.4 Varianza del error de predicción frente al período de muestreo.

mínima será mayor que la varianza de predicción si el proceso es inestable o tiene inverso inestable. Puesto que la estabilidad del inverso puede depender del período de muestreo, se complica la dependencia de la varianza y el período de muestreo. Otra complicación es que el retardo de tiempo puede no ser un múltiplo exacto del período de muestreo. Consideremos, por ejemplo, un sistema con un retardo de tiempo de una unidad. Se requiere un horizonte de predicción de dos períodos para un período de muestreo de $h=0,52$, pero se requiere un horizonte de predicción de tres períodos para $h=0,48$. De este modo, la dependencia real de la varianza mínima se complica. Se puede calcular fácilmente para casos específicos. Un ejemplo ilustra este principio.

Ejemplo 12.11—Dependencia de la varianza mínima del período de muestreo

Consideremos el sistema

$$\begin{aligned} dx(t) &= u(t-1)dt + dv \\ y(kh) &= x(kh) + \epsilon(kh) \quad k = 0, 1, 2, \dots \end{aligned}$$

donde $\{v(t)\}$ es un proceso de Wiener con covarianza incremental $\sigma_v^2 dt$ y $\{\epsilon(kh)\}$ es una secuencia de variables aleatorias independientes con media cero y varianza σ_e^2 . Muestreando el sistema con período h se obtiene

$$y(kh+h) = y(kh) + b_0u(kh-dh+h) + b_1u(kh-dh) + e(kh+h) + ce(kh)$$

donde

$$\begin{aligned} 1 &= (d-1)h + \tau' \\ b_0 &= h - \tau' \\ b_1 &= \tau' \end{aligned}$$

El número c está definido por (12.100) y la varianza σ^2 de e viene dada por (12.101). Una aplicación directa del Teorema 12.3 da

$$\min E y^2 = f(h)$$

donde

$$f(h) = \begin{cases} 1 + (d-1)(1+c)^2 & \tau' \leq h/2 \\ 1 + (d-1)(1+c)^2 + \frac{(1+c)^2(2\tau' - h)}{h} & \tau' > h/2 \end{cases}$$

En la Figura 12.5 se muestra una representación de la función f . Obsérvese que las partes de la curva con pendiente negativa corresponden a aquellos períodos de muestreo donde el sistema tiene inverso inestable. El ejemplo indica claramente la complejidad de la relación entre la varianza mínima y el período de muestreo. Para analizar el resultado, también es útil considerar el rizado de la función de coste entre los instantes de muestreo (compárese con el Ejemplo 12.10).

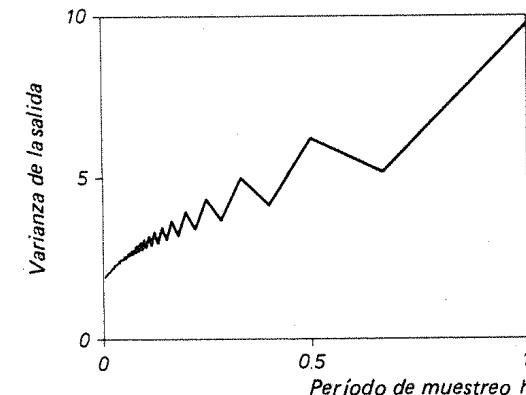


Figura 12.5 Varianza mínima frente al período de muestreo para el sistema del Ejemplo 12.11.

Para las leyes de control de tiempo mínimo o varianza mínima, la varianza de la señal de control se incrementa cuando el período de muestreo disminuye. Para elegir el período de muestreo de estos sistemas es útil dibujar la varianza de entrada en función de la varianza de salida tomando el período de muestreo como parámetro. En la Figura 12.6 se muestra un ejemplo de esta curva. La curva es válida para el sistema del Ejemplo 12.11.

Aspectos del cálculo

La ley de control LQ puede determinarse como una combinación de factorización espectral y solución de ecuaciones lineales diofánticas. Sin embargo, recordemos la dificultad fundamental que se presentaba por la mala condición numérica de las ecuaciones polinómicas (véase la Sección 15.6).

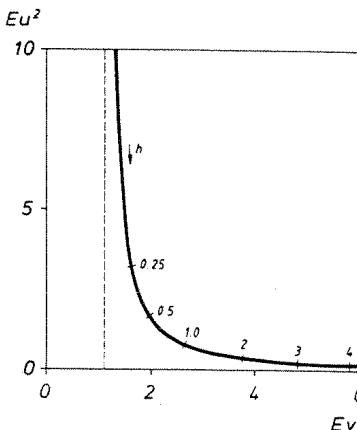


Figura 12.6 Varianza de la salida frente a la covarianza de la entrada para el regulador de varianza mínima del Ejemplo 12.11.

12.7. Conclusiones

En este capítulo se han resuelto los problemas de control óptimo para sistemas descritos por modelos de entrada-salida. Los resultados se han limitado a los sistemas de una entrada y una salida. En primer lugar se deduce un modelo canónico para el sistema, la ecuación (12.5). Este modelo está caracterizado por tres polinomios, A , B y C . El modelo continuo fundamental puede describirse como una combinación de un retardo de tiempo y un sistema con función de transferencia racional. Las perturbaciones se caracterizan como ruido blanco filtrado. Hay muchos sistemas físicos que pueden describirse con estos modelos.

Se resuelven para este tipo de sistemas los problemas de control óptimo caracterizados por funciones de coste cuadráticas. En primer lugar se considera un caso especial donde la función de coste simplemente es la varianza de la salida. Despues se trata el problema general, en el cual también hay una penalización de la variable de control. Estos métodos están estrechamente relacionados al problema de predicción para un proceso aleatorio con densidad espectral racional. También se resuelve este problema y se estudian aspectos prácticos, como la selección del período de muestreo.

La solución de los problemas de control óptimo proporcionan herramientas de diseño. Las soluciones también aclaran el carácter de las soluciones óptimas. En particular, nos dicen que el regulador óptimo siempre cancela los polos estables del proceso que también son ceros de las perturbaciones del proceso. Los polos estables del proceso se cancelan sólo si no están excitados por perturbaciones. El resultado también aclara la relación entre los diferentes métodos de diseño. Por ejemplo, las soluciones LQG pueden interpretarse como reguladores por asignación de polos, cuando los polos del proceso y los polos del observador se eligen de un modo especial.

El cálculo de la solución óptima se expresa en función de dos operaciones polinómicas, factorización espectral y solución de ecuaciones diofánticas.

12.8. Problemas

12.1. Considerar el proceso

$$y(k) = \frac{q^2 - 1.4q + 0.5}{q^2 - 1.2q + 0.4} e(k)$$

donde $e(k)$ es un ruido blanco con media cero y varianza unidad. Determinar el predictor óptimo de m pasos de anticipación y la varianza del error de predicción cuando $m=1, 2$ y 3 .

12.2. Determinar el predictor de m pasos de anticipación para el proceso

$$y(k) + ay(k-1) = e(k) + ce(k-1)$$

Determinar también la varianza del error de predicción como una función de m .

12.3. Un proceso estocástico se describe por

$$y(k) - 0.9y(k-1) = e(k) + 5e(k-1)$$

a) Determinar una descripción equivalente tal que el cero del polinomio correspondiente C esté dentro del círculo unidad.

b) Determinar el predictor de dos pasos de anticipación para el proceso y la varianza del error de predicción.

12.4. Suponer que la demanda de un producto en un inventario, $z(k)$, puede describirse como

$$z(k) = 300 + 10k + y(k)$$

donde la unidad de tiempo son meses e $y(k)$ se describe por el proceso

$$y(k) - 0.7y(k-1) - 0.1y(k-2) = 5e(k)$$

donde $e(k)$ es un ruido blanco con media cero y varianza unidad. Realizar una predicción y determinar la desviación estándar esperada del error de predicción de agosto a noviembre cuando se disponen de los siguientes datos:

Meses	k	$z(k)$
Enero	1	320
Febrero	2	320
Marzo	3	325
Abril	4	330
Mayo	5	350
Junio	6	370
Julio	7	375

12.5. Considerar el proceso

$$\begin{aligned} y(k) - y(k-1) + 0.5y(k-2) \\ = u(k-2) + 0.5u(k-3) + 0.5[e(k) + 0.8e(k-1) + 0.25e(k-2)] \end{aligned}$$

Determinar el controlador de varianza mínima y la varianza mínima que puede lograrse.



- 12.6. Determinar el controlador de varianza mínima para el sistema

$$y(k) - 0.5y(k-1) = u(k-2) + e(k) - 0.7e(k-1)$$

donde $e(k)$ es ruido blanco de media 2 y varianza unidad.

- 12.7. Considerar el proceso

$$y(k) + ay(k-1) = u(k-2) + e(k) + ce(k-1)$$

- a) Determinar el controlador de varianza mínima.

- b) Discutir el caso especial $a=0$.

- 12.8. Dado el sistema

$$\begin{aligned} y(k) &= 1.7y(k-1) + 0.7y(k-2) \\ &= u(k-d) + 0.5u(k-d-1) + e(k) + 1.5e(k-1) + 0.9e(k-2) \end{aligned}$$

- a) Determinar el controlador de varianza mínima y la varianza de la salida para $d=1$ y 2.
b) Simular el sistema en lazo abierto y el sistema controlado con el controlador de varianza mínima. Comparar la salida y las señales de control en los diferentes casos.

- 12.9. Considerar el proceso de la Figura 12.7. La perturbación z tiene la densidad espectral

$$\phi_z(\omega) = \frac{1}{2\pi} \frac{1}{1.36 + 1.2 \cos \omega}$$

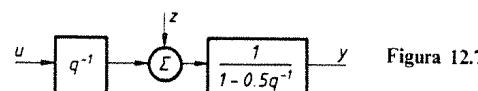


Figura 12.7

- a) Determinar la función de transferencia discreta $H(z)$, que da una salida con densidad espectral ϕ_z cuando se alimenta con ruido blanco de media nula y varianza unidad.
b) ¿Cuál es la varianza de y en estado estacionario cuando

$$u(k) = -Ky(k)$$

para $K=1$?

- c) ¿Cuál es la mínima varianza que puede conseguirse para un controlador proporcional y cuál es el correspondiente valor de K ?
d) ¿Cuál es la varianza de y cuando se utiliza un controlador de varianza mínima?

- 12.10. Sea el sistema

$$y(k) - 0.25y(k-1) + 0.5y(k-2) = u(k-1) + e(k) + 0.5e(k-1)$$

donde $e(k)$ es un ruido blanco con varianza unidad. Supongamos que el proceso se controla con el controlador proporcional

$$u(k) = -Ky(k)$$

- a) Demostrar que la varianza de la salida es

$$\frac{2.125 - K}{0.5(1.75 - K)(1.25 + K)}$$

y que el control de varianza mínima se obtiene para $K=1$, el cual proporciona la varianza $\frac{1}{2}$.

- b) La expresión anterior es cero para $K=2.125$. Explicar la paradoja.

- 12.11. Dado el proceso

$$y(k) - 1.5y(k-1) + 0.7y(k-2) = u(k-2) - 0.5u(k-3) + v(k)$$

- a) Suponer que $v(k)=0$ y calcular el controlador de tiempo finito para el sistema.

- b) Suponer que

$$v(k) = e(k) - 0.2e(k-1)$$

donde $e(k)$ es un ruido blanco. Calcular la ley de control de varianza mínima.

- c) ¿Cuál es la varianza en estado estacionario de y cuando se utilizan los controladores de tiempo finito y de varianza mínima si v es la misma que en b)?
d) Simular el sistema utilizando los diferentes controladores. Estudiar la salida y las pérdidas acumuladas; es decir, la suma de los cuadrados de la salida.

- 12.12. Considerar el sistema dinámico

$$y(k) = \frac{B(q)}{A(q)}u(k) + \lambda \frac{C(q)}{D(q)}e(k)$$

donde $e(k)$ es un ruido blanco. Se supone que los polinomios A , C y D son mónicos. Determinar el controlador de varianza mínima para el sistema.

- 12.13. Utilizar el resultado del Problema 12.12 para determinar el controlador de varianza mínima para el sistema

$$y(k) = \frac{bq^{-1}}{1+aq^{-1}}u(k) + (1+cq^{-1})e(k)$$

- 12.14. Considerar el proceso del Problema 12.13. Suponer que se dobla el período de muestreo; es decir, la señal de control puede cambiarse sólo cada dos unidades de tiempo. Determinar el controlador de varianza mínima y comparar con el caso en que el período de control es una unidad de tiempo.

- 12.15. Considerar el sistema de la Figura 12.8 donde e es un ruido blanco de media cero y varianza unidad. Además,

$$\begin{aligned} A(q) &= q - 0.7 & B(q) &= q \\ C(q) &= 1 - 0.5q & \alpha &= -0.8 \end{aligned}$$

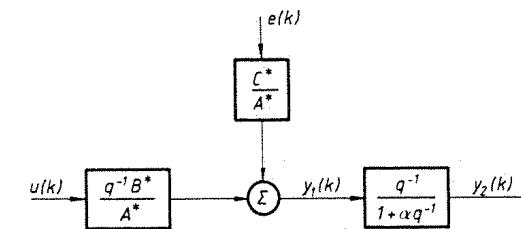


Figura 12.8



- a) Determinar un controlador que minimice la varianza de y_1 .
 b) Determinar la varianza de y_1 e y_2 cuando se utiliza el controlador de a).
 c) Determinar un controlador que minimice la varianza de y_2 si solamente es medible y_2 y calcular las varianzas de y_1 e y_2 .
 d) Determinar un controlador que minimice la varianza de y_2 si y_1 e y_2 son medibles.
 e) ¿Cuáles son las varianzas de y_1 e y_2 cuando se utiliza el controlador de d)?

12.16. Dados los procesos

$$A(q)y(k) = B(q)u(k) + C(q)e(k) + D(q)v(k)$$

donde $v(k)$ es una perturbación conocida. Determinar el controlador de varianza mínima para el proceso cuando $\text{gra } B = \text{gra } D$.

12.17. Determinar el controlador LQG definido por el Teorema 12.4 para el proceso

$$(1 - 0.9q^{-1})y(k) = u(k) + (1 - 0.5q^{-1})e(k)$$

cuando $\rho = 1$. Calcular la varianza de la salida y de la entrada para diferentes valores de ρ .

- 12.18. Considerar un sistema con inverso estable. Deducir el controlador de varianza mínima, donde se permite que la señal de control $u(k)$ sea una función de $y(k-1)$, $y(k-2)$, ... $u(k-1)$, ... Deducir la ecuación característica del sistema en lazo cerrado.
 12.19. Demostrar que la función de transferencia discreta entre e e y para (12.5) y (12.56) viene dada por (12.55). Utilizar (12.45) para deducir el controlador de varianza mínima para un sistema donde

$$A(q) = q^2 - 1.5q + 0.7$$

$$B(q) = q + 0.5$$

$$C(q) = q^2 - q + 0.24$$

Comparar con el controlador obtenido con (12.17).

- 12.20. Determinar para qué sistemas un controlador digital PID tiene la misma estructura que el controlador óptimo de varianza mínima.
 12.21. Considerar un sistema descrito por

$$y(k) = \frac{1}{q-a}[bu(k) + e(k)] + \frac{1}{q-1}w(k)$$

donde e y w son procesos de ruido blanco con media cero y desviación estándar σ_e y σ_w , respectivamente.

- a) Reducir el sistema a la forma estándar y determinar el controlador de varianza mínima.
 b) Interpretar el controlador a) como un controlador PI y determinar cómo dependen la ganancia y el tiempo de integración de la relación σ_w^2/σ_e^2 .
 12.22. Considerar la ley de control de varianza mínima de (12.36) para un sistema con inverso inestable. La salida del sistema en lazo cerrado está definida por

$$y(k) = \frac{F(q)}{q^{d-1}B^{-*}(q)}e(k)$$

Demostrar que la función F/B^{-*} tiene el desarrollo en serie

$$\frac{F(q)}{B^{-*}(q)} = q^{d-1} + f_1q^{d-2} + \dots + f_{d-1} + \frac{F_2(q)}{B^{-*}(q)}$$

donde $\text{gra } F_2(q) < \text{gra } B^{-*}$ y

$$F_1(q) = q^{d-1} + f_1q^{d-2} + \dots + f_{d-1}$$

es el cociente de $q^{d-1}C(q)$ y $A(q)$. Encontrar un modo adecuado de calcular F_2 . Utilizar el resultado del problema para determinar el incremento de la varianza mínima debida a los ceros inestables del sistema.

- 12.23. Determinar el rizado entre períodos de muestreo de la función de coste cuando el proceso

$$dx_1 = x_2 dt$$

$$dx_2 = u dt + dv$$

$$y(t_k) = x(t_k) + \epsilon(t_k)$$

está controlado por el regulador de varianza mínima. El proceso $\{v(t)\}$ es un proceso de Wiener con covarianza incremental $\sigma_v^2 dt$ y $\{\epsilon(t_k)\}$ es un ruido de medida blanca con media cero y varianza σ_ϵ^2 .

- 12.24. Considerar el proceso del Ejemplo 12.10. Determinar la ley de control con un período de muestreo h que minimiza

$$\lim_{T \rightarrow \infty} E \frac{1}{T} \int_0^T y^2(s) ds$$

y compararla con la ley de control de varianza mínima.

- 12.25. Considerar un proceso sujeto a perturbaciones que se caracteriza como un proceso de Wiener con covarianza incremental dt . Determinar el error de predicción de la varianza mínima en cada caso. Utilizar diferentes límites de predicción y períodos de muestreo.

- a) El proceso tiene un cero inestable $z = b > 1$.
 b) El proceso tiene un polo inestable $z = a > 1$.

- 12.26. Considerar el sistema del Problema 12.23 con un retardo de tiempo extra de 1 s. Determinar la varianza mínima en función del período de muestreo.

- 12.27. Considerar el sistema del Problema 12.23. Determinar la varianza de la salida como función de la covarianza de la entrada para diferentes períodos de muestreo.

- 12.28. Considerar el sistema

$$y(k) = \frac{1}{q - 0.999}u(k) + \frac{q}{q - 0.7}e(k)$$

Determinar la ley de control de varianza mínima para el sistema. Compararla con una realimentación proporcional que produce la misma respuesta. Discutir las ventajas de las leyes de control calculando las ganancias del lazo y las diferencias de retorno. Explicar por qué es inferior el control de varianza mínima. (Sugerencia: Un problema de optimización malo produce un regulador óptimo malo.)

12.9. Referencias

- El tratamiento del caso lineal cuadrático es el fundamento del trabajo de Wiener; véase WIENER, N. (1949): *Extrapolation, Interpolation & Smoothing of Stationary Time Series*. Cambridge, Mass.: MIT Press.



NEWTON, G. C., L. A. GOULD y J. F. KAISER (1957): *Analytical Design of Linear Feedback Controls*. New York: John Wiley.

YOULA, D. C., J. J. BONGIORNO, JR., y H. A. JABR (1976): «Modern Wiener-Hopf Design of Optimal Controllers. Part. I: The Single-Input-Single-Output Case; Part II: The Multivariable Case.» *IEEE Trans. Autom. Control*, AC-21, 3-13 y 319-38.

Un tratamiento completo del control de predicción y control de varianza mínima se encuentra en

ÅSTRÖM, K. J. (1970): *Introduction to Stochastic Control Theory*. New York: Academic Press.

el cual se basa en

ÅSTRÖM, K. J. (1965): *Notes on the Regulation Problem*, Report CT211, IBM Nordic Laboratory.

ÅSTRÖM, K. J. (1967): «Computer Control of a Paper Machine: An Application of Linear Stochastic Control Theory.» *IBM Journal Res. Develop.*, II, 389-405.

Un tratamiento similar del problema de control estocástico se encuentra en

BOX, G. E. P. y G. M. JENKINS (1970): *Time Series Analysis, Forecasting, and Control*, San Francisco: Holden-Day.

El teorema para control de varianza mínima de sistemas con inversos inestables se publicó inicialmente en

PETERKA, V. (1972): «On Steady-State Minimum Variance Control Strategy.» *Kybernetika*, 8, 219-32.

Un tratamiento algebraico de los problemas de control multivariable LQ y varianza mínima se da en

KUČERA, V. (1979): *Discrete Linear Control*. Prague: Academia.

La elección del intervalo de muestreo para el control estocástico se expone en los libros mencionados, y también en

MACGREGOR, J. F. (1976): «Optimal Choice of the Sampling Interval for Discrete Process Control.» *Technometrics*, 18, no. 2, 151-60.

Identificación

Objetivo:

Exponer cómo se pueden obtener modelos de procesos a partir de datos experimentales.
Presentar el método de mínimos cuadrados.

13.1. Introducción

La noción de modelo matemático es fundamental para la ciencia y la ingeniería. Un modelo es una forma muy útil y compacta de sintetizar el conocimiento que se tiene de un proceso. Un modelo es también una herramienta muy eficaz para la educación y la comunicación. Los métodos de diseño de los capítulos anteriores suponen que se dispone de los modelos de los procesos y de las perturbaciones. Algunas veces se pueden obtener los modelos de los procesos a partir de los principios de la Física. Es más difícil conseguir los modelos de las perturbaciones, que son igualmente importantes. A menudo estos modelos hay que obtenerlos mediante experimentos. Los tipos de modelos necesarios para los métodos de diseño expuestos aquí son tanto modelos en el espacio de estados (modelos internos) como modelos de entrada-salida (modelos externos). Para los modelos internos los modelos de las perturbaciones se dan como sistemas dinámicos gobernados por un ruido blanco. Para los modelos externos las perturbaciones se dan en términos de densidades espectrales y funciones de covarianza. Sin embargo, sólo en raras ocasiones se pueden deducir los modelos de las perturbaciones a partir de principios físicos. Entonces con frecuencia la única forma de construir modelos de las perturbaciones es mediante experimentos.

Un proceso no puede caracterizarse por un único modelo matemático. Debe representarse por una jerarquía de modelos que van desde los detallados y complejos de simulación hasta los muy sencillos, fáciles de manipular analíticamente. Los modelos sencillos se utilizan para exploraciones y para obtener las características más

destacadas del comportamiento del sistema. Los modelos detallados se emplean para comprobaciones precisas del comportamiento del sistema de control. Lleva mucho tiempo desarrollar los modelos complejos. Entre los dos extremos, puede haber muchos tipos de modelos. El distintivo de una buena ingeniería es la elección del modelo apropiado para cada propósito específico.

Ejemplo 13.1

Para describir una unidad de alimentación de una caldera de vapor, pueden hacer falta varios modelos diferentes. Para planificación de la producción y control de frecuencia, puede ser suficiente caracterizar la unidad mediante dos o tres estados describiendo la energía almacenada en el colector de vapor y los recalentadores. Para construir los sistemas de seguridad y los de control, puede hacer falta un modelo con 20 a 50 estados. Finalmente, para modelar la temperatura y las tensiones en la unidad de turbina, deben emplearse varios cientos de estados.

En principio, hay dos formas diferentes de obtener los modelos: a partir de un conocimiento previo —es decir, en términos de leyes físicas— o por experimentación sobre un proceso. Cuando se intenta obtener un modelo específico, a menudo es beneficioso combinar ambos enfoques.

En la Sección 13.2 se presenta brevemente la *elaboración matemática de modelos*. En la mayor parte de los casos no es posible construir un modelo matemático a partir solamente del conocimiento físico. Algunos parámetros es preciso determinarlos mediante experimentos. Este enfoque es lo que se conoce como *identificación de sistemas* y se trata en la Sección 13.3. Hay muchos métodos para analizar los datos obtenidos mediante experimentación. Una forma básica es el *principio de mínimos cuadrados* (LS, del inglés *least squares*), expuesto en la Sección 13.4. En la sección 13.5 se presentan formas recursivas para realizar los cálculos. En la Sección 13.6 se analizan dos ejemplos.

13.2. Construcción matemática del modelo

No existen métodos generales que se puedan utilizar siempre para obtener un modelo completo. Cada proceso o problema tiene sus propias características. Se pueden dar algunas líneas generales de orientación, pero en ningún caso pueden sustituir a la experiencia. La elaboración de modelos a partir de las leyes físicas requiere conocimiento y comprensión del proceso.

El principal problema cuando se construye un modelo matemático es encontrar los estados del sistema. En esencia, las variables de estado describen un almacenamiento de energía y masa en el sistema. Variables típicas que se suelen escoger como estados son posiciones y velocidades (sistemas mecánicos); tensiones e intensidades (sistemas eléctricos); niveles y flujos (sistemas hidráulicos), y temperaturas, presiones y densidades (sistemas térmicos). Las relaciones entre los estados se determina mediante ecuaciones de balance para fuerza, momento, masa, energía y ecuaciones de composición.

La ventaja de la construcción de modelos a partir de la Física es que ayuda a comprender el problema; también, los diferentes parámetros y variables tienen una interpretación física. La desventaja es que puede ser difícil y llevar mucho tiempo hacer el modelo a partir de principios físicos. A menudo la construcción matemática de modelos debe combinarse con experimentos. Las referencias dan un tratamiento más detallado de la construcción matemática de modelos.

13.3. Identificación de sistemas

La identificación de sistemas es un planteamiento experimental de la construcción de modelos. La identificación de sistemas incluye lo siguiente:

- Planificación experimental.
- Selección de la estructura del modelo.
- Estimación de parámetros.
- Validación.

En la práctica, el procedimiento de identificación de sistemas es iterativo. Cuando se investiga un proceso del que se tiene un conocimiento previo pobre, lo razonable es comenzar con un análisis de transitorios o de respuesta en frecuencia para lograr una estimación grosera de la dinámica y de las perturbaciones. Entonces se pueden utilizar los resultados para planificar posteriores experiencias. Basándose en los resultados se puede mejorar la estructura del modelo y pueden ser necesarios nuevos experimentos.

Planificación experimental

Frecuentemente es difícil y costoso realizar experimentos con procesos industriales. Sin embargo, es de desear disponer de métodos de identificación que no requieren señales de entrada especiales. Muchos métodos «clásicos» dependen en gran medida de que la entrada esté en una forma específica, por ejemplo, sinusoides o impulsos. Otras técnicas pueden utilizar prácticamente cualquier tipo de señal de entrada, a cambio de muchos más cálculos. Un requisito de la señal de entrada es que debe excitar de forma suficiente todos los modos del proceso. Un buen método de identificación debe, por tanto, ser insensible a las características de la señal de entrada.

Algunas veces se puede basar la identificación de sistemas en datos de los procesos obtenidos bajo control en lazo cerrado. Esto es útil desde el punto de vista de las aplicaciones. Por ejemplo, los controladores adaptativos se basan en su mayor parte en identificación en lazo cerrado. La principal dificultad que tienen los datos de un proceso obtenidos cuando existe una realimentación es que puede ser imposible determinar todos los parámetros del modelo deseado; es decir, el sistema no es identificable, incluso aunque se puedan determinar todos los parámetros en un experimento en lazo abierto. La identificabilidad se puede recuperar si la realimentación es suficientemente compleja. Por esto supone una ayuda el hacer la realimentación no lineal y variable y cambiar las referencias.

Estructuras del modelo

Las estructuras del modelo se sacan del conocimiento previo del proceso y de las perturbaciones. En algunos casos el único conocimiento previo que se tiene es que el proceso se puede describir como un sistema lineal en un rango de operación concreto. Entonces es natural utilizar representaciones de sistemas lineales de tipo general. Estas representaciones se conocen como *modelos de caja negra*. Un ejemplo típico es el modelo de ecuación en diferencias

$$A(q)y(k) = B(q)u(k) + C(q)e(k) \quad (13.1)$$

donde u es la entrada, y es la salida y e es una perturbación de tipo ruido blanco. Los parámetros, así como el orden de los modelos, se consideran como parámetros desconocidos.

Algunas veces es posible aplicar leyes físicas para obtener modelos de procesos que contienen solamente algunos parámetros de valor desconocido. El modelo puede ponerse en la forma

$$\begin{cases} \dot{x} = f(x, u, v, \theta) \\ y = g(x, u, e, \theta) \end{cases} \quad (13.2)$$

donde θ es un vector de los parámetros desconocidos, x es el estado del sistema y v y e son perturbaciones.

Criterios

Cuando se formula un problema de identificación se introduce un criterio para tener una medida de hasta qué punto un modelo se ajusta a los datos experimentales. El criterio se puede postular. Mediante hipótesis estadísticas se pueden deducir criterios de base probabilística. Con frecuencia los criterios para los sistemas discretos se expresan en la forma

$$J(\theta) = \sum_{k=1}^N g(\epsilon(k))$$

donde ϵ es el error de entrada, el error de salida o un error generalizado. El error de predicción es un ejemplo típico de error generalizado. Frecuentemente se elige la función g de tipo cuadrático, pero puede ser de muchas otras formas.

La primera formulación, solución y aplicación de un problema de identificación se debe a Gauss en su famosa determinación de la órbita del asteroide Ceres. Gauss formuló el problema de identificación como un problema de optimización e introdujo el principio de los mínimos cuadrados, un método basado en la minimización de la suma de los cuadrados del error. Desde entonces, el criterio de mínimos cuadrados se ha utilizado ampliamente.

El método de mínimos cuadrados es muy sencillo y fácil de entender. En algunas circunstancias da estimaciones con valores medios erróneos (sesgados). Sin embargo,

esto se puede soslayar usando diferentes extensiones. El método de los mínimos cuadrados está restringido a estructuras de modelo que son lineales en los parámetros desconocidos.

Cuando las perturbaciones de un proceso se describen como procesos estocásticos, el problema de identificación se puede formular como un problema estadístico de estimación de parámetros. Entonces es posible emplear el método de máxima probabilidad, por ejemplo; este método tiene muchas propiedades estadísticas atractivas. Puede interpretarse como un criterio de mínimos cuadrados si se toma como cantidad a minimizar la suma de los cuadrados del error de predicción. El método de máxima probabilidad es una técnica muy general que se puede aplicar a una gran variedad de estructuras de modelo.

Métodos de estimación de parámetros

Resolver el problema de la estimación de parámetros requiere lo siguiente:

- * Datos de entrada-salida del proceso.
- * Una clase de modelos.
- * Un criterio.

La estimación de parámetros se puede formular como un problema de optimización, en el que el mejor modelo es aquel que mejor se ajusta a los datos de acuerdo con un criterio dado.

El resultado de un problema de estimación depende, naturalmente, de cómo se formule el problema. Por ejemplo, el modelo obtenido depende de la amplitud y del contenido de frecuencias de la señal de entrada. Hay muchas posibilidades de combinar condiciones experimentales, clases de modelos y criterios. También hay muchas formas distintas de organizar los cálculos. Por consiguiente, hay disponibles un gran número de métodos de identificación. Una distinción general se realiza entre métodos *en línea* (*on-line*) y *fuerza de línea* (*off-line*). Los métodos en línea dan las estimaciones en forma recursiva, cuando se obtienen las medidas, y son la única alternativa si la identificación se va a utilizar en un controlador adaptativo o si el proceso es de tipo variable. En muchos casos los métodos fuera de línea dan estimaciones con mayor precisión y son más fiables, por ejemplo, en términos de convergencia.

Para un ingeniero de la industria, que está interesado principalmente en disponer de una herramienta para obtener un modelo, la gran cantidad de métodos le induce a confusión. Se han hecho varios intentos de comparar los diferentes métodos de identificación. Las comparaciones apenas han llegado a ninguna conclusión en el sentido de que no hay un método que sea el mejor en todas las circunstancias. Afortunadamente, parece que la elección del método no es crucial. Por tanto, es aconsejable que un usuario potencial conozca los métodos clásicos (análisis de frecuencia, de la respuesta transitoria, correlación y análisis espectral), el método de mínimos cuadrados con sus extensiones, y el método de la máxima probabilidad.



Validación del modelo

Cuando se ha obtenido un modelo a partir de datos experimentales, es necesario comprobar el modelo para encontrar sus inadecuaciones. Para la validación de un modelo es útil determinar factores tales como respuesta a saltos escalón, respuesta a impulsos, polos y ceros, errores de modelo y errores de predicción. Como el propósito de la validación del modelo es explorarle para detectar sus inadecuaciones, es útil buscar cantidades que sean sensibles a cambios en el modelo.

13.4. El principio de mínimos cuadrados

Según Gauss el principio de mínimos cuadrados es tal que los parámetros desconocidos de un modelo se deben elegir de forma que

la suma de los cuadrados de las diferencias entre los valores observados realmente y los calculados multiplicada por números que midan el grado de precisión sea un mínimo.

Para ser capaces de encontrar una solución analítica, los valores calculados deben ser funciones lineales de los parámetros desconocidos. En el encuadre de la formulación general del problema de identificación dado en las secciones anteriores, la clase de modelos es lineal en los parámetros y el criterio es una función cuadrática. El propósito de esta sección es formular el problema de mínimos cuadrados y darle su solución.

El problema general

En el problema general de mínimos cuadrados se supone que «la variable calculada», \hat{y} , en la terminología de Gauss viene dada por el modelo

$$\hat{y} = \theta_1\varphi_1(x) + \theta_2\varphi_2(x) + \cdots + \theta_n\varphi_n(x) \quad (13.3)$$

donde $\varphi_1, \varphi_2, \dots, \varphi_n$ son funciones conocidas y $\theta_1, \theta_2, \dots, \theta_n$ son parámetros desconocidos. Mediante un experimento se obtienen pares de observaciones $\{(x_i, y_i), i = 1, 2, \dots, N\}$. El problema es determinar los parámetros de tal forma que las variables \hat{y}_i calculadas mediante el modelo (13.3) y los valores experimentales x_i coincidan lo más posible con las variables medidas y_i . Suponiendo que todas las medidas tienen la misma precisión, el principio de mínimos cuadrados dice que los parámetros se deben elegir de tal forma que la función de coste

$$J(\theta) = \frac{1}{2} \sum_{i=1}^N \epsilon_i^2$$

sea mínima, donde

$$\epsilon_i = y_i - \hat{y}_i = y_i - \theta_1\varphi_1(x_i) - \cdots - \theta_n\varphi_n(x_i) \quad i = 1, 2, \dots, N$$

Para simplificar los cálculos, se introduce la siguiente notación vectorial:

$$\begin{aligned}\varphi &= [\varphi_1 \quad \varphi_2 \quad \cdots \quad \varphi_n] \\ \theta &= [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_n]^T \\ y &= [y_1 \quad y_2 \quad \cdots \quad y_N]^T \\ \epsilon &= [\epsilon_1 \quad \epsilon_2 \quad \cdots \quad \epsilon_N]^T \\ \Phi &= \begin{bmatrix} \varphi(x_1) \\ \vdots \\ \varphi(x_N) \end{bmatrix}\end{aligned}$$

Ahora se puede formular el problema de mínimos cuadrados en una forma compacta. La función de coste J se puede escribir

$$J(\theta) = \frac{1}{2}\epsilon^T\epsilon = \frac{1}{2}\|\epsilon\|^2 \quad (13.4)$$

donde

$$\epsilon = y - \hat{y} \quad (13.5)$$

y

$$\hat{y} = \Phi\theta$$

Hay que determinar el parámetro θ de tal forma que $\|\epsilon\|^2$ sea mínimo. La solución al problema de mínimos cuadrados se da en el Teorema 13.1.

Teorema 13.1. La función (13.4) es mínima para los parámetros $\hat{\theta}$ tales que

$$\Phi^T\Phi\hat{\theta} = \Phi^Ty \quad (13.6)$$

Si la matriz $\Phi^T\Phi$ es no singular, el mínimo es único y está dado por

$$\hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^Ty = \Phi^Ty \quad (13.7)$$

Demostración. La función de coste (13.4) puede escribirse en la forma

$$\begin{aligned}2J(\theta) &= \epsilon^T\epsilon = [y - \Phi\theta]^T[y - \Phi\theta] \\ &= y^Ty - y^T\Phi\theta - \theta^T\Phi^Ty + \theta^T\Phi^T\Phi\theta.\end{aligned}$$

Como la matriz $\Phi^T\Phi$ siempre es definida no negativa, la función J tiene un mínimo. Empleando (11.15) el mínimo se obtiene para

$$\theta = \hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^Ty$$

con lo que se ha demostrado el teorema. \square

Nota 1. La ecuación (13.6) recibe el nombre de *ecuación normal*.

Nota 2. La matriz $\Phi^\dagger = (\Phi^T \Phi)^{-1} \Phi^T$ se conoce como la *pseudoinversa* de Φ si la matriz $\Phi^T \Phi$ es no singular.

Identificación de sistemas

El método de mínimos cuadrados se puede emplear para identificar parámetros en sistemas dinámicos. Sea el sistema descrito por (13.1) con $C(q)=1$. Además, supongamos que A y B son de orden n y $n-1$, respectivamente. Supongamos que se ha aplicado al sistema una secuencia de entradas $\{u(1), u(2), \dots, u(N)\}$ y que se ha observado la correspondiente secuencia de salida $\{y(1), y(2), \dots, y(N)\}$. Entonces los parámetros desconocidos son

$$\theta = [a_1 \dots a_n \ b_1 \dots b_n]^T \quad (13.8)$$

Además

$$\varphi(k+1) = [-y(k) \dots -y(k-n+1) \ u(k) \dots u(k-n+1)] \quad (13.9)$$

y

$$\Phi = \begin{bmatrix} \varphi(n+1) \\ \vdots \\ \varphi(N) \end{bmatrix}$$

Entonces la estimación de mínimos cuadrados es la dada por (13.7) si $\Phi^T \Phi$ es no singular. Por ejemplo, este es el caso si la señal de entrada es, hablando en términos generales, suficientemente rica.

Ejemplo 13.2

Determinar la estimación de mínimos cuadrados de los parámetros a y b en el modelo

$$\hat{y}(k) = -ay(k-1) + bu(k-1)$$

de tal forma que el criterio

$$J(a, b) = \frac{1}{2} \sum_{k=2}^N \epsilon(k)^2$$

sea mínimo, donde

$$\begin{aligned} \epsilon(k) &= y(k) - \hat{y}(k) = y(k) + ay(k-1) - bu(k-1) \\ &= y(k) - \varphi(k)\theta \end{aligned}$$

Comparándolo con el caso general se tiene

$$y = \begin{bmatrix} y(2) \\ y(3) \\ \vdots \\ y(N) \end{bmatrix}, \quad \Phi = \begin{bmatrix} -y(1) & u(1) \\ -y(2) & u(2) \\ \vdots & \vdots \\ -y(N-1) & u(N-1) \end{bmatrix}, \quad \epsilon = \begin{bmatrix} \epsilon(2) \\ \epsilon(3) \\ \vdots \\ \epsilon(N) \end{bmatrix}$$

y

$$\theta = [a \ b]^T$$

Por tanto,

$$\begin{aligned} \Phi^T \Phi &= \begin{bmatrix} \sum_{k=1}^{N-1} y(k)^2 & \sum_{k=1}^{N-1} y(k)u(k) \\ \sum_{k=1}^{N-1} y(k)u(k) & \sum_{k=1}^{N-1} u(k)^2 \end{bmatrix} \\ \Phi^T y &= \begin{bmatrix} \sum_{k=1}^{N-1} y(k+1)y(k) \\ \sum_{k=1}^{N-1} y(k+1)u(k) \end{bmatrix} \end{aligned}$$

Suponiendo que la matriz $\Phi^T \Phi$ es no singular, se puede obtener fácilmente la estimación de mínimos cuadrados de los parámetros a y b . La matriz $\Phi^T \Phi$ será no singular si se imponen condiciones especiales (*suficiente riqueza o excitación persistente*) a la señal de entrada.

Interpretación estadística

Para analizar las propiedades del estimador de mínimos cuadrados es necesario hacer algunas hipótesis. Sean los datos generados por el proceso

$$y = \Phi\theta_0 + \epsilon \quad (13.10)$$

donde θ_0 es el vector de parámetros «verdaderos» y ϵ es un vector de ruido con valor de media cero. Se enuncia el siguiente teorema sin demostración.

Teorema 13.2. Considerese la estimación (13.7) y supóngase que se generan los datos mediante (13.10) donde ϵ_i es un ruido blanco con varianza σ^2 . Entonces, si n es el número de parámetros de $\hat{\theta}$ y θ_0 y N es el número de datos, se verifican las siguientes condiciones.

1. $E\hat{\theta} = \theta_0$.
2. $\text{Var } \hat{\theta} = \sigma^2(\Phi^T \Phi)^{-1}$.
3. $s^2 = 2J(\hat{\theta})/(N-n)$ es una estimación no sesgada de σ^2 .

El Teorema 13.2 implica que los parámetros en (13.1) se pueden estimar sin sesgo si $C(q)=q^n$. Si $C(q) \neq q^n$, entonces las estimaciones estarán sesgadas. Esto es debido a la correlación entre el ruido $C^*(q^{-1})\epsilon(k)$ y los datos en $\varphi(k)$.

Extensiones del método de mínimos cuadrados

El método de mínimos cuadrados da resultados no sesgados de los parámetros en (13.1) sólo si $C(q)=q^n$. Sin embargo, se puede utilizar el método de máxima probabilidad para el caso general. Se puede demostrar que maximizar la función de

probabilidad es equivalente a minimizar la función de coste (13.4), donde los residuos, ϵ , están relacionados con las entradas y las salidas por la expresión

$$C(q)\epsilon(k) = A(q)y(k) - B(q)u(k)$$

Los residuos se pueden interpretar como el error de predicción con un paso de antelación. Sin embargo, la función de coste no es lineal en los parámetros y tiene que minimizarse numéricamente. Se puede hacer esto empleando una rutina de gradiente de Newton-Raphson, que incluye el cálculo del gradiente de J con respecto a los parámetros, así como la matriz de derivadas parciales segundas. El método de máxima probabilidad es, por tanto, un método de tipo fuera de línea. Se pueden hacer aproximaciones del método de máxima probabilidad que permitan el cálculo en línea de los parámetros del modelo (13.1). Algunos métodos corrientes son Mínimos Cuadrados Ampliado (ELS, por Extended Least Squares) y Máxima Probabilidad Recursivo (RML, por Recursive Maximum Likelihood).

13.5. Cálculos recursivos

En muchos casos las observaciones se obtienen de forma secuencial. Puede desearse calcular la estimación de mínimos cuadrados para diferentes valores de N . Si se ha resuelto el problema de mínimos cuadrados para N observaciones, parece un desperdicio de recursos de cálculo comenzar desde el principio cuando se ha hecho una nueva observación. Por tanto, se desea organizar los cálculos de tal forma que se puedan utilizar los resultados obtenidos para N observaciones con el fin de calcular las estimaciones para $N+1$ observaciones. Un problema similar se presenta cuando no se conoce previamente el número de parámetros. En este caso puede ser necesaria la estimación LS para un número diferente de parámetros. En esta sección se intenta la posibilidad de calcular la estimación de mínimos cuadrados de forma recursiva.

Recursión en el número de observaciones

Se pueden obtener ecuaciones recursivas para el caso en que las observaciones se consigan de forma secuencial. A menudo se hace referencia a este procedimiento con el nombre de *identificación recursiva*. Se puede volver a escribir la solución dada en (13.7) al problema LS para disponer de ecuaciones recursivas. Representemos por $\hat{\theta}(N)$ la estimación de mínimos cuadrados basada en N medidas. Para deducir las ecuaciones se introduce N como un parámetro formal en las funciones; es decir,

$$\Phi(N) = \begin{bmatrix} \varphi(x_1) \\ \vdots \\ \varphi(x_N) \end{bmatrix}, \quad y(N) = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$$

Se supone que la matriz $\Phi^T\Phi$ es regular para todo N . Entonces la estimación de mínimos cuadrados $\hat{\theta}(N)$ viene dada por la ecuación (13.7):

$$\hat{\theta}(N) = [\Phi^T(N)\Phi(N)]^{-1}\Phi^T(N)y(N)$$

Cuando se obtiene una medida adicional se añade una fila a la matriz Φ y un elemento al vector y . Por tanto,

$$\Phi(N+1) = \begin{bmatrix} \Phi(N) \\ \varphi(N+1) \end{bmatrix}, \quad y(N+1) = \begin{bmatrix} y(N) \\ y_{N+1} \end{bmatrix} \quad (13.11)$$

La estimación $\hat{\theta}(N+1)$ dada por (13.7) se puede entonces escribir

$$\begin{aligned} \hat{\theta}(N+1) &= [\Phi^T(N+1)\Phi(N+1)]^{-1}\Phi^T(N+1)y(N+1) \\ &= [\Phi^T(N)\Phi(N) + \varphi^T(N+1)\varphi(N+1)]^{-1}[\Phi^T(N)y(N) + \varphi^T(N+1)y_{N+1}] \end{aligned} \quad (13.12)$$

La solución se da en el siguiente teorema.

Teorema 13.3 – Estimación de mínimos cuadrados recursiva. Supóngase que la matriz $\Phi^T(N)\Phi(N)$ es definida positiva. La estimación de mínimos cuadrados $\hat{\theta}$ satisface la ecuación recursiva

$$\hat{\theta}(N+1) = \hat{\theta}(N) + K(N)[y_{N+1} - \varphi(N+1)\hat{\theta}(N)] \quad (13.13)$$

$$\begin{aligned} K(N) &= P(N+1)\varphi^T(N+1) \\ &= P(N)\varphi^T(N+1)[1 + \varphi(N+1)P(N)\varphi^T(N+1)]^{-1} \end{aligned} \quad (13.14)$$

$$P(N+1) = [I - K(N)\varphi(N+1)]P(N) \quad (13.15)$$

El siguiente lema es útil para demostrar el teorema:

Lema 13.1 – Lema de inversión de una matriz. Sean las matrices no singulares A , C y $C^{-1} + DA^{-1}B$; entonces

$$[A + BCD]^{-1} = A^{-1} - A^{-1}B[C^{-1} + DA^{-1}B]^{-1}DA^{-1}$$

Demostración. Por sustitución directa,

$$\begin{aligned} &[A + BCD][A^{-1} - A^{-1}B[C^{-1} + DA^{-1}B]^{-1}DA^{-1}] \\ &= I + BCDA^{-1} - B[C^{-1} + DA^{-1}B]^{-1}DA^{-1} - BCDA^{-1}B[C^{-1} + DA^{-1}B]^{-1}DA^{-1} \\ &= I + BCDA^{-1} - BC[C^{-1} - DA^{-1}B][C^{-1} + DA^{-1}B]^{-1}DA^{-1} \\ &= I + BCDA^{-1} - BCDA^{-1} \\ &= I \end{aligned}$$

Demostración del Teorema 13.3. Para simplificar la notación en las manipulaciones que siguen se suprimirá el argumento N de $\Phi(N)$ e $y(N)$ y el argumento $N+1$ de $\varphi(N+1)$. Entonces la ecuación (13.12) se puede escribir

$$\begin{aligned}\hat{\theta}(N+1) &= [\Phi^T\Phi + \varphi^T\varphi]^{-1}[\Phi^Ty + \varphi^Ty_{N+1}] \\ &= (\Phi^T\Phi)^{-1}\Phi^Ty + [(\Phi^T\Phi + \varphi^T\varphi)^{-1} - (\Phi^T\Phi)^{-1}]\Phi^Ty + (\Phi^T\Phi + \varphi^T\varphi)^{-1}\varphi^Ty_{N+1}\end{aligned}\quad (13.16)$$

Obsérvese que

$$\hat{\theta}(N) = (\Phi^T\Phi)^{-1}\Phi^Ty$$

y

$$\begin{aligned}[(\Phi^T\Phi + \varphi^T\varphi)^{-1} - (\Phi^T\Phi)^{-1}]\Phi^Ty &= (\Phi^T\Phi + \varphi^T\varphi)^{-1}(\Phi^T\Phi - \Phi^T\Phi - \varphi^T\varphi)(\Phi^T\Phi)^{-1}\Phi^Ty \\ &= -(\Phi^T\Phi + \varphi^T\varphi)^{-1}\varphi^T\varphi(\Phi^T\Phi)^{-1}\Phi^Ty \\ &= -(\Phi^T\Phi + \varphi^T\varphi)^{-1}\varphi^T\varphi\theta\end{aligned}$$

La ecuación (13.16) se puede escribir

$$\hat{\theta}(N+1) = \hat{\theta}(N) + K(N)[y_{N+1} - \varphi(N+1)\hat{\theta}(N)]$$

donde

$$\begin{aligned}K(N) &= [\Phi^T(N)\Phi(N) + \varphi^T(N+1)\varphi(N+1)]^{-1}\varphi^T(N+1) \\ &= [\Phi^T(N+1)\Phi(N+1)]^{-1}\varphi^T(N+1)\end{aligned}$$

Para obtener una ecuación recursiva para el factor de peso $K(N)$ conviene introducir la cantidad P definida por

$$P(N) = [\Phi^T(N)\Phi(N)]^{-1}$$

P es proporcional a la varianza de las estimaciones (compárese con el Teorema 13.2). Aplicando el Lema 13.1 a la matriz $P(N+1)$ se tiene

$$\begin{aligned}P(N+1) &= [\Phi^T(N+1)\Phi(N+1)]^{-1} = [\Phi^T\Phi + \varphi^T\varphi]^{-1} \\ &= (\Phi^T\Phi)^{-1} - (\Phi^T\Phi)^{-1}\varphi^T[I + \varphi(\Phi^T\Phi)^{-1}\varphi^T]^{-1}\varphi(\Phi^T\Phi)^{-1}\end{aligned}$$

Por tanto,

$$P(N+1) = P(N) - P(N)\varphi^T(N+1)[I + \varphi(N+1)P(N)\varphi^T(N+1)]^{-1}\varphi(N+1)P(N)$$

Ahora unos sencillos cálculos dan

$$\begin{aligned}K(N) &= P(N+1)\varphi^T(N+1) \\ &= P(N)\varphi^T(N+1)[I + \varphi(N+1)P(N)\varphi^T(N+1)]^{-1}\end{aligned}$$

Nótese que la inversión de la matriz es necesaria para calcular P . Sin embargo, la matriz que hay que invertir es de la misma dimensión que el número de medidas; es decir, para un sistema de una sola salida es un escalar. \square

Nota 1. La ecuación (13.13) tiene una apariencia muy intuitiva. La estimación $\hat{\theta}(N+1)$ se obtiene añadiendo una corrección a la estimación anterior $\hat{\theta}(N)$. La corrección es proporcional a $y_{N+1} - \varphi(N+1)\hat{\theta}(N)$, donde el último término se puede interpretar como el valor de y en el instante $N+1$ predicho por el modelo. Entonces el término de corrección es proporcional a la diferencia entre el valor medido de y_{N+1} y la predicción de y_{N+1} basada en las estimaciones previas de los parámetros. Los componentes del vector $K(N)$ son factores de ponderación que nos dicen cómo hay que combinar la corrección y las estimaciones previas. Nótese que la componente $K_i(N)$ es proporcional a $\varphi_i(N+1)$.

Nota 2. La estimación de mínimos cuadrados se puede interpretar como un filtro de Kalman para el proceso

$$\begin{aligned}\theta(k+1) &= \theta(k) \\ y(k) &= \varphi(k)\theta(k) + e(k)\end{aligned}$$

Véase la Sección 11.3.

Nótese que la matriz $P(N)$ está definida sólo cuando la matriz $\Phi^T(N)\Phi(N)$ es no singular. Como

$$\Phi^T(N)\Phi(N) = \sum_{k=1}^N \varphi^T(k)\varphi(k)$$

se deduce que $\Phi^T\Phi$ será siempre singular si N es suficientemente pequeño. Para obtener una condición inicial para P es necesario escoger un $N=N_0$ tal que $\Phi^T(N_0)\Phi(N_0)$ sea no singular y determine

$$\begin{aligned}P(N_0) &= [\Phi^T(N_0)\Phi(N_0)]^{-1} \\ \hat{\theta}(N_0) &= P(N_0)\Phi^T(N_0)y(N_0)\end{aligned}$$

Las ecuaciones recursivas se pueden utilizar para $N \geq N_0$. Sin embargo, a menudo es conveniente utilizar las ecuaciones recursivas en todos los pasos. Si se comienzan las ecuaciones recursivas con la condición inicial

$$P(0) = P_0$$

donde P_0 es definida positiva, entonces

$$P(N) = [P_0^{-1} + \Phi^T(N)\Phi(N)]^{-1}$$

Esto puede hacerse arbitrariamente próximo a $[\Phi^T(N)\Phi(N)]^{-1}$ con sólo escoger P_0 suficientemente grande.

Utilizando la interpretación estadística del método de mínimos cuadrados se ve que esta forma de comenzar la recursión corresponde a la situación en que los parámetros tienen una covarianza previa proporcional a P_0 .

Sistemas variables con el tiempo

Utilizando la función de coste (13.4) a todos los puntos de datos se les da el mismo peso. Si los parámetros son variables, es necesario eliminar la influencia de los datos antiguos. Esto se puede hacer empleando una función de coste con ponderación exponencial; es decir,

$$J(\theta) = \sum_{k=1}^N \lambda^{N-k} [y(k) - \theta\varphi(k)]^2 \quad (13.17)$$

El «factor de olvido», λ , es menor que la unidad y es una medida de la rapidez con que se olvidan los datos antiguos. La estimación de mínimos cuadrados cuando se utiliza la función de coste (13.17) viene dada por

$$\begin{aligned}\hat{\theta}(k+1) &= \hat{\theta}(k) + K(k)[y_{k+1} - \varphi(k+1)\hat{\theta}(k)] \\ K(k) &= P(k)\varphi^T(k+1)[\lambda + \varphi(k+1)P(k)\varphi^T(k+1)]^{-1} \\ P(k+1) &= [I - K(k)\varphi(k+1)]P(k)/\lambda\end{aligned}\quad (13.18)$$

También es posible modelar los parámetros variables por un proceso de Markov,

$$\theta(k+1) = \Phi\theta(k) + v(k)$$

y entonces utilizar un filtro de Kalman para estimar θ . Véase la nota 2 del Teorema 13.3.

Recursión en el número de parámetros

Cuando se introducen parámetros extra, el vector $\hat{\theta}$ tendrá más componentes y habrá filas adicionales en la matriz Φ . Se pueden organizar los cálculos de forma que sea posible realizar una recursión en el número de parámetros en el modelo. La recursión entraña una inversión de una matriz de la misma dimensión que el número de parámetros añadidos.

Factorización de covarianza $U-D$

La ecuación (13.18) es una forma de mecanizar la actualización recursiva de las estimaciones y de la matriz de covarianza. Sin embargo, estas ecuaciones no están bien planteadas, desde un punto de vista numérico. Una forma mejor de realizar los cálculos es actualizar la raíz cuadrada de P en lugar de actualizar P . Otra forma de hacer los cálculos es utilizar el algoritmo $U-D$ de Bierman y Thornton. Este método se basa en una factorización de P de la forma

$$P = UD U^T$$

donde D es diagonal y U es una matriz triangular superior. Este método es de tipo raíz cuadrada porque $UD^{1/2}$ es la raíz cuadrada de P . El método de factorización $U-D$ no incluye cálculos de raíces cuadradas y por tanto es muy apropiado para pequeños computadores y aplicaciones en tiempo real. En las referencias se pueden encontrar más detalles sobre el algoritmo.

En el Listado 13.1 se da un programa en Pascal para la estimación basado en la factorización $U-D$. El programa da la estimación de los parámetros del proceso

$$y(k) + a_1y(k-1) + \cdots + a_ny(k-na) = b_1u(k-1) + \cdots + b_nu(k-nb) + e(k) \quad (13.19)$$

LISTADO 13.1 Programa en Pascal para la estimación por mínimos cuadrados de los parámetros del proceso (13.19) empleando la factorización $U-D$.

```
const npar=10; {número máximo de parámetros estimados}
noff=45; {noff=npar*(npar-1)/2}
type vec1=array[1..npar] of real;
vec2=array[1..noff] of real;
estpartyp = record
  n, na:integer;
  theta:vec1;
  fi:vec1;
  diag:vec1;
  offdiag:vec2;
end;
var y,u,lambda:real;
eststate:estpartyp;

Procedure LS(u,y,lambda:real; var eststate:estpartyp);
{Calcula la estimación por mínimos cuadrados utilizando el método U-D
según Bierman y Thornton}
var kf,ku,i,j:integer;
perr,fj,vj,alphaj,ajlast,pj,w:real;
k:vec1;

begin
  with eststate do {Calcular el error de predicción}
begin
  perr:=y;
  for i:=1 to n do perr:=perr-theta[i]*fi[i];
  {Calcular ganancia y covarianza utilizando el método U-D}
  fj:=fi[1];
  vj:=diag[1]*fj;
  k[1]:=vj;
  alphaj:=1.0+vj*fj;
  diag[1]:=diag[1]/alphaj/lambda;
  if n>1 then
```

LISTADO 13.1 (CONT.)

```

begin
  kf:=0;
  ku:=0;
  for j:=2 to n do
  begin
    fj:=fi[j];
    for i:=1 to j-1 do
    begin {f:=fi*i}
      kf:=kf+1;
      fj:=fj+fi[i]*offdiag[kf]
    end; {i}
    vj:=fj*diag[j];{v=D*f}
    k[j]:=vj;
    ajlast:=alphaj;
    alphaj:=ajlast+vj*fj;
    diag[j]:=diag[j]*ajlast/alphaj/lambda;
    pj:=-fj/ajlast;
    for i:=1 to j-1 do
    begin
      {kj+1:=kj+vj*u(j)}
      {uj:=uj+pj*kj}
      ku:=ku+1;
      w:=offdiag[ku]+k[i]*pj;
      k[i]:=kj+offdiag[ku]*vj;
      offdiag[ku]:=w
    end; {i}
  end; {j}
end; {si n>1 entonces}
{Actualizar los parámetros estimados}
for i:=1 to n do theta[i]:=theta[i]+perr*k[i]/alphaj;
{Actualización de fi}
for i:=1 to n-1 do fi[n+1-i]:=fi[n-i];
fi[1]:=-y;
fi[n+1]:=u
end {con eststate hacer}
end; {LS}

```

Las notaciones usadas en el programa son

Variable	Notación en el programa
$u(k)$	u
$y(k)$	y
na	na
$na+nb$	n
$n(n-1)/2$	noff
$\hat{\theta}(k)$ compárese con (13.8)	theta
$\varphi(k)$ compárese con (13.9)	fi
λ	lambda

13.6. Ejemplos

Dos ejemplos mostrarán el empleo de los métodos de identificación.

Ejemplo 13.3

Sea el sistema descrito por el modelo

$$\begin{aligned}
 y(k) &= 1.5y(k-1) + 0.7y(k-2) \\
 &= u(k-1) + 0.5u(k-2) + e(k) - e(k-1) + 0.2e(k-2)
 \end{aligned} \tag{13.20}$$

donde e tiene media cero y desviación estándar 0,5. Se trata de un sistema «típico» que se ha empleado mucho en la literatura para probar diferentes métodos de identificación. En (13.20), $C(q) \neq q^n$, lo que implica que el método de mínimos cuadrados dará una estimación sesgada. Sin embargo, la relación entrada-salida del proceso se puede aproximar usando el método de mínimos cuadrados para un modelo de orden superior. La Figura 13.1 muestra una simulación del sistema. La entrada es una secuencia de señal

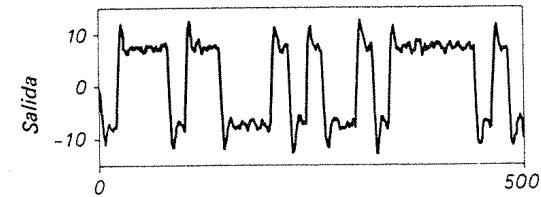
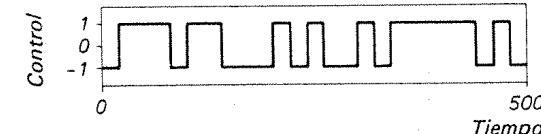


Figura 13.1 Entrada y salida cuando se simula el sistema (13.20). La entrada es una secuencia PRBS.



binaria pseudo-aleatoria (PRBS, por Pseudo Random Binary Signal) con amplitud ± 1 . Se han usado los datos para identificar modelos de diferentes órdenes mediante los métodos de mínimos cuadrados y de máxima probabilidad. La Figura 13.2 muestra las respuestas a un salto escalón del sistema real (13.20) junto con la de los modelos estimados cuando se emplea el método de mínimos cuadrados con modelos de órdenes $n = 1, 2$ y 4 y el método de máxima probabilidad con un modelo de orden 2. El método de mínimos cuadrados da una descripción pobre para un modelo de segundo orden, mientras que se obtiene un buen modelo cuando el orden se aumenta a cuatro. El método de máxima probabilidad da muy buenas estimaciones de las características dinámica y de ruido para un modelo de segundo orden. En la Tabla 13.1 se pueden ver los parámetros estimados para modelos de segundo orden usando los métodos de mínimos cuadrados y de máxima probabilidad.

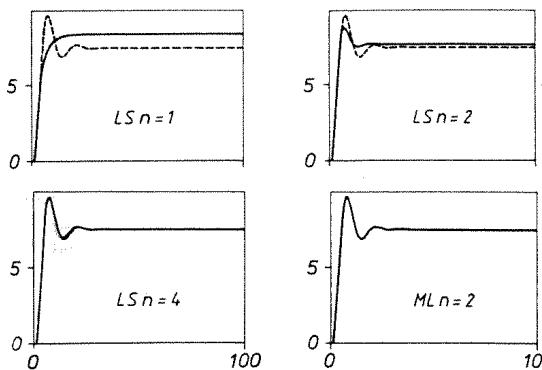


Figura 13.2 Respuestas salto de la parte determinística del sistema (13.20) y de los modelos estimados obtenidos cuando se utiliza el método de mínimos cuadrados (LS) con $n=1$, 2 y 4, y el método de máxima probabilidad (ML) con $n=2$.

TABLA 13.1 Parámetros estimados y desviaciones estándar para modelos de segundo orden del proceso (13.20) cuando se utilizan los métodos de mínimos cuadrados (LS) y de máxima probabilidad (ML).

Parámetro	Valor real	LS $n=2$	ML $n=2$
a_1	-1,5	$-1,285 \pm 0,027$	$-1,497 \pm 0,009$
a_2	0,7	$0,540 \pm 0,021$	$0,699 \pm 0,006$
b_1	1	$1,056 \pm 0,091$	$1,019 \pm 0,051$
b_2	0,5	$0,913 \pm 0,121$	$0,497 \pm 0,075$
c_1	-1	—	$-0,964 \pm 0,045$
c_2	0,2	—	$0,174 \pm 0,044$

Ejemplo 13.4 – Estimación recursiva

Considérese el proceso

$$y(k) + ay(k-1) = bu(k-1) + e(k)$$

con $a = -0,8$ y $b = 1$. La varianza del ruido e es 1. Se supone que la señal de entrada es una señal PRBS con amplitud ± 1 . En la Figura 13.3 se muestran los datos de entrada y

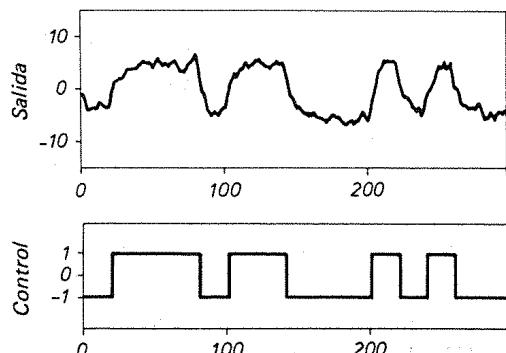


Figura 13.3 Datos de entrada-salida para el proceso del Ejemplo 13.4.

de salida. Para estimar a y b se han usado las ecuaciones recursivas (13.13)-(13.15). En la Figura 13.4 se pueden ver las estimaciones cuando los valores iniciales de los parámetros se han tomado cero y cuando $P(0)$ es 10 veces la matriz unidad. Después de unas pocas observaciones, las estimaciones están muy próximas a los valores reales.

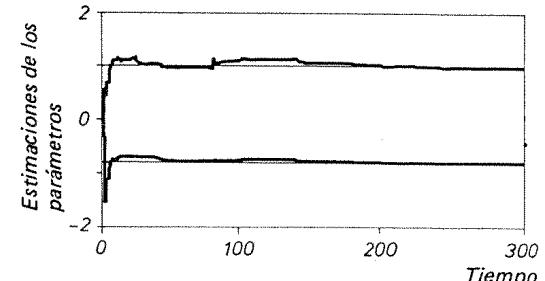


Figura 13.4 Estimación recursiva de los parámetros cuando se emplean (13.13)-(13.15) sobre los datos de la Figura 13.3.

13.7. Resumen

Este capítulo da una rápida presentación del problema de la identificación. La exposición se ha concentrado en el método de mínimos cuadrados, porque es la base de muchos otros métodos. En muchas ocasiones es importante hacer las estimaciones en tiempo real, y se ha visto cómo se puede obtener la estimación de mínimos cuadrados de forma recursiva. En el Capítulo 14 se la usará en conexión con los controladores adaptativos.

13.8. Problemas

- 13.1. Se ha realizado el siguiente experimento para determinar la aceleración de la gravedad, g . Se ha dejado caer con velocidad inicial nula una bola de acero desde una antena de televisión muy elevada. Se ha determinado la posición, l , de la bola en diferentes instantes de tiempo, obteniendo las siguientes medidas:

Tiempo, s	Longitud de caída, en metros
1	8,49
2	20,05
3	50,65
4	72,19
5	129,85
6	171,56

Los tiempos de las medidas son exactos, pero hay un error en la medida de la posición. Determíñese la aceleración de la gravedad mediante el método de mínimos cuadrados a partir del modelo

$$l = \frac{gt^2}{2} + e$$



- 13.2. Dedúzcanse las ecuaciones recursivas para un aumento en el número de parámetros para el método de mínimos cuadrados. (*Sugerencia:* Utilícese la misma idea que cuando se hacen las observaciones de forma recursiva.)

- 13.3. Considérese el proceso

$$y(k) + ay(k-1) = bu(k-1) + e(k) + ce(k-1)$$

donde u y e son procesos ruido blanco independientes con media cero y varianza unidad. Supóngase que se emplea el método de mínimos cuadrados para estimar a y b como en el Ejemplo 13.2. Determinense los valores esperados de \hat{a} y \hat{b} como funciones de a , b y c .

- 13.4. En el sistema

$$y(k) = b_1 u(k-1) + b_2 u(k-2) + e(k)$$

se determinan los parámetros b_1 y b_2 mediante el método de mínimos cuadrados. Sea la entrada un salto escalón en el instante $k=0$. ¿Se pueden determinar los parámetros b_1 y b_2 con precisión arbitraria cuando aumenta el número de observaciones? ¿Se producirá algún cambio si se sabe que $b_2=0$?

- 13.5. El matemático americano Richardson ha propuesto el siguiente modelo sencillo para una carrera de armamento entre dos países:

$$x(k+1) = ax(k) + by(k) + f$$

$$y(k+1) = cx(k) + dy(k) + g$$

donde $x(k)$ e $y(k)$ son los gastos en armas de las dos naciones y a , b , c , d , f y g son constantes. Los siguientes datos han sido recopilados por SIPRI. (World Armaments and Disarmaments, SIPRI Yearbook 1982.)

Millones de dólares U.S. en precios de 1979 e índice de cambio de divisas de 1979

Año	Irán	Irak	OTAN	Pacto de Varsovia
1972	2.891	909	216.478	112.893
1973	3.982	1.123	211.146	115.020
1974	8.801	2.210	212.267	117.169
1975	11.230	2.247	210.525	119.612
1976	12.178	2.204	205.717	121.461
1977	9.867	2.303	212.009	123.561
1978	9.165	2.179	215.988	125.498
1979	5.080	2.675	218.561	127.185
1980	4.040		225.411	129.000
1981		233.957		131.595

Determinense los parámetros del modelo mediante mínimos cuadrados y estúdiese la estabilidad del modelo.

- 13.6. Considérese el modelo de la carrera de armamentos de Richardson en el Problema 13.5.

a) Determinense las estimaciones de los parámetros basándose en tres años consecutivos. Mírese la validez de las estimaciones.

b) Determinese una estimación recursiva de los parámetros. Comiéncese en 1975 con valores iniciales sacados del período 1972-74.

13.9. Referencias

Hay muchos libros y trabajos que tratan de los métodos de identificación. Algunas referencias básicas en forma de libro son

JENKINS, G. M. y D. G. WATTS (1968): *Spectral Analysis and Its Applications*. San Francisco: Holden-Day.

EYKHOFF, P. (1974): *System Identification: Parameter and State Estimation*. London: John Wiley.

BOX, G. E. P. y G. M. KENKINS (1976): *Time Series Analysis and Control* (rev. ed.). San Francisco: Holden-Day.

GOODWIN, G. C. y R. L. PAYNE (1977): *Dynamic System Identification: Experiment Design and Data Analysis*. New York: Academic Press.

LJUNG, L. y T. SÖDERSTRÖM (1983): *Theory and Practice of Recursive Identification*. Cambridge, Mass.: MIT Press.

Se puede encontrar una visión de conjunto de la identificación de sistemas en

ÅSTRÖM, K. J. y P. EYKHOFF (1971): «System Identification: A Survey.» *Automatica*, 123-62.

Son buenas fuentes para más referencias

BIERMAN, G. J. (1977): *Factorization Methods for Discrete Sequential Estimation*. New York: Academic Press.

EYKHOFF, P., ed. (1981): *Trends and Progress in System Identification*. Oxford: Pergamon Press.

ISERMANN, R., ed. (1981): «System Identification,» Tutorial presentado en el 5th IFAC Symposium on Identification and System Parameter Estimation, Darmstadt, September 1979, Oxford: Pergamon Press.

LAWSON, C. L. y R. J. HANSON (1974): *Solving Least Squares Problems*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.

Número especial sobre Identification and System Parameter Estimation, *Automatica*, 17, No. 1 (January 1981).

Para los interesados en notas históricas, véase

GAUSS, K. F. (1809): *Theoria Motus Corporum Coelestium* (en Latín). Trad. inglesa (1963): *Theory of Motion of the Heavenly Bodies*. New York: Dover.

Se pueden encontrar notas históricas adicionales en

SORENSEN, H. W. (1970): «Least-squares Estimation: From Gauss to Kalman,» *IEEE Spectrum*, 7, No. 7, 63-68.

Control adaptativo

Objetivo:

Combinar el diseño de sistemas de control y los métodos de estimación recursiva para obtener reguladores auto-sintonizados. Esbozar los principios del control adaptativo.

14.1. Introducción

La aplicación de las técnicas expuestas en los capítulos previos a un problema práctico conlleva un esfuerzo sustancial. Es necesario realizar los pasos de modelado, identificación, diseño del control y análisis de sensibilidad. Incluso puede ser necesario repetir los pasos hasta obtener un resultado satisfactorio. Tiene un considerable interés explorar las posibilidades de simplificación del procedimiento para dar entrada a reguladores más complejos. Una posibilidad consiste es tratar de automatizar el procedimiento completo. Esto debe hacerse preparando el regulador con algoritmos para estimación de parámetros y diseño del control. Este método conduce al denominado Regulador Auto-Sintonizado (STR, del inglés Self-Tuning Regulator), el cual tiene la capacidad de sintonizarse él mismo. Esto se estudia en la Sección 14.2. Tales reguladores serán más complejos que los reguladores de ganancia constante. Sin embargo, pueden realizarse adecuadamente utilizando microprocesadores.

Los sistemas en lazo cerrado que se obtienen con reguladores auto-sintonizados son sistemas no lineales y variables con el tiempo. El análisis de tales sistemas requiere métodos bastante diferentes de los desarrollados en este libro. En la Sección 14.3 se esbozan algunos problemas claves. También se ilustran las propiedades con ejemplos simples.

Pequeñas modificaciones de un regulador auto-sintonizado producen un controlador adaptativo, el cual puede manipular sistemas con amplias variaciones en los parámetros. En la Sección 14.4 se da una revisión de diferentes métodos de control

adaptativo. Esto incluye planificación de la ganancia, control adaptativo de modelo de referencia y control dual.

La experiencia en el uso de técnicas adaptativas es muy limitada. En la Sección 14.5 se dan algunas notas sobre usos y abusos.

14.2. Control auto-sintonizado

Un método de modelado y diseño automático es el siguiente: determinar una estructura adecuada para el modelo. Estimar los parámetros del modelo recursivamente utilizando los métodos expuestos en el Capítulo 13. Utilizar la estimación para calcular la ley de control mediante un método adecuado de diseño, por ejemplo, una de las técnicas presentadas en los Capítulos 9, 10, 11 ó 12. Un diagrama de bloques de este sistema se muestra en la Figura 14.1. El regulador obtenido se denomina *regulador auto-sintonizado* porque tiene capacidad para sintonizar sus propios parámetros. Puede pensarse que el regulador está compuesto de dos lazos. El lazo interno consiste en el proceso y un regulador corriente de realimentación lineal. Los parámetros del regulador se ajustan con el lazo exterior, el cual está compuesto de un estimador recursivo de parámetros y cálculos de diseño. El bloque de cálculos de diseño de la Figura 14.1 representa la solución instantánea de un problema de diseño para un sistema de parámetros conocidos.

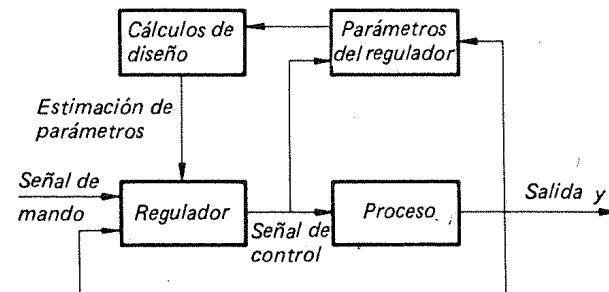


Figura 14.1 Diagrama de bloques de un regulador auto-sintonizado (STR).

El regulador auto-sintonizado es muy flexible respecto al método de diseño. Virtualmente tiene cabida cualquier técnica de diseño. Se han considerado auto-sintonizadores basados en los márgenes de fase y ganancia, asignación de polos, control de varianza mínima y control LQG. Pueden utilizarse diferentes esquemas de estimación de parámetros; por ejemplo, aproximación estocástica, mínimos cuadrados, mínimos cuadrados extendidos y generalizados, variables instrumentales, filtros de Kalman extendidos y máxima probabilidad.

Los algoritmos de auto-sintonizado pueden dividirse en dos clases principales: algoritmos implícitos y explícitos. En un algoritmo explícito hay una estimación de un

modelo *explícito* del proceso. Algunas veces es posible parametrizar el proceso para que pueda expresarse en función de los parámetros del regulador. Esto simplifica significativamente el algoritmo porque se eliminan los cálculos del diseño. Este auto-sintonizador se denomina regulador auto-sintonizado *implícito* porque se basa en la estimación de un modelo de proceso implícito.

El problema de diseño

Consideremos un problema de diseño por asignación de polos del tipo expuesto en la Sección 10.5. Supongamos que el proceso está descrito por el modelo

$$A(q)y(k) = B(q)u(k) \quad (14.1)$$

donde u es la señal de control, y es la señal medida y las perturbaciones son insignificantes. Además, $d = \text{gra } A - \text{gra } B$. Supongamos que se desea encontrar un regulador tal que la función de transferencia desde la señal de mando a la salida venga dada por

$$H_m = \frac{B_m}{A_m} \quad (14.2)$$

Se supone que el polinomio del observador es A_o . La solución al problema de diseño se da en la Sección 10.5. El regulador es

$$Ru = Tu_c - Sy \quad (14.3)$$

donde u_c es la señal de mando. Los polinomios R , S y T se obtienen resolviendo la ecuación diofántica

$$AR_1 + B^-S = A_mA_o \quad (14.4)$$

con respecto a R_1 y S . La realimentación deseada está definida por (14.3) con

$$R = R_1B^+ \quad y \quad T = A_oB'_m \quad (14.5)$$

donde $B_m = B^-B'_m$ y $B = B^-B^+$.

Hay diferentes modos de estimar recursivamente los parámetros del modelo (14.1). Una posibilidad es el método de mínimos cuadrados discutido en la Sección 13.4.

Un auto-sintonizador explícito

Puede describirse un auto-sintonizador simple con el siguiente algoritmo.

Algoritmo 14.1

Paso 1. Estimar los coeficientes de los polinomios A y B de (14.1) por mínimos cuadrados recursivos.

Paso 2. Sustituir A y B por las estimaciones obtenidas en el paso 1 y resolver la ecuación (14.4) para R_1 y S . Calcular R y T mediante (14.5).

Paso 3. Calcular la señal de control con (14.3).

Los pasos 1, 2 y 3 se repiten en cada período de muestreo.

Deben tomarse algunas precauciones cuando se utiliza este algoritmo. Para obtener buenas estimaciones es necesario que la señal de entrada al proceso sea suficientemente rica en frecuencias. Esto no siempre será así, porque la entrada se genera por realimentación. La solución de la ecuación diofántica (14.4) requiere que los polinomios A y B no tengan factores comunes.

Un auto-sintonizador implícito

Es posible eliminar el paso 2 del algoritmo 14.1 mediante reparametrización del modelo (14.1). De (14.4) se deduce que

$$A_oA_my = AR_1y + B^-Sy = BR_1u + B^-Sy = B^-[Ru + Sy] \quad (14.6)$$

donde la segunda igualdad se deduce de (14.1) y la tercera de (14.5). La ecuación (14.6) puede interpretarse como un modelo del proceso, el cual está parametrizado en B^- , R y S . Una estimación de los parámetros del modelo (14.6) da directamente los parámetros del regulador. También, el modelo (14.6) es lineal en los parámetros solamente si $B^- = 1$. El algoritmo implícito puede expresarse del siguiente modo.

Algoritmo 14.2

Paso 1. Estimar los coeficientes de los polinomios R , S y B^- de (14.6), sucesivamente.

Paso 2. Calcular la señal de control en (14.3), donde se reemplazan R y S por las estimaciones obtenidas en el paso 1. T se elige como en (14.5).

Los pasos 1 y 2 se repiten en cada período de muestreo.

En este caso también es necesario tomar algunas precauciones porque la ley de control de (14.3) no es causal si el primer coeficiente de la estimación del polinomio R es cero. Se requieren pocas modificaciones para evitar este inconveniente.

El algoritmo 14.2 se hace muy simple si $B^- = 1$, porque el problema de estimación puede resolverse por mínimos cuadrados recursivos. En esta suposición se basan la mayor parte de los auto-sintonizadores simples. De (14.5) se deduce que todos los ceros del proceso se cancelan en el diseño. Así, los algoritmos basados en esta suposición no funcionarán para procesos con inversos inestables; incluso en el caso ideal de que se conozcan los parámetros (compárese con la Sección 10.5).



14.3. Análisis

Los sistemas en lazo cerrado obtenidos con control auto-sintonizado no son lineales, lo cual dificulta la comprensión del funcionamiento. Un tratamiento razonado completo está fuera del propósito de este libro. Sin embargo, hay algunas propiedades de esta interesante e importante clase de sistemas controlados por computador dignas de discusión.

Para la exposición se necesita un modelo del proceso que se va a controlar. Se supone que el proceso está gobernado por el modelo

$$Ay = Bu + Ce \quad (14.7)$$

donde e es un ruido blanco.

Temas clave

El problema esencial consiste en averiguar cómo se comporta el sistema en lazo cerrado bajo control auto-sintonizado. Esto incluye un análisis de estabilidad, convergencia y funcionamiento. Otro objetivo importante es averiguar si un sistema de control como el auto-sintonizador es razonable o hay estructuras mejores.

Algoritmos explícitos

Un auto-sintonizador explícito, como el dado en el algoritmo 14.1, converge si los parámetros estimados convergen. Esto requiere que la estructura del modelo utilizado en la estimación sea correcta y que la señal de entrada sea suficientemente rica en frecuencias. Puesto que se utiliza el método de los mínimos cuadrados, es necesario que no haya correlación en las perturbaciones, es decir, $C = 1$ en el modelo de (14.7). Como la señal de control se genera por realimentación, no hay garantía de que sea suficientemente rica en frecuencias. Puede ser necesario introducir señales de perturbación para asegurar esto.

Análisis de un sistema de primer orden

Para ilustrar algunos principios del análisis de reguladores auto-sintonizados se considera un ejemplo de primer orden.

Supongamos que la dinámica del proceso y su entorno pueden describirse por el sistema de primer orden

$$y(k) + ay(k - 1) = bu(k - 1) + e(k) + ce(k - 1) \quad (14.8)$$

donde u es la variable de control, y es la salida y $\{e(k)\}$ es una secuencia de variables aleatorias gaussianas independientes. Además, supongamos que el criterio es minimizar la función de coste cuadrática

$$J = \lim_{N \rightarrow \infty} E \frac{1}{N} \sum_{k=1}^N y^2(k) \quad (14.9)$$

Se supone que los controladores admisibles son aquellos en los que $u(k)$ es una función de todas las salidas pasadas $y(k), y(k-1), \dots$. Si se conocen los parámetros, del Teorema 12.2 se deduce que el control óptimo es la realimentación proporcional

$$u(k) = \frac{a - c}{b} y(k) \quad (14.10)$$

Supongamos que $b = 1$ y consideremos un auto-sintonizador basado en la estimación por mínimos cuadrados del parámetro θ en el modelo

$$y(k) + \theta y(k - 1) = u(k - 1) + e(k) \quad (14.11)$$

La estimación $\hat{\theta}$ basada en los datos disponibles en el instante k —es decir, $y(k), y(k-1), \dots, y(1), u(k-1), u(k-2), \dots, u(1)$ — está definida por

$$\hat{\theta}(k) = \frac{-\left[\sum_{i=1}^{k-1} [y(i+1) - u(i)]y(i) \right]}{\sum_{i=1}^{k-1} y^2(i)} \quad (14.12)$$

La ley de control de varianza mínima para (14.11) es

$$u(k) = \theta(k)y(k)$$

si se conoce θ . Cuando no se conoce θ se reemplaza por su estimación. Entonces, la ley de control es

$$u(k) = \hat{\theta}(k)y(k) \quad (14.13)$$

Puede esperarse que el algoritmo de control definido por (14.12) y (14.13) trabaje con precisión para el sistema (14.8) si $c = 0$ y $b = 1$. En este caso la estimación por mínimos cuadrados $\hat{\theta}$ convergerá hacia a cuando $k \rightarrow \infty$, y la ley de control (14.13) convergerá a

$$u(k) = ay(k)$$

que es la ley de control deseada. Compárese con (14.10).

Una propiedad a destacar de la ley de realimentación descrita por (14.12) y (14.13) es que también converge hacia la ley óptima (14.10) cuando $c \neq 0$. Esto se ilustra en la Figura 14.2, la cual muestra la estimación de parámetros $\hat{\theta}$ para $a = -0.9$, $b = 3$ y $c = -0.3$.

Obsérvese que la estimación $\hat{\theta}$ parece converger a $\hat{\theta} = -0.2$ y no al valor de a , el cual es -0.9 . El valor -0.2 corresponde al control de varianza mínima. Para comparar el regulador auto-sintonizado con el regulador óptimo de parámetros conocidos se ha calculado la función de coste acumulada

$$J(k) = \sum_{i=1}^k y^2(i)$$

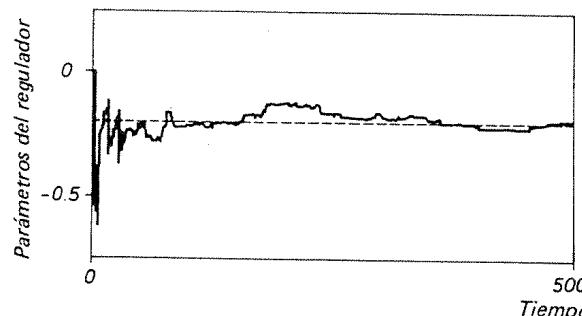


Figura 14.2 Estimaciones del parámetro $\hat{\theta}$ obtenidas en una simulación de la ley de control (14.12) y (14.13) aplicada al sistema (14.8) con $a = -0.9$, $b = 3$ y $c = -0.3$.

para el regulador auto-sintonizado y el regulador óptimo

$$u(k) = -0.2y(k)$$

El resultado se muestra en la Figura 14.3. Se observa claramente en la figura que no hay una gran diferencia en el funcionamiento de los dos reguladores.

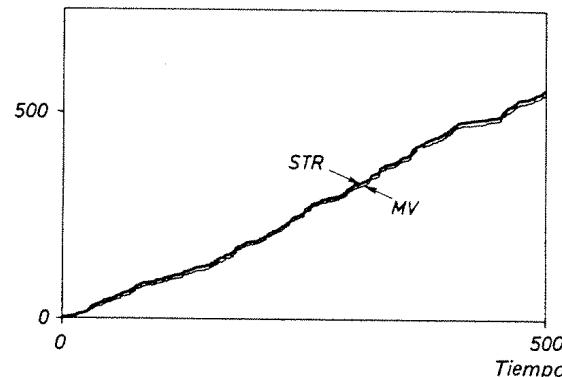


Figura 14.3 Funciones de coste acumulado para el regulador auto-sintonizado dado en (14.12) y (14.13) y el regulador óptimo basado en parámetros conocidos.

El ejemplo muestra que el regulador auto-sintonizado simple de (14.12) y (14.13) funcionará muy bien. Después de un período transitorio corto, tendrá casi el mismo funcionamiento que un regulador basado en los parámetros exactos del sistema. El parámetro θ parece converger al valor que corresponde al regulador de varianza

mínima. Estas observaciones empíricas se investigarán más adelante. Supongamos en principio que el sistema está gobernado por

$$y(k+1) + ay(k) = u(k) + n(k) \quad (14.14)$$

donde n es una perturbación. Si la perturbación está limitada en el sentido

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=1}^k n^2(i) < \infty \quad (14.15)$$

entonces el valor medio del cuadrado de la salida del sistema en lazo cerrado

$$\frac{1}{t} \sum_{k=1}^t y^2(k) \quad (14.16)$$

también está acotado.

Esta afirmación se demuestra por contradicción a lo largo de los siguientes párrafos. Si y no está acotada, entonces la influencia de $n(k)$ en (14.14) puede no tenerse en cuenta. Esto implica que $\hat{\theta}(k)$ convergerá hacia a . Sin embargo, utilizando la ley de control

$$u(k) = ay(k)$$

daría una salida acotada porque n está acotado; esto es una contradicción.

El regulador auto-sintonizado de (14.12) y (14.13) siempre estabilizará el sistema de (14.14) en el sentido de la media de los cuadrados. Puede demostrarse en este caso que la estimación de parámetros $\hat{\theta}(k)$ también converge. Si $\hat{\theta}(k)$ converge cuando $k \rightarrow \infty$, es fácil encontrar el punto de convergencia. La ecuación normal puede escribirse

$$\sum_{i=1}^k y(i+1)y(i) = \hat{\theta}(k+1) \sum_{i=1}^k y^2(i) + \sum_{i=1}^k y(i)u(i)$$

La ecuación (14.13) ahora da

$$\frac{1}{k} \sum_{i=1}^k y(i+1)y(i) = \frac{1}{k} \sum_{i=1}^k [\hat{\theta}(k+1) - \hat{\theta}(i)]y^2(i)$$

El miembro de la derecha converge a cero cuando $k \rightarrow \infty$ porque $\hat{\theta}(k)$ converge y $y(k)$ tiene limitada la media de los cuadrados. De este modo, si la estimación de parámetros converge, entonces

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=1}^k y(i+1)y(i) = 0 \quad (14.17)$$

El regulador auto-sintonizado intenta hacer cero la correlación de la salida del sistema en lazo cerrado, $r_y(\tau)$, para $\tau = 1$. Suponiendo que el proceso que se va a controlar está definido por (14.8), se deduce que (14.16) está acotado y (14.17) se verifica sólo para un valor de $\hat{\theta}$, precisamente para

$$\theta = a - c$$

De este modo se ha establecido que el regulador de (14.12) y (14.13) está auto-sintonizado para el sistema (14.8) y el criterio de varianza mínima. El análisis puede extenderse al caso en que $b \neq 1$. El resultado puede extenderse para controlar un proceso de orden n -ésimo de la forma de (14.7). Entonces se requieren condiciones adicionales para estabilidad y convergencia. También hay casos en que la estimación de parámetros no converge.

Resultados generales

El análisis del sencillo ejemplo anterior puede generalizarse. Las propiedades asintóticas se resumen en dos teoremas.

Teorema 14.1. Considérese el algoritmo 14.2 y supóngase que $B^- = 1$. Supóngase, además, que los parámetros estimados s_i , $i=0, \dots, n_s$ y r_i , $i=1, \dots, n_r$, convergen y que el sistema en lazo cerrado es ergódico (en los segundos momentos); entonces, el sistema en lazo cerrado tiene las propiedades

$$\begin{aligned} E(y(k+\tau)y(k)) &= r_y(\tau) = 0, \quad \tau = d, \dots, d+n_r, \\ E(y(k+\tau)u(k)) &= r_{yu}(\tau) = 0, \quad \tau = d, \dots, d+n_r, \end{aligned} \quad \square$$

Teorema 14.2. El sistema que se va a controlar está gobernado por (14.7), donde se conoce $d = \text{gra } A - \text{gra } B$. Supongamos que se usa el algoritmo 14.2 con $n_s \geq n-1$ y $n_r \geq n+d-1$. Si el parámetro estimado converge, entonces el regulador de (14.3) con $T=0$ converge al regulador de varianza mínima para el proceso (14.7). \square

El Teorema 14.1 implica que los puntos estacionarios del algoritmo están caracterizados por la condición de que algunos valores de $r_y(\tau)$ y $r_{yu}(\tau)$ sean iguales a cero. El Teorema se mantiene independientemente de lo que parezca el proceso. El Teorema 14.2 implica que si el proceso está gobernado por (14.7), si se conoce d , y si hay suficientes parámetros en el regulador, entonces el algoritmo convergerá al controlador de varianza mínima si converge de alguna forma.

También pueden analizarse las propiedades de estabilidad y convergencia de un regulador auto-sintonizado. Puede demostrarse que un auto-sintonizador basado en la estimación de mínimos cuadrados y control de varianza mínima converge a un regulador de varianza mínima para (14.7) con la condición de que

$$\frac{1}{C(z)} - \frac{1}{2}$$

sea real estrictamente positivo. El algoritmo será estable localmente con la condición de que

$$C(z_i) > 0$$

para todos los z_i tales que $B(z_i)=0$.

14.4. Otros métodos de control adaptativo

El auto-sintonizador expuesto en la Sección 14.2 está motivado por el deseo de obtener una sintonía automática para un lazo de control. De hecho, el nombre de *auto-sintonizado* se introdujo para destacar esto. Parece razonable, sin embargo, que el regulador de la Figura 14.1 pueda hacerse para controlar un proceso con parámetros variables; es decir, se llevaría a cabo un regulador adaptativo. Para hacer esto es necesario cambiar el algoritmo para que el estimador de parámetros pueda seguir la variación de los parámetros del proceso. Un método de realización consiste en olvidar los datos antiguos, como se estudió en la Sección 13.5. También hay otros esquemas para control adaptativo que están estrechamente relacionados con el auto-sintonizador. Tres de ellos se describirán en esta sección. El punto inicial es un lazo de control con realimentación corriente con un proceso y un regulador de parámetros ajustables. El esquema representa diferentes modos de alterar los parámetros del regulador en respuesta a cambios en las dinámicas del proceso y las perturbaciones. El esquema difiere únicamente en el modo de ajustar los parámetros del regulador.

Planificación de la ganancia

A veces es posible encontrar variables auxiliares del proceso que se correlacionan bien con los cambios en la dinámica del proceso. Entonces es posible eliminar la influencia de las variaciones de los parámetros cambiando los parámetros del regulador mediante funciones de las variables auxiliares (véase la Fig. 14.4). Este método se denomina *planificación de la ganancia* porque el sistema se utilizó originalmente para acomodar solamente cambios en la ganancia del proceso.

Una desventaja de la planificación de la ganancia está en que es una compensación en lazo abierto. No hay realimentación que compense una planificación incorrecta. De este modo, la planificación de la ganancia puede verse como una extensión de la compensación por prealimentación. Otro inconveniente radica en que

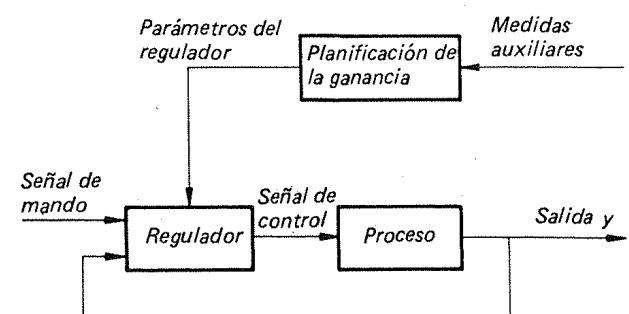


Figura 14.4 Diagrama de bloques de un sistema en el que se han eliminado las variaciones de los parámetros por planificación de la ganancia.

es un diseño que necesita tiempo. Los parámetros del regulador deben determinarse para muchas condiciones de operación. El funcionamiento debe comprobarse mediante una gran cantidad de simulaciones.

La planificación de la ganancia tiene la ventaja de que los parámetros pueden cambiar rápidamente en respuesta a los cambios en el proceso. Las limitaciones dependen de la rapidez con que las medidas auxiliares responden a los cambios en el proceso.

Hay una controversia en la nomenclatura concerniente a si la planificación de ganancia debe considerarse como un método adaptativo o no, porque los parámetros se cambian en lazo abierto. Independientemente de esta discusión, la planificación de ganancia es una técnica muy utilizada para reducir los efectos de la variación de parámetros.

Sistemas adaptativos de modelo de referencia (MRAS)

En la Figura 14.5 se ilustra otro modo de ajustar los parámetros del regulador. Este esquema se desarrolló originalmente para el problema del servo. Las especifica-

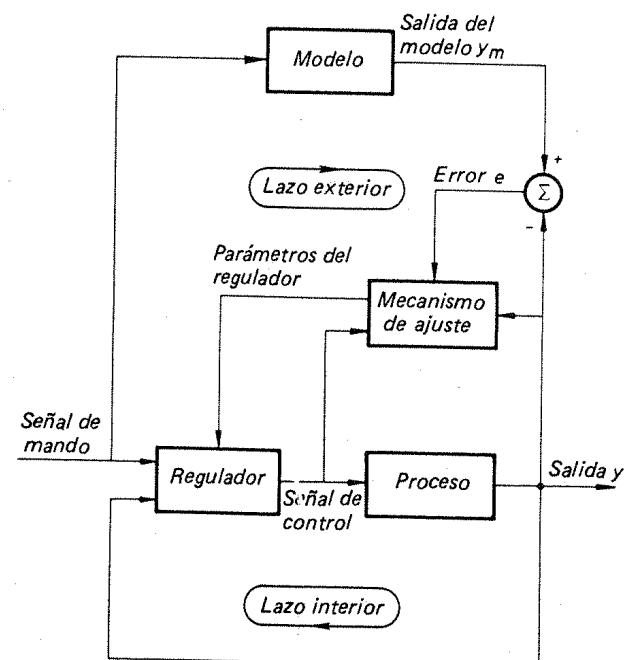


Figura 14.5 Diagrama de bloques de un sistema adaptativo de modelo de referencia.

ciones se dan en función de un modelo de referencia, el cual muestra idealmente cómo debe responder la salida del proceso a la señal de mando. Obsérvese que el modelo de referencia es parte del sistema de control. Puede considerarse que el regulador está formado por dos lazos. El lazo interno es un lazo de control corriente compuesto por el proceso y el regulador. Los parámetros del regulador se ajustan mediante el lazo exterior de tal modo que el error e entre la salida del modelo y_m y la salida del proceso y es pequeño. De este modo, el lazo exterior también es similar a un lazo del regulador. El problema fundamental consiste en determinar el mecanismo de ajuste para obtener un sistema estable, el cual lleve el error a cero. Este problema no es trivial. Es fácil demostrar que no puede resolverse con una simple realimentación desde el error a los parámetros del controlador.

El siguiente mecanismo de ajuste de parámetros, denominado *regla MIT*, se utilizó en el MRAS (del inglés Model-Reference Adaptive Systems) original:

$$\frac{d\theta}{dt} = -\alpha e \text{ grad}_\theta e \quad (14.18)$$

donde e es el error del modelo y las componentes del vector θ son los parámetros ajustables. El número α es un parámetro que determina la velocidad de adaptación. La ecuación (14.18) representa un mecanismo de ajuste, el cual está compuesto de tres partes: un filtro lineal para calcular las variaciones de la sensibilidad de las entradas y salidas del proceso, un multiplicador y un integrador. Esta configuración es típica en algunos sistemas adaptativos.

La regla MIT funcionará bien si el parámetro α es pequeño. La medida admisible depende de la magnitud de la señal de referencia. Por consiguiente, no es posible fijar unos límites que garanticen la estabilidad. Así, la regla MIT puede dar un sistema en lazo cerrado inestable. Pueden obtenerse reglas de ajuste modificadas utilizando la teoría de estabilidad. Estas reglas son similares a la regla MIT. La sensibilidad deducida en (14.18) se sustituirá por otras funciones.

Relaciones entre MRAS y STR

Está claro de las Figuras 14.1 y 14.5 que MRAS y STR están estrechamente relacionadas. Ambos sistemas tienen dos lazos de realimentación. El lazo interno es un lazo de realimentación normal con un proceso y un regulador. El regulador tiene parámetros ajustables que se definen mediante el lazo exterior. El ajuste se basa en la realimentación desde las entradas y salidas del proceso. Sin embargo, los métodos de diseño del lazo interno y las técnicas utilizadas para ajustar los parámetros del lazo exterior pueden ser diferentes. Cuando los métodos fundamentales de diseño son los mismos, el auto-sintonizador implícito y el MRAS son idénticos.

El regulador mostrado en la Figura 14.1 también puede deducirse del método MRAS si la estimación de parámetros se hace actualizando un modelo de referencia. El esquema entonces se denomina un *MRAS indirecto* porque los parámetros del regulador se actualizan indirectamente a través de los cálculos del diseño. El *MRAS directo*, donde los parámetros del regulador se actualizan directamente, está



estrechamente relacionado con el STR implícito; el MRAS indirecto está relacionado con el STR explícito.

Teoría de control estocástico

Las estructuras de los reguladores como el MRAS y el STR están basadas en argumentos heurísticos. Debería resultar atractivo obtener reguladores a partir de estructuras teóricas unificadas. Esto puede hacerse utilizando teoría de control estocástico no lineal. Entonces, el sistema y su entorno se describen por un modelo estocástico. El criterio se formula para minimizar el valor esperado de la función de coste, la cual es una función escalar de estados y controles.

El problema de encontrar un control que minimice la función de coste esperada es difícil. No se conocen las condiciones para la existencia del control óptimo. Bajo la suposición de que existe solución puede deducirse una ecuación funcional para la función de coste óptima utilizando programación dinámica. Esta ecuación, denominada *ecuación de Bellman*, puede resolverse numéricamente sólo en casos muy simples. La estructura del regulador óptimo obtenido se muestra en la Figura 14.6. Se puede considerar al controlador compuesto de dos partes: un estimador y un regulador de realimentación. El estimador genera la distribución de probabilidad condicional del estado a partir de las medidas. Esta distribución se denomina *Hiperestado* del problema. El regulador de realimentación es una función no lineal que aplica el hiperestado en el espacio de las variables de control.

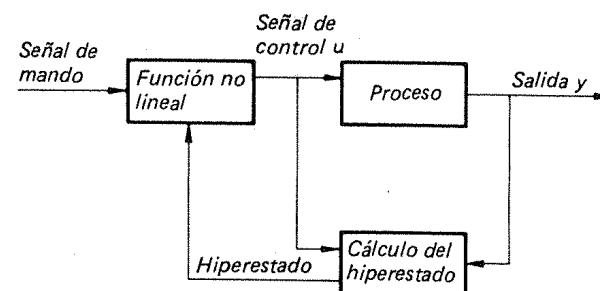


Figura 14.6 Diagrama de bloques de un regulador adaptativo obtenido mediante la teoría de control estocástico.

La simplicidad estructural de la solución se obtiene pagando el precio de introducir el hiperestado, el cual es una cantidad de gran dimensión. Obsérvese que el auto-sintonizador de la Figura 14.1 puede considerarse como una aproximación al regulador de la Figura 14.6, donde el hiperestado se approxima por el estado del proceso y el parámetro estimado.

La ley de control óptimo tiene una propiedad interesante. El control no sólo tratará de conducir la salida a su valor deseado; cuando hay incertidumbre en los

parámetros, el regulador también introduce perturbaciones, las cuales mejoran la estimación y los futuros controles. El control óptimo da el equilibrio correcto entre el mantener pequeñas las señales de control y los errores de estimación. Esta propiedad se denomina *control dual*.

Ejemplo 14.1

Se utiliza un ejemplo simple como ilustración. Considérese un sistema descrito por

$$y(k+1) = y(k) + bu(k) + e(k)$$

donde u es el control, y es la salida, e es un ruido blanco y b es un parámetro constante o un proceso de Markov. Supongamos que el criterio es minimizar la desviación cuadrática media de la salida y .

Si el parámetro b tiene una distribución gaussiana, puede demostrarse que la distribución condicional de b , dadas las entradas y salidas hasta el instante k , es gaussiana con distribución media $\hat{b}(k)$ y desviación estándar $\sigma(k)$. Las ecuaciones para actualizar \hat{b} y σ son las mismas que las ecuaciones del filtro corriente de Kalman.

La ley de control óptimo puede calcularse numéricamente utilizando programación dinámica. Una aproximación a la ley de control dual es

$$u(k) = -y(k) \frac{\hat{b}(k) + 0.56\sigma(k)}{\hat{b}^2(k) + 0.08\hat{b}(k)\sigma(k) + 2.2\sigma^2(k)} - \frac{1.9\sigma^3(k)}{\hat{b}^4(k) + 1.7\sigma^4(k)} \quad (14.19)$$

para $\hat{b} > 0$ e $y \geq 0$. La ley de control es simétrica en \hat{b} e y . Esta ley de control tiene acción de control dual debido al último término.

Pueden hacerse algunas aproximaciones a la ley de control dual. El *controlador de certidumbre*

$$u(k) = -y(k)/\hat{b}(k) \quad (14.20)$$

se obtiene simplemente resolviendo el problema del control en el caso de parámetros conocidos y sustituyendo los parámetros conocidos por sus estimaciones. El regulador auto-sintonizado puede interpretarse como un controlador de certidumbre.

La ley de control

$$u(k) = -\frac{1}{\hat{b}(k)} \cdot \frac{\hat{b}^2(k)}{\hat{b}^2(k) + \sigma^2(k)} y(k) \quad (14.21)$$

es otra aproximación, denominada *controlador preventivo*, porque compensa y usa menor ganancia cuando la estimación es incierta. Los controladores (14.19) y (14.21) coinciden con (14.20) cuando no hay incertidumbre en b .

Obsérvese que el controlador de certidumbre y el controlador preventivo no tienen la propiedad dual.

14.5. Usos y abusos de las técnicas adaptativas

Las técnicas auto-sintonizadas y adaptativas se encuentran aún en su infancia. Estos métodos no se pudieron utilizar a costos razonables antes de la aparición de los microprocesadores. La posibilidad de usar estas técnicas es interesante porque ofrecen ventajas muy importantes para obtener nuevos tipos de reguladores y funciones en los sistemas de control.

Auto-sintonizado

Es posible sintonizar manualmente reguladores con tres o cuatro parámetros si no hay interacción entre los ajustes de los diferentes parámetros. Sin embargo, para reguladores más complejos, es necesario disponer de herramientas para sintonizarlos adecuadamente. Tradicionalmente, el sintonizado de los reguladores más complejos ha seguido la línea de modelado o identificación y diseño del regulador. Este es con frecuencia, un procedimiento lento y costoso, el cual sólo puede aplicarse en lazos importantes o a sistemas que se fabrican en grandes cantidades.

Los controles MRAS o STR se transforman en controles con realimentación de ganancia constante cuando los parámetros estimados son constantes. Por tanto, el lazo adaptativo puede utilizarse como *sintonizador* para un lazo de control. En tales aplicaciones simplemente se conecta el lazo de adaptación; pueden añadirse señales de perturbación. El regulador adaptativo trabaja hasta que el funcionamiento es satisfactorio. En tales casos el lazo de identificación también funciona con la ganancia disminuida. Entonces, el lazo de adaptación se desconecta y se deja funcionar al sistema con los parámetros del regulador fijos. El auto-sintonizado se puede considerar como un modo adecuado de incorporar modelado y diseño automático dentro del regulador. Esto puede aumentar la clase de problemas para los que pueden utilizarse métodos sistemáticos de diseño a costos eficaces. El autosintonizado es particularmente útil para métodos de diseño, tales como prealimentación, la cual depende críticamente de la bondad del modelo. El sintonizado automático puede aplicarse a controladores PID simples, así como a modelos de reguladores más complicados.

Es conveniente introducir auto-sintonizado en los paquetes para DDC. Entonces, un algoritmo de sintonizado puede servir para muchos lazos. Como un buen algoritmo de sintonizado necesita sólo unos pocos cientos de localizaciones de memoria en un computador de control, se obtienen beneficios sustanciales a costos despreciables.

El auto-sintonizado también se puede incluir en reguladores de un lazo. Por ejemplo, es posible diseñar reguladores con tres modos conmutables: manual, automático y sintonizado. Sin embargo, el algoritmo de sintonizado representa la mayor parte de la programación de un regulador de un lazo. Cuando se introduce el auto-sintonizado las necesidades de memoria normalmente se duplican.

El auto-sintonizado resuelve algunas de las dificultades teóricas comúnmente asociadas al control adaptativo. La única dificultad teórica que se presenta es la identificación bajo condiciones de lazo cerrado. Esto puede evitarse introduciendo señales de perturbación. Como el sintonizado puede realizarse bajo supervisión de un operario, a menudo es posible hacerlo con seguridad.

Botones de ajuste relacionados con las prestaciones

La interfase con el operador es muy importante porque los reguladores adaptativos también pueden tener parámetros que deben fijarse. Cuando se aplican

técnicas adaptativas, es deseable disponer de una caja absolutamente negra sin ningún ajuste externo. Tales reguladores pueden diseñarse para aplicaciones específicas, donde pueden establecerse a priori los objetivos del control; sin embargo, en algunos casos no es posible especificarlos a priori. Al menos es necesario decir al regulador qué se espera que haga. Esto puede hacerse introduciendo *selectores relacionados con las prestaciones* que fijan las propiedades deseadas del sistema en lazo cerrado. Utilizando este concepto se pueden diseñar nuevos tipos de reguladores. Por ejemplo, es posible tener reguladores con un selector que especifica el ancho de banda deseado en lazo cerrado. Otra posibilidad consiste en tener un selector, que indique el peso relativo de la desviación del estado y la acción de control en un problema LQG. Una tercera posibilidad consiste en disponer de un selector para seleccionar los márgenes de fase y ganancia. Las características de un regulador con ajustes relativos a las prestaciones se ilustra con un ejemplo.

Ejemplo 14.2 – El auto-sintonizador de ancho de banda

El auto-sintonizador de ancho de banda es un regulador que tiene un dispositivo ajustable en el panel de mandos, marcado con el *ancho de banda deseado en lazo cerrado*. La realización particular es en forma de un auto-sintonizador por asignación de polos.

La Figura 14.7 muestra la respuesta del servo ante una onda cuadrada como señal de mando. Al principio el regulador no tiene información sobre el sistema. Se ve en la figura que el servo funciona muy bien después del primer paso de la señal de mando. La Figura 14.7 muestra que la señal de control da unos «saltos» moderados después de un escalón en la señal de mando cuando el ancho de banda exigido es bajo (1,5 rad/s). Entonces la señal de control decrece gradualmente a un nivel estacionario. Cuando se exige un ancho de banda grande (4,5 rad/s), como en la Figura 14.7b), el valor de la señal de control inmediatamente después del escalón es más de 30 veces superior al valor estacionario. Un ancho de banda bajo produce una respuesta lenta, señales de control bajas y baja sensibilidad al ruido de las medidas. Un ancho de banda grande produce respuestas rápidas. Sin embargo, las acciones de control serán grandes y el sistema será sensible al ruido. Es fácil para un operador determinar experimentalmente el ancho de banda adecuado a una aplicación particular. Obsérvese que excluyendo las especificaciones del ancho de banda, todos los ajustes necesarios —por ejemplo, la selección del período de muestreo— se realizan automáticamente por el auto-sintonizador.

Abusos del control adaptativo

Un regulador adaptativo, que es esencialmente no lineal, es más complicado que el regulador de ganancia fija. Antes de intentar utilizar control adaptativo, es muy importante asegurarse que el problema de control no puede resolverse por realimentación de ganancia constante. En la amplia literatura sobre control adaptativo, hay muchos casos en los cuales una realimentación de ganancia constante puede funcionar tan bien como un regulador adaptativo. Obsérvese que no es posible juzgar la necesidad de un control adaptativo por las variaciones de la dinámica en lazo abierto sobre el rango de operación. Se conocen muchos casos en los que una realimentación de ganancia constante puede trabajar bien en sistemas dinámicos con variaciones considerables.

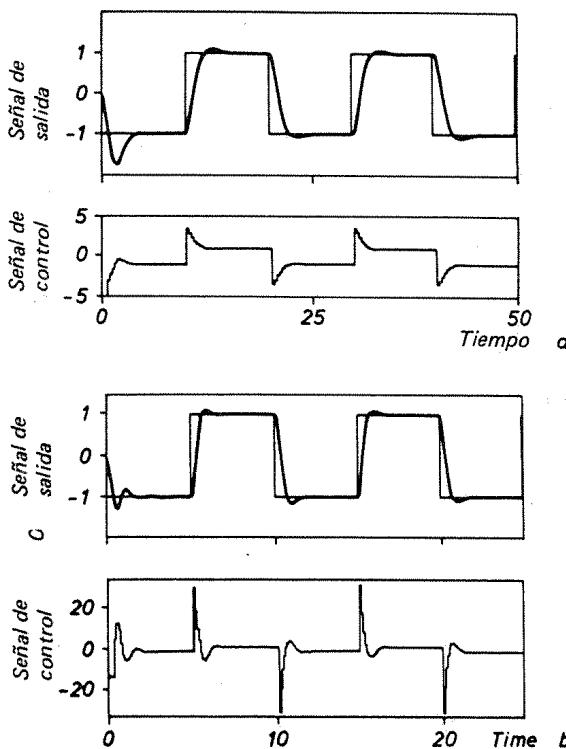


Figura 14.7 Simulación del auto-sintonizador de ancho de banda. El proceso tiene como función de transferencia $1/(s+1)^2$. El ancho de banda exigido es a) 1,5 rad/s y b) 4,5 rad/s.

teóricos sobre estabilidad y estructura. Aún queda mucho trabajo teórico por desarrollar. La aparición de los microprocesadores ha impulsado fuertemente las aplicaciones. Los experimentos en laboratorios y la posibilidad de estudios industriales han contribuido a una mejor comprensión de los aspectos prácticos del control adaptativo. También están apareciendo un gran número de reguladores adaptativos en el mercado.

14.7. Problemas

14.1. Considerar el proceso

$$y(k) + \alpha y(k-1) = bu(k-1) + e(k) + ce(k-1)$$

y utilizar el método de mínimos cuadrados para estimar el parámetro α en el modelo

$$y(k+1) + \alpha y(k) = \beta_0 u(k) + \epsilon(k+1)$$

donde β_0 se supone que es conocido. Genérese la señal de control mediante

$$u(k) = \alpha/\beta_0 y(k)$$

- a) Demostrar que un posible punto de convergencia para la estimación de mínimos cuadrados está caracterizado por

$$r_y(1) = 0$$

$$r_{yu}(1) = 0$$

- b) Demostrar que el controlador de varianza mínima dado por

$$\alpha = \frac{a - c}{b} \beta_0$$

es un posible punto de convergencia.

- 14.2. Simular el sistema del Problema 14.1 donde α se estima recursivamente. Estudiar la influencia de β_0 . ¿Pueden b y β_0 , por ejemplo, tener diferentes signos?

(Sugerencia: Es útil restringir la señal de control, por ejemplo, a 3-5 veces la desviación estándar de la señal de control, cuando se utiliza el controlador óptimo de varianza mínima.)

- 14.3. Deducir las condiciones para que

$$y(k) + a_1 y(k-1) + \cdots + a_n y(k-n) = n(k)$$

sea de valor cuadrático medio acotado cuando n es de valor cuadrático medio acotado; es decir,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{j=1}^k n(j)^2 < \infty$$

- 14.4. Considérese el proceso

$$y(k+1) - y(k) = u(k-1) + e(k+1) - 0.9e(k)$$

- a) Determinar la ley de control de varianza mínima para el sistema.

Lazos de supervisión

Los sistemas adaptativos de las Figuras 14.1 y 14.5 pueden considerarse como sistemas jerárquicos de dos niveles. El nivel inferior es el lazo de realimentación usual con el proceso y el regulador. El lazo de adaptación, el cual ajusta los parámetros, representa el nivel superior. El lazo de adaptación en MRAS o STR típicos requiere parámetros tales como los órdenes de los modelos, factores de olvido y períodos de muestreo. Puede añadirse un tercer nivel a esta jerarquía para inicializar estos parámetros. Por ejemplo, puede determinarse un factor de olvido monitorizando la excitación del proceso. Almacenando las entradas y salidas del proceso, también es posible estimar modelos con diferentes períodos de muestreo, diferentes órdenes y diferentes estructuras.

14.6. Conclusiones

Las técnicas adaptativas están emergiendo lentamente después de muchos años de investigación y experimentación. Se han establecido importantes resultados

b) Determinar un auto-sintonizador para el modelo

$$y(k+1) = s_0 y(k) + s_1 y(k-1) + r_0 [u(k) + r_1 u(k-1)] + \epsilon(k)$$

basado en el control de varianza mínima y estimación de mínimos cuadrados. Demostrar que el regulador es compatible con el regulador de varianza mínima para el sistema de a) y comparar las características del auto-sintonizador con el regulador de varianza mínima basado en parámetros conocidos.

c) Determinar los posibles puntos de equilibrio para el auto-sintonizador de b) utilizando los Teoremas 14.1 y 14.2.

14.5. Simular un auto-sintonizador basado en la estimación de los parámetros del modelo

$$y(k) + a_1 y(k-1) + a_2 y(k-2) = b_1 u(k-1) + b_2 u(k-2)$$

y el diseño por asignación de polos dado en el Capítulo 10. Verificar los resultados de la Figura 14.7.

14.6. Probar que (14.14) tendrá una salida limitada si la perturbación está limitada cuando se utiliza el auto-sintonizador simple de (14.12) y (14.13).

14.8. Referencias

Estudios sobre teoría y aplicaciones de reguladores auto-sintonizados pueden encontrarse en

ÅSTRÖM, K. J. (1980): «Self-Tuning Regulators: Design Principles and Applications,» en *Applications of Adaptive Control*, eds. Narendra and Monopoli. New York: Academic Press.

ISERMANN, R. (1982): «Parameter Adaptive Control Algorithms-A Tutorial,» *Automatica*, 18, 513-528.

ÅSTRÖM, K. J. (1983): «Theory and Applications of Adaptive Control,» *Automatica*, 19.

El regulador auto-sintonizado se propuso originalmente en

KALMAN, R. E. (1958): «Design of a Self-optimizing Control System,» *Trans. ASME*, 80, 468-78.

Desde entonces los reguladores auto-sintonizados han recibido mucha atención. Algunas referencias básicas son

PETERKA, V. (1970): *Adaptive Digital Regulation of Noisy Systems*. Preprint 2nd IFAC Symposium on Identification and Process Parameter Estimation, June 1970, Prague, Czechoslovakia.

ÅSTRÖM, K. J. y WITTENMARK (1973): «On Self-tuning Regulators,» *Automatica*, 9, 185-99.

CLARKE, D. W. y P. J. GAWTHROP (1975): «A Self-tuning Controller,» *Proc. IEE*, 122, 929-34.

WELLSTEAD, P. E., J. M. EDMUND, D. PRAGER, y P. ZANKER (1979): «Self-tuning Pole-Zero Assignment Regulators,» *Int. Journal Control*, 30, 1-26.

ÅSTRÖM, K. J. y WITTENMARK (1980): «Self-tuning Controllers Based on Pole-Zero Placement,» *Proc. IEE*, 127, 120-30.

El control adaptativo se trata en general en

BELLMAN, R. (1961): *Adaptive Control: A Guide Tour*. Princeton, N. J.: Princeton University Press.

SARIDIS, G. N. (1977): *Self-organizing Control of Stochastic Systems*. New York: Marcel Dekker.
WITTENMARK, B. (1975): «Stochastic Adaptive Control Method: A Survey,» *Int. Journal Control*, 21, 705-30.

El control adaptativo de modelo de referencia se discute en

WHITAKER, H. P., J. YAMROM, y A. KEZER (1958): *Design of Model-reference Adaptive Control Systems for Aircraft*. Report R-164, Instrumentation Laboratoru, MIT, Cambridge, Mass.

PARKS, P. C. (1966): «Lyapunov Redesign of Model Reference Adaptive Control Systems,» *IEEE Trans. Autom. Control*, AC-11, 326-67.

MONOPOLI, R. V. (1974): «Model Reference Adaptive Control with an Augmented Error Signal,» *IEEE Trans. Autom. Control*, AC-19, 474-84.

LANDAU, Y. D. (1979): *Adaptive Control: The Model Reference Approach*. New York: Marcel Dekker.

El análisis de las propiedades de estabilidad y convergencia de reguladores auto-sintonizados se encuentra en

LJUNG, L. (1977): «On Positive Real Transfer Functions and the Convergence of Some Recursive Schemes,» *IEEE Trans. Autom. Control*, AC-22, 539-51.

EGARDT, B. (1979): *Stability of Adaptive Controllers*. Berlin: Springer-Verlag.

MORSE, A. S. (1980): «Global Stability of Parameter Adaptive Systems,» *IEEE Trans. Autom. Control*, AC-25, 433-39.

NARENDRA, K. S., Y. H. LIN, y L. S. VALAVANI (1980): «Stable Adaptive Controller Design. Part II: Proof of Stability,» *IEEE Trans. Autom. Control*, AC-25, 440-48.

GOODWIN, G. C., P. J. RAMADGE, y P. E. CAINES (1980): «Discrete Time Multivariable Adaptive Control,» *IEEE Trans. Autom. Control*, AC-25, 449-56.

GOODWIN, G. C., P. J. RAMADGE, y P. E. CAINES (1981): «Discrete Time Stochastic Adaptive Control,» *SIAM Journal on Control and Optimization* 19, 829-53.

El control dual lo introdujo Feldbaum en

FELDBAUM, A. A. (1960): «Dual Control Theory I-IV,» *Automation & Remote Control*, 21, 874-80, 1033-39, 22, 1-12, 109-21.

Más exposiciones sobre el problema del control dual se encuentran en

ÅSTRÖM, K. J. y B. WITTENMARK (1971): «Problems of Identification and Control,» *J. Math. Analysis and Appl.*, 34, 90-113.

BAR-SHALOM, Y. y E. TSE (1974): «Dual Effect, Certainty Equivalence and Separation in Stochastic Control,» *IEEE Trans. on Autom. Control*, AC-19, 494-500.

ÅSTRÖM, K. J. y A. HELMERSSON (1982): «Dual Control of a Low Order System,» *Proc. CNRS Colloque National*, Belle Île, France, September.

Las aplicaciones de control adaptativo se discuten en

NARENDRA, K. S. y R. V. MONOPOLI, eds. (1980): *Applications of Adaptive Control*. New York: Academic Press.

UNBEHAUEN, H. ed. (1980): *Methods and Applications in Adaptive Control*. Berlin: Springer-Verlag.

Realización práctica de controladores digitales

Objetivo:

Exponer la realización de controladores digitales.

15.1. Introducción

En los capítulos anteriores se ha expuesto el diseño de algoritmos para control con computador. En este capítulo se analizará el problema de realización de un algoritmo de control mediante un computador digital.

Los algoritmos de control obtenidos en los capítulos anteriores son sistemas dinámicos discretos. El problema clave es cómo realizar un sistema dinámico discreto mediante un computador digital. En la Sección 15.2 se da una visión global de este problema, que muestra que es inmediato obtener un programa de computador para un algoritmo discreto. Sin embargo, hay que considerar varios aspectos. Es necesario tener en cuenta las interfaces con los sensores, los actuadores y los operadores humanos. También hay que considerar la precisión numérica requerida.

La interfase de los sensores se trata en la Sección 15.3. Abarca el prefiltrado y los retardos de cálculo y muestra que estos últimos dependen de forma crítica de la organización del algoritmo. Se exponen diferentes formas de acortar el retardo de cálculo mediante una reorganización del programa. Se presentan métodos muy buenos para filtrar las señales de forma eficaz mediante la introducción de no linealidades, que pueden reducir la influencia de sensores poco fiables. Esta es una de las mayores ventajas del control con computador. La mayor parte de la teoría de este libro se refiere a la teoría lineal. Sin embargo, hay unas pocas no linealidades que deben tenerse en cuenta. La saturación del actuador es una de ellas. En la Sección 15.4 se exponen diferentes formas de tratarla. Esto conduce a la extensión de los métodos para evitar la saturación del término integral usados en el control clásico de procesos.

La interfase con el operador es otro factor importante; se trata en la Sección 15.5. Incluye el tratamiento de los modos de operación y diferentes formas de impedir los transitorios por conmutación. También se aborda la información que se debe presentar y las diferentes maneras de influir sobre el lazo de control. Los computadores digitales ofrecen muchas posibilidades interesantes; hasta el momento sólo se han empleado en un grado muy modesto. Hay muchas oportunidades para innovaciones en este campo.

Ciertamente es muy importante disponer de operaciones numéricas seguras en el algoritmo de control, que es el tema de la Sección 15.6. También se exponen los efectos de la longitud de palabras finita. La programación de los algoritmos de control se trata en la Sección 15.7. Es muy fácil realizar sistemas sencillos. Para sistemas más ambiciosos, en los que se cambian sobre la marcha los parámetros y los algoritmos de control, hace falta comprender los pormenores de la programación concurrente. En la Sección 15.7 se da una breve introducción, pero un tratamiento más detallado queda fuera del alcance de este libro.

15.2. Visión de conjunto

Esta sección da una visión global de la realización de leyes de control digitales. En primer lugar se dan diferentes representaciones de las leyes de control obtenidas en los Capítulos 8-12 a partir de los métodos de diseño; después se realizan los algoritmos. Se da también una lista de algunos problemas importantes. Estos problemas se expondrán con mayor detalle en las secciones siguientes.

Diferentes representaciones del regulador

Los métodos de diseño de los capítulos anteriores dan leyes de control en la forma de un sistema dinámico discreto. Dependiendo del enfoque usado, se obtienen representaciones diferentes.

Los métodos de diseño basados en la asignación de polos mediante realimentación de estados (Capítulo 9) y en el control LQG (Capítulo 11) dan un regulador de la forma

$$\begin{aligned}\hat{x}(k|k) &= \hat{x}(k|k-1) + K[y(k) - \hat{y}(k|k-1)] \\ u(k) &= L[x_m(k) - \hat{x}(k|k)] + D_c u_c(k) \\ \hat{x}(k+1|k) &= \Phi \hat{x}(k|k) + \Gamma u(k) \\ x_m(k+1) &= f(x_m(k), u_c(k)) \\ \hat{y}(k+1|k) &= C\hat{x}(k+1|k)\end{aligned}\quad (15.1)$$

En esta representación el estado del regulador es \hat{x} y x_m , donde \hat{x} es una estimación del proceso y x_m es el estado del modelo que genera la respuesta deseada a las señales de mando u_c . La expresión (15.1) recibe el nombre de *representación de estado con un observador explícito* por la interpretación física del estado del regulador. Es fácil incluir en esta representación un modelo no lineal para el estado que se desee.

Si la función f en (15.1) es lineal, el regulador en (15.1) es un sistema lineal con entradas y y u_c y salida u . Un regulador de este tipo siempre se puede representar como:

$$\begin{aligned} u(k) &= Cx(k) + Dy(k) + D_cu_c(k) \\ x(k+1) &= Fx(k) + Gy(k) + G_cu_c(k) \end{aligned} \quad (15.2)$$

(véase el Problema 9.7). La ecuación (15.2) es una *representación de estado general* de un sistema dinámico discreto. Esta forma es más compacta que (15.1). El estado no tiene necesariamente una interpretación física simple.

Los métodos de diseño para sistemas de una salida y una entrada expuestos en los Capítulos 10 y 12, que se basan en modelos externos, dan un regulador en forma de una *representación entrada-salida* general

$$R(q)u(k) = T(q)u_c(k) - S(q)y(k) \quad (15.3)$$

donde $R(q)$, $S(q)$ y $T(q)$ son polinomios del operador de desplazamiento hacia delante q . Existen transformaciones sencillas entre las diferentes representaciones (compárese con el Capítulo 3).

Realización de un sistema discreto

La realización de un sistema discreto definido por (15.1), (15.2) o (15.3) mediante un computador digital es inmediata. Los detalles dependen del software y del hardware disponibles. Para mostrar los principios en que se basa, supóngase que hay que realizar el sistema descrito por (15.2) empleando un computador digital con convertidores A-D y D-A y un reloj de tiempo real. En la Figura 15.1 puede verse una representación gráfica del programa. El reloj controla la ejecución del programa. La línea horizontal indica que se detiene la ejecución hasta que llega una interrupción desde el reloj. El reloj se dispone de forma que se obtiene una interrupción cada instante de muestreo. El programa dentro del rectángulo se ejecuta después de cada interrupción. En el Listado 15.1 se da el contenido del programa necesario para realizar el algoritmo.

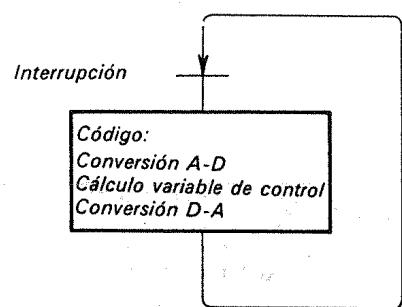


Figura 15.1 Representación gráfica de un programa empleado para realizar en la práctica un sistema discreto.

LISTADO 15.1 Esbozo de un programa empleando la ley de control dada en (15.2). Los números de línea se dan solamente como referencia.

```

Procedure Regulate
begin
1   Adin y uc
2   u := C*x + D*y + Dc*u
3   x := F*x + G*y + Gc*u
4   Daout u
end

```

En la primera línea se ordena la conversión analógica a digital. Los valores obtenidos se almacenan en los vectores y y u_c . La señal de control u se calcular en la segunda línea empleando la multiplicación de vectores y matrices y la suma de vectores. En la tercera línea se actualiza el vector de estado, y en la cuarta se realiza la conversión digital a analógica. Para obtener el programa completo hay que incluir también las declaraciones de tipo para los vectores u , uc , x e y , y para las matrices F , G , Gc , C , D y Dc . También hace falta asignar valores a las matrices y el valor inicial al estado x . Cuando se emplean lenguajes de programación que no tienen operaciones con matrices es necesario escribir los procedimientos apropiados para generar las operaciones con matrices usando operaciones con escalares. Nótese que las líneas segunda y tercera corresponden exactamente al algoritmo dado en (15.2).

Para obtener un buen sistema de control hay que considerar también los siguientes temas:

- Aspectos numéricos.
- Sensores.
- Actuadores.
- Aspectos operacionales.
- Aspectos de programación.

De ellos se tratará en las siguientes secciones.

15.3. Prefiltrado y retardo de cálculo

Cuando se realiza un sistema de control son muy importantes las interacciones entre el computador y su entorno. En esta sección se trata de la interfase del sensor.

En el Capítulo 6 se trataron las consecuencias de las perturbaciones. En el Capítulo 2 se expuso la importancia de utilizar un prefiltro analógico para evitar el enmascaramiento. También se ve claramente en la Sección 15.2 que siempre hay un retardo de tiempo asociado con los cálculos.



El prefiltro y el retardo de cálculo dan lugar a una dinámica adicional, que puede ser importante cuando se realiza un controlador digital. Ahora se van a tratar estos efectos.

Prefiltrado analógico

Para evitar el enmascaramiento (véase la Sección 2.5) es necesario usar un prefiltro analógico para eliminar las perturbaciones, con frecuencias mayores que la frecuencia de Nyquist asociada con el período de muestreo. En la Sección 2.5 se han examinado diferentes prefiltros desde el punto de vista del procesamiento de señal. El análisis se basó en el contenido de frecuencias de la señal. En un problema de control, normalmente hay disponible mucha más información acerca de las señales en forma de ecuaciones diferenciales para los modelos del proceso y posiblemente también para las perturbaciones. Un filtro de Kalman analógico sería un buen prefiltro porque se puede basar en una descripción detallada de la señal. Hay varias ventajas en realizar un filtro de Kalman en un computador. En este caso es útil muestrear las señales analógicas a una frecuencia alta en comparación con la del resto del sistema y eliminar el enmascaramiento mediante un filtro analógico ordinario diseñado desde el punto de vista del procesamiento de señal.

El ancho de banda ω_B del prefiltro es inversamente proporcional al período de muestreo h . Una regla empírica es

$$\omega_B h \approx 0.5 - 1$$

La elección exacta depende del orden del filtro y del carácter de la señal medida. Cuando se diseña el sistema hay que tener en cuenta la dinámica del sistema.

Si se cambia la frecuencia de muestreo hay que cambiar también el prefiltro. Con valores razonables de componentes es posible construir prefiltros analógicos para períodos de muestreo menores de algunos segundos. Para frecuencias de muestreo menores, a menudo es más sencillo muestrear una vez por segundo o más rápido mediante un prefiltro analógico adecuado y aplicar un filtrado digital a la señal muestreada. Además este enfoque posibilita el cambio del período de muestreo de los cálculos de control sólo por software.

Como el prefiltro analógico tiene su dinámica, es preciso incluir la dinámica del filtro en el modelo del proceso. Si se cambia el prefiltro o la frecuencia de muestreo, hay que recalcular los coeficientes de la ley de control. El siguiente ejemplo ilustra una forma sencilla de evaluar cuándo hay que considerar la dinámica de prefiltro en el diseño del control.

Ejemplo 15.1 – Cuándo se puede despreciar la dinámica de prefiltro

Consideremos un servo con ancho de banda ω_B (rad/s), que se ha realizado mediante un computador digital con una frecuencia de muestreo ω_s (rad/s). Tomemos como prefiltro un filtro Butterworth de segundo orden. La función de transferencia es (compárese con la Tabla 2.1).

$$G(s) = \frac{\omega_f^2}{s^2 + 2\zeta\omega_f s + \omega_f^2}$$

con $\zeta = 1/\sqrt{2}$. A la frecuencia ω_B , el filtro da un retardo de fase de

$$\alpha = \arctan \frac{2\zeta\omega_f\omega_B}{\omega_f^2 - \omega_B^2} \approx \frac{2\zeta\omega_B}{\omega_f} \text{ cuando } \omega_f \gg \omega_B$$

La atenuación a la frecuencia de Nyquist es

$$n = \frac{1}{|G(i\omega_N)|} \approx \left(\frac{\omega_N}{\omega_f}\right)^2 = \left(\frac{\omega_s}{2\omega_f}\right)^2$$

Eliminando ω_f entre las dos ecuaciones tenemos

$$\omega_s = \omega_B \frac{4\zeta\sqrt{n}}{\alpha} \quad (15.4)$$

Con $\alpha = 0.1$ rad (5.7°) y $n = 10$, la fórmula (15.4) da una frecuencia de muestreo ω_s que es 90 veces mayor que el ancho de banda.

El ejemplo indica que con frecuencias de muestreo normales —es decir, de 10 a 20 veces por período— es necesario considerar la dinámica del prefiltro (compárese con la Sección 10.9).

Retardo de cálculo

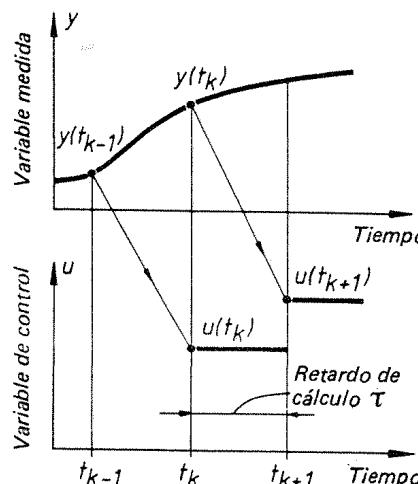
Como las conversiones A-D y D-A y los cálculos llevan un tiempo, siempre habrá un retardo cuando se realiza una ley de control utilizando un computador. El retardo, que recibe el nombre de *retardo de cálculo*, depende de cómo se realice el algoritmo de control. Hay básicamente dos formas de hacerlo (véase la Fig. 15.2). En el caso A, las variables medidas leídas en el instante t_k se pueden utilizar para calcular la señal de control que se aplicará en el instante t_{k+1} . Otra posibilidad es el caso B, que consiste en leer las variables medidas en el instante t_k y hacer la conversión D-A lo más pronto posible.

La desventaja del caso A es que las acciones de control se retrasan innecesariamente; la desventaja del caso B es que el retardo será variable dependiendo de la programación. En ambos casos es necesario tener en cuenta el retardo de cálculo cuando se calcule la ley de control. Esto se lleva a cabo fácilmente incluyendo un retardo h (caso A) o τ (caso B) en el modelo del proceso. Una práctica aconsejable es leer las entradas antes de enviar las salidas. Si no se hace así, siempre existe el riesgo de un acoplamiento eléctrico cruzado.

En el caso B es aconsejable hacer el retardo de cálculo lo más pequeño posible. Esto sólo se puede conseguir haciendo el menor número posible de operaciones entre las conversiones A-D y D-A.

Considérese el programa del Listado 15.1. Como la señal de control u está disponible después de ejecutar la segunda línea de código, la conversión D-A puede hacerse antes de actualizar el estado. El retardo se puede reducir todavía más calculando el producto $C*x$ después de la conversión D-A. El algoritmo del Listado 15.1 se modifica para obtener el Listado 15.2.

Caso A



Caso B

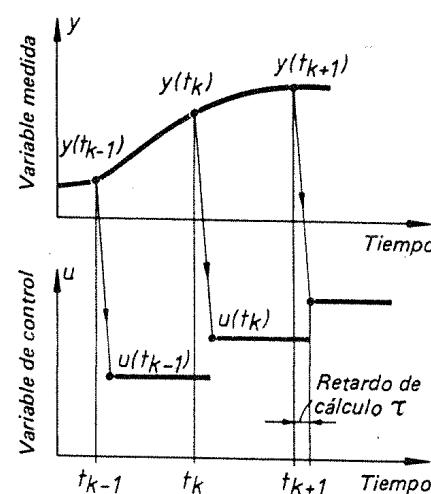


Figura 15.2 Dos formas de sincronizar las entradas y las salidas. En el caso A las señales medidas en el instante t_k se usan para calcular la señal de control a aplicar en el instante t_{k+1} . En el caso B las señales de control se cambian en cuanto se han calculado.

LISTADO 15.2 Esbozo de un programa que realiza prácticamente el algoritmo de control (15.2). Este programa tiene menor retraso de cálculo que el del Listado 15.1.

```

Procedure Regulate
begin
  1   Adin y uc
  2   u := u1 + D*y + Dc*u
  3   Daout u
  4   x := F*x + G*y + Gc*u
  5   u1 := C*x
end

```

Es útil disponer de buenas estimaciones de los tiempos de cálculo para diferentes algoritmos de control. Una buena manera de obtenerlos es ejecutar programas de prueba. Para las leyes de control lineal, a menudo es posible estimar los tiempos a partir de los resultados del cálculo de un producto escalar. Como las operaciones clave en las leyes de control lineal son operaciones matriciales, se pueden lograr buenas estimaciones de tiempos de cálculo midiendo el tiempo de cálculo del siguiente producto escalar:

```

Procedure Scapro;
begin
  s := 0;
  for i := 1 to n do
    s := s + a[i]*b[i];
end

```

El algoritmo requiere el cálculo de las direcciones de los elementos de la matriz así como de las sumas y multiplicaciones.

En los microcomputadores sencillos, que no tienen aritmética en punto flotante por hardware, habrá una diferencia de tiempo de cálculo sustancial entre las operaciones en punto fijo y punto flotante. Si se dispone de hardware para operaciones en punto flotante la diferencia será mucho menor.

Para juzgar las consecuencias de los retardos de cálculo, también es útil conocer la sensibilidad del sistema en lazo cerrado respecto a un retardo temporal. Una forma todavía más sencilla es evaluar cuánto cambian los polos en lazo cerrado cuando se introduce un retardo de un período de muestreo.

Detección de datos erróneos y de funcionamiento defectuoso de la medida

La teoría de filtrado lineal presentada en la Sección 11.3 es muy útil para reducir la influencia del ruido de medida. Sin embargo, hay otros muchos tipos de errores, como los defectos de funcionamiento de los instrumentos y los errores de conversión. Estos se caracterizan típicamente por grandes desviaciones, que se presentan con probabilidades reducidas. Es muy importante intentar eliminar este tipo de errores de forma que no entren en los cálculos de la ley de control. Existen muchas formas muy buenas de lograrlo al emplear el control con computador.

Los errores se pueden localizar en su origen. En los sistemas con requisitos de fiabilidad muy altos se puede lograr mediante la duplicación de los sensores. Entonces se combinan los dos sensores mediante una lógica muy sencilla que da una alarma si la diferencia entre las señales de los sensores es mayor que un umbral. Un par de sensores redundantes puede verse como un sensor que da una medida muy fiable o una señal de que no funciona correctamente.

En los casos más extremos se pueden emplear tres sensores. Se acepta una medida cuando dos de los tres sensores están de acuerdo (decisión por votación o lógica dos de tres). Por supuesto, también es posible emplear combinaciones más elaboradas de sensores y filtros.

También se puede emplear un filtro de Kalman para detección de errores. Por ejemplo, considérese el algoritmo de control de (15.1) con un observador explícito. Nótese que el error de predicción de un paso

$$\epsilon(k) = y(k) - \hat{y}(k|k-1) = y(k) - C\hat{x}(k|k-1) \quad (15.5)$$

aparece de forma explícita en el algoritmo. Si se dispone de la estimación de la matriz de covarianza del error de predicción, es fácil comprobar si una medida particular es razonable.



Una forma de obtener la matriz de covarianza es actualizar en línea la ecuación de covarianza del filtro de Kalman.

Para un filtro de la forma de (15.1), las ecuaciones son

$$\begin{aligned} R(k) &= CP(k|k-1)C^T + R_2 \\ K(k) &= P(k|k-1)C^T R^{-1}(k) \\ P(k|k) &= P(k|k-1) - P(k|k-1)C^T R^{-1}(k)CP(k|k-1) \\ P(k+1|k) &= A^T P(k|k)A + R_1 \end{aligned} \quad (15.6)$$

donde $R(k)$ es la covarianza de $\epsilon(k)$, $P(k|k-1)$ es la covarianza de $\hat{x}(k|k-1)$ y $P(k|k)$ es la covarianza de $\hat{x}(k|k)$ (véase la Sección 11.3). Si se detecta un error de medida, hay que modificar las ecuaciones de actualización. Si se rechazan todas las medidas, no hay que restar el segundo término en la ecuación para actualizar $P(k|k)$ porque representa la reducción en la varianza debido a la medida $y(k)$. También es posible rechazar de forma selectiva medidas de las diferentes componentes de $y(k)$. También se pueden combinar filtros de Kalman y parejas redundantes de sensores. Si se comprueban de esta forma los errores de medida es posible obtener un sistema muy flexible. El esquema se debe aumentar con comprobaciones para asegurar la observabilidad. En este caso se puede obtener un sistema que puede proporcionar los diferentes tipos de diagnósticos.

En el control con computador hay también otras muchas posibilidades de detectar diferentes tipos de errores de hardware y software. Con algunos canales de más en el convertidor A-D, conectados a tensiones fijas, se pueden realizar comprobaciones y calibraciones del sistema. Para comprobar y calibrar el convertidor D-A se conecta un canal del convertidor D-A a un canal A-D. Para comprobar el computador mismo se le puede pedir que haga unos cálculos cuyo resultado se conoce y se compara lo que se obtiene con los valores conocidos.

15.4. Actuadores no lineales

Todos los métodos de diseño de los Capítulos 9 al 12 se basan en la hipótesis de que se puede describir el proceso mediante un modelo lineal. Aunque la teoría lineal tiene una aplicabilidad muy amplia, a menudo hay que tener en cuenta alguna no linealidad. Por ejemplo, sucede con frecuencia que los actuadores son no lineales, como se ve en la Figura 15.3. Cuando la válvula no se satura, el sistema de la Figura

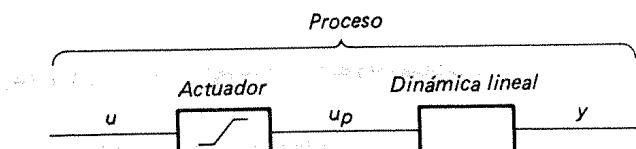


Figura 15.3 Diagrama de bloques de un proceso con un actuador no lineal que tiene saturación.

15.3 se puede describir de forma lineal. Entonces cuando se hacen cambios muy grandes es cuando la no linealidad tiene importancia. Si no se tiene en cuenta la no linealidad, durante el arranque y la parada pueden presentarse dificultades para el sistema de control, lo mismo que durante las grandes variaciones. Un ejemplo típico es el de la saturación del término integral, que se mostró en el Ejemplo 8.4.

La forma más racional de tratar la saturación es desarrollar una teoría de diseño que tenga en cuenta las no linealidades. Esto se puede llevar a cabo mediante la teoría de control óptimo. Sin embargo, un método de diseño de este tipo es bastante complicado. Además la correspondiente ley de control es compleja. Por tanto, lo más práctico es utilizar métodos heurísticos sencillos.

Las dificultades se presentan porque el regulador es un sistema dinámico. Cuando la variable de control se satura es necesario estar seguro de que el estado del regulador se comporta de forma adecuada. A continuación se exponen diferentes formas de lograrlo.

Reguladores en el espacio de estados con un observador explícito

Consideremos en primer lugar el caso en el que se describe una ley de control mediante un observador combinado con una realimentación del estado (15.1). El regulador es un sistema dinámico cuyo estado se representa en (15.1) mediante el estado estimado \hat{x} . En este caso es inmediato ver cómo se pueden evitar las dificultades provenientes de la saturación.

El estimador de (15.1) dará la estimación correcta si se escoge la variable u como la variable de control real u_p de la Figura 15.3. Si se mide la variable u_p entonces la estimación dada por (15.1) y el estado del regulador son correctos incluso aunque la variable de control se sature. Si no se mide la salida del actuador se puede estimar, suponiendo que se conocen las características no lineales. Para el caso de una saturación sencilla, la ley de control se puede escribir en la forma

$$\begin{aligned} \hat{x}(k|k) &= \hat{x}(k|k-1) + K[y(k) - C\hat{x}(k|k-1)] \\ &= [A - KC]\hat{x}(k|k-1) + B\hat{u}_p(k-1) \\ \hat{u}_p(k) &= \text{sat}\{L[x_m(k) - \hat{x}(k|k)] + u_m\} \\ \hat{x}(k+1|k) &= A\hat{x}(k|k) + B\hat{u}_p(k) \end{aligned} \quad (15.7)$$

donde la función sat se define por

$$\text{sat } u = \begin{cases} u_{\inf} & u \leq u_{\inf} \\ u & u_{\inf} < u < u_{\sup} \\ u_{\sup} & u \geq u_{\sup} \end{cases} \quad (15.8)$$



para un escalar y por

$$\text{sat } u = \begin{bmatrix} \text{sat } u_1 \\ \text{sat } u_2 \\ \vdots \\ \text{sat } u_n \end{bmatrix} \quad (15.9)$$

para un vector. Los valores u_{\inf} y u_{\sup} se escogen según las limitaciones del actuador. En la Figura 15.4 se presenta un diagrama de bloques de un regulador con un modelo de la no linealidad del actuador. Obsérvese que aunque la función de transferencia entre y y u para (15.1) sea inestable, el estado del sistema en (15.7) estará acotado si la matriz $A - KC$ es estable. También es evidente que \hat{x} será una buena estimación del estado del proceso incluso aunque la válvula se sature, suponiendo que se han escogido adecuadamente u_{\inf} y u_{\sup} .

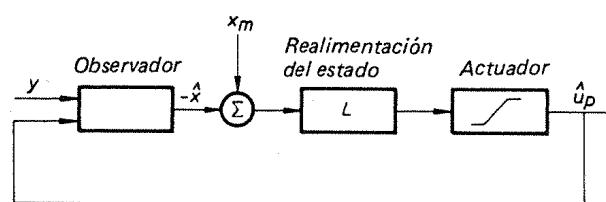


Figura 15.4 Regulador basado en un observador y una realimentación del estado con compensación de la saturación del término integral.

El modelo general en el espacio de estados

También se puede especificar el regulador como un modelo en el espacio de estados de la forma dada en (15.2):

$$x(k+1) = Fx(k) + Gy(k) \quad (15.10)$$

$$u(k) = Cx(k) + Dy(k) \quad (15.11)$$

que no incluye un observador explícito. Por simplificar se han despreciado las señales de mando. Si la matriz F tiene valores propios fuera del círculo unidad y la variable de control se satura, es evidente que puede presentarse una saturación del término integral. Supongamos, por ejemplo, que la salida ha llegado a su límite y hay un error de control y . El estado y la señal de control continuarán creciendo aunque la influencia sobre el proceso esté limitada debido a la saturación.

Para evitar esta dificultad interesa estar seguro de que el estado de (15.10) toma los valores apropiados cuando la variable de control se satura. En los controladores

de proceso convencionales esto se lleva a cabo introduciendo un *modo de seguimiento* especial, que garantiza que el estado del sistema corresponde a la secuencia de entrada-salida $\{u_p(k), y(k)\}$. El diseño de un modo de seguimiento puede formularse como un problema de observador. En el caso de realimentación del estado con un observador explícito, el seguimiento se hace automáticamente proporcionando al observador la salida del actuador u_p o su estimación \hat{u}_p . En el regulador de (15.10) y (15.11) no hay un observador explícito. Para conseguir un regulador que elimine el problema de la saturación del término integral imitaremos la solución del regulador con un observador explícito. Primero se reescribe la ley de control como se indica en la Figura 15.5. Los sistemas en a) y b) tienen la misma relación entrada-salida. El

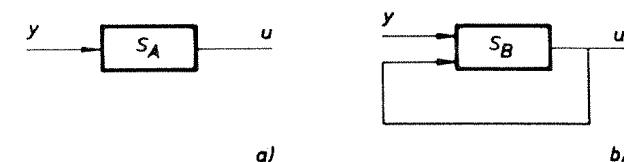


Figura 15.5 Diferentes representaciones de la ley de control.

sistema S_B también es estable. Introduciendo una saturación en el lazo de realimentación de b), el estado del sistema S_B estará siempre acotado si y y u lo están. Este argumento se puede expresar formalmente de la siguiente manera. Se multiplica (15.11) por K y se suma a (15.10). Esto da

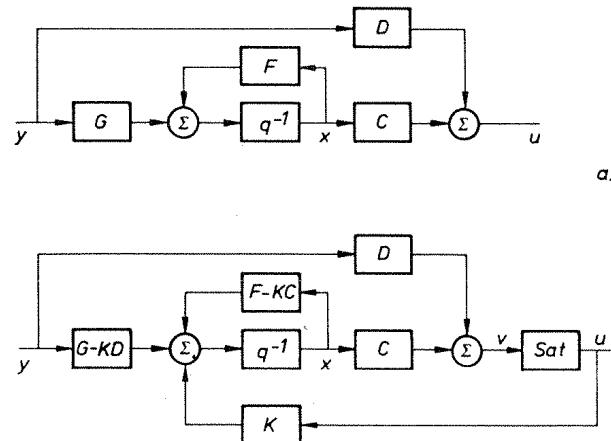
$$\begin{aligned} x(k+1) &= Fx(k) + Gy(k) + K[u(k) - Cx(k) - Dy(k)] \\ &= [F - KC]x(k) + [G - KD]y(k) + Ku(k) \\ &= F_0x(k) + G_0y(k) + Ku(k) \end{aligned}$$

Si el sistema dado por (15.10) y (15.11) es observable, la matriz K siempre se puede escoger de forma que $F_0 = F - KC$ tenga unos valores propios prefijados dentro del círculo unidad. Nótese que esta ecuación es análoga a (15.7). Aplicando los mismos argumentos que para el regulador con un observador explícito, la ley de control se convierte en

$$\begin{aligned} x(k+1) &= F_0x(k) + G_0y(k) + Ku(k) \\ u(k) &= \text{sat}[Cx(k) + Dy(k)] \end{aligned} \quad (15.12)$$

La función de saturación se escoge de forma que se corresponda con la saturación real del actuador. Comparando con el caso del observador explícito se ve que (15.12) corresponde a un observador con dinámica dada por la matriz F_0 . Para señales pequeñas el sistema (15.12) es equivalente a (15.2).

En la Figura 15.6 se puede ver un diagrama de bloques de un regulador con compensación de saturación del término integral.



a)

Figura 15.6 Diagrama de bloques del regulador (15.2) y la modificación de (15.12) que evita la saturación del término integral.

Forma entrada-salida

También se puede realizar una construcción equivalente para reguladores caracterizados por modelos entrada-salida. Considérese un regulador descrito por

$$R(q)u(k) = T(q)u_e(k) - S(q)y(k) \quad (15.13)$$

donde R , S y T son polinomios en el operador desplazamiento. El problema es reescribir la ecuación de forma que se asemeje a un sistema dinámico con la dinámica del observador gobernado por tres entradas, la señal de mando u_c , la salida del proceso y , y la señal de control u . Esto se puede realizar de la siguiente forma.

Sea $A_o(q)$ el polinomio característico deseado del observador. Sumando $A_o(q)u(k)$ a ambos lados de (15.13) se tiene

$$A_o u = Tu_e - Sy + (A_o - R)u$$

Entonces un regulador con compensación de saturación del término integral viene dado por

$$\begin{cases} A_o v = Tu_e - Sy + (A_o - R)u \\ u = \text{sat } v \end{cases} \quad (15.14)$$

Este regulador es equivalente a (15.13) cuando no se satura. Cuando la variable de control se satura se puede interpretar como un observador con dinámica dada por el polinomio A_o .

En la Figura 15.7 se da un diagrama de bloques del regulador lineal (15.13) y de la modificación no lineal (15.14) que evita la saturación del término integral. Un caso

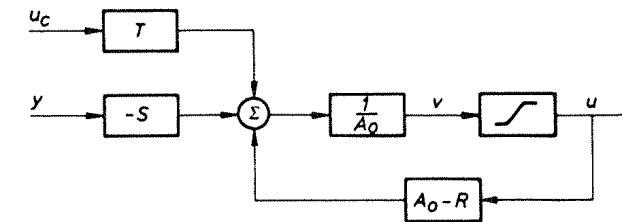
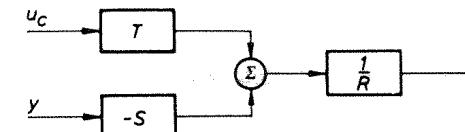


Figura 15.7 Diagrama de bloques del regulador (15.13) y la modificación (15.14) que evita la saturación del término integral.

especialmente simple lo constituye el observador de tiempo finito; es decir, $A_o^* = 1$. Entonces el regulador se puede escribir como

$$u(k) = \text{sat}[T^*(q^{-1})u_e(k) - S^*(q^{-1})y(k) + (1 - R^*(q^{-1}))u(k)] \quad (15.15)$$

Un ejemplo ilustra la realización práctica.

Ejemplo 15.2 – Regulador PI con compensación de saturación del término integral

Un regulador PI discreto tiene como función de transferencia

$$H(z) = \frac{s_0 z + s_1}{z - 1} = \frac{s_0(z - 1) + s_0 + s_1}{z - 1} \quad (15.16)$$

donde

$$s_0 = k$$

$$s_1 = k\left(-1 + \frac{h}{t_i}\right)$$

Sean e el error de control y u la variable de control. Comencemos con la representación de (15.16) en el espacio de estados

$$u(k) = s_0 e(k) + i(k)$$

$$i(k+1) = i(k) + (s_0 + s_1)e(k)$$

Empleando (15.12) con un observador de tiempo finito —es decir, $K=1$ — se tiene

$$i(k+1) = s_1 e(k) + u(k) \quad (15.17)$$

$$u(k) = \text{sat}[s_0 e(k) + i(k)]$$

Se puede obtener el mismo resultado comenzando con (15.15). El algoritmo (15.17) concuerda con el dado en el Ejemplo 8.4.

15.5. Aspectos operacionales

En esta sección se trata la interfase entre el regulador y el operador. Esto incluye una evaluación de la información presentada al operador y los mecanismos de que dispone el operador para cambiar los parámetros del regulador. En el caso de los reguladores analógicos convencionales para control de procesos, es costumbre presentar la consigna, la salida medida y la señal de control. El regulador también se puede pasar de control manual a automático. El operador puede cambiar la ganancia (o banda proporcional), y las constantes de tiempo de integración y de derivación. Esta organización se debía a las propiedades del hardware analógico antiguo. Cuando se emplean computadores para realizar los reguladores, existen muchas otras posibilidades. Hasta ahora el potencial de los computadores se ha usado en una proporción muy baja.

Para tratar la interfase con el operador es preciso considerar cómo se va a utilizar el sistema durante su funcionamiento. De ello se habló en la Sección 7.2 y aquí haremos algunos comentarios adicionales. En primer lugar, es importante caer en la cuenta de la amplia variedad de aplicaciones de los sistemas de control. No hay forma de hacer un tratamiento que los abarque a todos, por lo que se darán algunos ejemplos.

Ejemplo 15.3 – Piloto automático para un avión

Consideremos un piloto automático para un avión de altas prestaciones. En este caso el operador interacciona con el sistema dando señales de mando mediante una palanca y unos pedales. Hay un interruptor de modo para seleccionar entre manual y automático. En modo manual, el piloto automático puede realizar la evaluación de la estabilidad para aumentar el amortiguamiento natural del avión. Puede haber varios modos automáticos, como mantenimiento de la orientación y de la altitud. Al piloto no le interesan los errores de control ni los parámetros del regulador.

Ejemplo 15.4 – Control de procesos

Considérese un sistema de control de procesos con varios lazos que mantienen unas variables de proceso en los niveles deseados. Tradicionalmente se presenta y se almacena el error de regulación. También es corriente llevar a cabo las actualizaciones y cambios muy grandes mediante un control manual. Por tanto, a los reguladores se les dota de un interruptor de modo para escoger entre las operaciones manual o automática.

Para ayudar en la operación normal también se dispone de dispositivos de visualización de referencias, valores medidos y de la variable de control. En el campo de control de procesos hay también diferentes actores: operadores encargados del funcionamiento de la planta, ingenieros de instrumentación que mantienen y ajustan el sistema de control, e ingenieros de proceso que diseñan y modifican las plantas y los instrumentos.

Ejemplo 15.5 – Planta piloto

En la operación de una planta piloto hay necesidad de realizar frecuentes cambios y modificaciones del proceso y del sistema de control. Es necesario entonces disponer de un sistema flexible que permita almacenar gran cantidad de datos, así como llevar a cabo fácilmente cambios y modificaciones en el sistema de control.

Modos de operación

A menudo interesa tener la posibilidad de hacer funcionar un sistema con un control manual. Una sencilla forma de hacerlo es disponer de la estructura mostrada en la Figura 15.8, donde se puede ajustar la variable de control de forma manual. Frecuentemente el control manual se realiza mediante pulsadores para aumentar o disminuir el valor de la variable de control. También es útil disponer de algún artificio que permita modificar la velocidad de variación de la consigna.

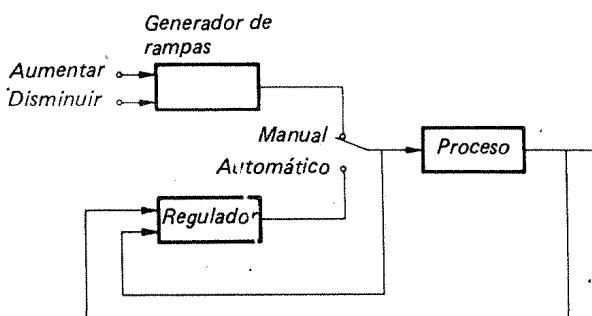


Figura 15.8 Sistema de control con modos manual y automático.

Como el regulador es un sistema dinámico, el estado del regulador debe tener el valor correcto cuando se pasa del modo manual al automático. Si no ocurre así se producirá un transitorio de cambio. Una transición suave recibe el nombre de *transición o transferencia sin salto*.

En los reguladores analógicos convencionales, lo normal es lograr la transición sin salto por la introducción de un *modo de seguimiento*, que ajusta el estado del regulador para que sea compatible con las entradas y salidas dadas para el regulador. Se puede ver el modo de seguimiento como la realización de un observador.

En los reguladores de (15.7) (15.12) y (15.15) el modo de seguimiento se obtiene automáticamente porque llevan incluido en ellos un observador. Para hacerlos funcionar en modo de seguimiento, no hay más que poner

$$u_{\text{inf}} = u_{\text{sup}} = u_{\text{manual}}$$

Esto implica que la señal de control siempre será igual a la señal de entrada manual. El estado del regulador se cambiará automáticamente debido a la realimentación interna del regulador. La saturación introducida en el regulador para resolver la saturación del actuador lleva consigo a la vez la producción de la transferencia sin salto. Hay también formas de lograr un control semiautomático mediante algunas realimentaciones destinadas a estabilizar el sistema.



Con el control con computador también es posible tener otros muchos modos de operación. Por ejemplo, es posible incluir en el regulador algoritmos de estimación de parámetros y de diseño de control. Se puede introducir un *modo de estimación*, en el que se estima un modelo del proceso. El modelo estimado se puede utilizar en el algoritmo de diseño para suministrar en un *modo de sintonizado* una actualización de los parámetros del regulador. También se pueden añadir modos de control adaptativo, en los que se actualizan los parámetros de forma continua. Compárese esto con el regulador auto-sintonizado presentado en la Figura 14.1.

Inicialización

Como un regulador es un sistema dinámico, es importante establecer adecuadamente el estado del regulador cuando se arranca éste. Si no se hace, pueden producirse grandes transitorios de arranque. En el control de procesos convencional con reguladores PI, el regulador tiene un único estado —esto es, el integrador. Lo que se suele hacer es inicializar el regulador manejándole en control manual hasta que la salida del proceso se aproxima lo suficiente al valor deseado.

Para un algoritmo con un observador explícito, el estado del regulador puede inicializarse manteniendo fija la señal de control el tiempo suficiente para que el observador alcance el valor estacionario.

Presentación de información y vigilancia del comportamiento

El control con computador ofrece muchas posibilidades interesantes. Por ejemplo, se puede presentar el comportamiento del lazo de control en lugar de la consigna y del error.

Del Teorema 12.2 se deduce que la salida de un lazo de control bajo un control de varianza mínima tiene una media móvil. Entonces es posible determinar si un lazo de control está bajo un control de varianza mínima con sólo mirar la función de covarianza del error de regulación. Para lazos de control de este tipo lo razonable es presentar al operador la función de covarianza del error.

De forma similar se ha mostrado en la Sección 12.5 que la salida de un regulador LQG viene dada por la ecuación (12.55). Sea y la salida del proceso. La cantidad

$$\epsilon(t) = \frac{P(q)}{R(q)}y(k)$$

será un ruido blanco si el modelo es correcto y el regulador está sintonizado correctamente. Nótese que los polinomios $P(q)$ y $R(q)$ se conocen por el diseño, por lo que tiene sentido presentar la función de covarianza de la señal ϵ . Por supuesto, también será útil presentar valores estadísticos sencillos como el valor medio, la varianza, y los valores máximo y mínimo. También es de utilidad poseer dispositivos para almacenar las entradas y las salidas.

Parametrización y cambios de parámetros

En los controladores convencionales de procesos el operador puede manipular la consigna y los parámetros básicos de la ley de control (ganancia, y constantes de integración y derivación). Con el control con computador existen muchas otras alternativas interesantes. Debido a la sencillez de la realización de los cálculos, es posible emplear una parametrización en el algoritmo de control y otra en la comunicación con el operador. Los parámetros presentados al operador pueden estar relacionados con el comportamiento del sistema en lugar de con el algoritmo de control. La conversión entre los distintos tipos de parámetros se encarga de hacerla un algoritmo en el computador.

Para ilustrar la idea de los parámetros relacionados con el comportamiento, considérese el diseño de un servo empleando el método de asignación de polos descrito en el Capítulo 10. Las propiedades del lazo cerrado se pueden especificar mediante el amortiguamiento relativo ζ y el ancho de banda ω_B . Para realizar el diseño, también hace falta disponer de un modelo de la dinámica en lazo abierto. Una posibilidad es que el ingeniero de procesos indique el ancho de banda y el amortiguamiento deseados y un modelo continuo. El computador puede encargarse de hacer las conversiones necesarias para obtener la ley de control. Si el computador dispone también de un algoritmo de estimación recursivo, no es necesario introducir el modelo. Está claro que hay muchas posibilidades muy interesantes si se incluyen en el regulador algoritmos de estimación y diseño.

Hay dos problemas operacionales con los cambios en línea de los parámetros. Uno está relacionado con la programación en tiempo real. Los datos que representan los parámetros están compartidos por varios programas. Entonces es preciso asegurarse de que no hay un programa utilizando los datos que otro programa está cambiando. Esto se trata en la Sección 15.7.

El otro problema es de naturaleza algorítmica. Pueden aparecer transitorios de cambio cuando se modifican los parámetros en un algoritmo de control. Para comprender qué puede ocurrir, considérese el sencillo algoritmo PI siguiente

```
e := uc - y
u := k*(e + i/ti)
i := i + e*h
```

Es evidente que un cambio en la constante de integración ti producirá un salto en la señal de control a menos que la parte integral, i , sea cero. Se puede evitar el problema cambiando el estado de i a $i*ti/ti'$, donde ti' es el nuevo valor de la constante de integración. Otra forma más sencilla es escribir el algoritmo así

```
e := uc - y
u := k*(e + i)
i := i + e*h/ti'
```

La necesidad de cambiar el estado cuando se cambian los parámetros viene impuesta por el hecho de que el estado del regulador depende de sus parámetros. Una forma de



lograr el cambio de parámetros sin saltos es almacenar un conjunto de datos de entrada-salida antiguos y hacer funcionar un observador cuando se cambian los parámetros del regulador. Sin embargo, suele ser posible emplear una solución más sencilla.

Para ver qué es lo que se debe hacer, considérese el algoritmo (15.7) con un observador explícito y una realimentación del estado. En primer lugar hay que escoger una realización en la que las matrices C y D no dependan de los parámetros a variar. Si el estado x representa una estimación de variables físicas de estado, habrá muy pocas dificultades porque el estado estimado no cambiará drásticamente cuando cambien los parámetros del modelo. Si el error $e = x - \hat{x}$ es distinto de cero no se pueden evitar los transitorios debidos a cambios en la ganancia de realimentación. De forma parecida, con el algoritmo (15.12) no se puede producir ningún transitorio de cambio, porque se trata de una representación en la que las matrices C y D no contienen parámetros que se puedan modificar.

Más complicado de ver es lo que debe hacerse con el algoritmo (15.14). En la representación de (15.14), el estado consiste en las entradas y salidas retardadas. Este estado no es mínimo. Aunque el estado no depende de los coeficientes de los polinomios R , S y T no hay garantía de que los R , S , T , u e y dados sean compatibles con la ecuación (15.14). Con la representación (15.14) habrá transitorios de cambio cuando se varíen los valores de los parámetros.

Seguridad

Es muy importante garantizar de que un sistema de control con computador opera de forma segura. Idealmente, esto significa que el sistema debe dar el resultado correcto o una alarma en caso de que no funcione adecuadamente. Los sistemas con exigencias extremadamente altas pueden triplicarse (o cuadruplicarse) y aceptarse la salida si dos de los subsistemas dan el mismo resultado. Para sistemas más sencillos es suficiente con apoyarse en autocomprobaciones. Existen muchas formas de hacerlas empleando el control con computador. Las unidades aritméticas se pueden comprobar calculando funciones con resultados conocidos. La memoria y la transmisión de datos se pueden comprobar mediante sumas de control. Los convertidores A-D y D-A se pueden comprobar empleando algunos canales adicionales conectados. Se ordena una conversión D-A y se comprueba el resultado de una conversión A-D del mismo canal. También se puede comprobar el temporizado conectando un circuito de constante de tiempo conocida entre un convertidor D-A y otro A-D.

15.6. Aspectos numéricos

Cuando se realiza prácticamente un sistema de control con computador es necesario responder cuestiones como las siguientes: ¿Qué precisión deben tener los convertidores?, ¿qué precisión se requiere en los cálculos?, ¿hay que hacer los cálculos en aritmética de punto flotante o de punto fijo? Para responder a estas preguntas, es

necesario comprender los efectos de la limitación y estimar sus consecuencias para el sistema en lazo cerrado. No se trata de un asunto trivial porque la respuesta depende de una interacción compleja entre la realimentación, el algoritmo y la frecuencia de muestreo. Afortunadamente, sólo es necesario realizar estimaciones groseras. Por ejemplo, ¿la resolución debe ser de 10 ó 12 bits y el tamaño de palabra de 24 ó 32 bits? Estas preguntas se pueden responder mediante un análisis simplificado.

Fuentes de error

Las fuentes de error más importantes son las siguientes:

Cuantificación en los convertidores A-D.

Cuantificación de los parámetros.

Redondeo, desbordamiento y desbordamiento negativo en sumas, restas, multiplicaciones, divisiones, evaluación de funciones y otras operaciones.

Cuantificación en los convertidores D-A.

Los tipos corrientes de convertidores A-D tienen precisiones de 8, 10, 12 y 14 bits, lo que corresponde a una resolución de 0,4 %, 0,1 %, 0,025 % y 0,006 %. Los tantos por ciento están dados respecto a la totalidad de la escala. Los convertidores D-A también tienen una precisión limitada. Lo típico es una exactitud de 10 bits. El error por la cuantificación de los parámetros depende de forma crítica del período de muestreo y de la realización escogida para la ley de control.

Longitud de palabra

Los algoritmos de control digital se suelen realizar mediante microcomputadores y minicomputadores, que tienen longitudes de palabra de 8, 16 ó 32 bits. En aplicaciones como la lanzadera espacial o en productos especiales que se fabrican en cantidades muy grandes, se pueden emplear computadores de propósito especial, en los que se puede elegir la longitud de palabra que se deseé.

Hay muchas diferencias en las representaciones de números. Son corrientes las siguientes:

Punto fijo, simple precisión, 16 bits.

Punto fijo, doble precisión, 32 bits.

Punto flotante, simple precisión, exponente de 8 bits, mantisa de 24 bits.

Punto flotante, doble precisión, exponente de 8 bits, mantisa de 56 bits.

Mediante un ejemplo se ilustra la esencia de la dificultad numérica.

Ejemplo 15.6 – Cálculo de un producto escalar

Considérense los vectores

$$a = [100 \ 1 \ 100]$$

$$b = [100 \ 1 \ -100]$$

El producto escalar es $\langle a, b \rangle = 1$. Si se calcula el producto escalar con una representación en punto flotante y una precisión correspondiente a tres decimales, el resultado será cero porque $100 \cdot 100 + 1 \cdot 1$ se redondeará a 10.000. Nótese que el resultado obtenido depende del orden de las operaciones. Las operaciones con longitud de palabra finita no son ni asociativas ni distributivas.

Se puede evitar esta dificultad sin emplear todo el cálculo en doble precisión sumando los términos en doble precisión y después redondeando a simple precisión. Este método se puede emplear en los cálculos de punto fijo y punto flotante. Nótese que en muchos computadores la instrucción de multiplicar se realiza de tal forma que el producto se obtiene en doble precisión. También muchos lenguajes de alto nivel tienen estructuras que permiten este tipo de cálculos. En términos generales, el redondeo y la cuantificación dan lugar a errores pequeños mientras que los efectos del desbordamiento son desastrosos.

Visión general de los efectos de redondeo y cuantificación

Las consecuencias del redondeo y de la cuantificación dependen del sistema de realimentación y de los detalles del algoritmo. Se puede influir de forma considerable sobre las propiedades cambiando la representación de la ley de control o los detalles del algoritmo. Por eso es importante comprender el fenómeno.

Una descripción detallada del redondeo y de la cuantificación implica el empleo de modelos no lineales complicados, que son muy difíciles de analizar. Sin embargo, el estudio de casos muy sencillos muestra que la cuantificación y el redondeo pueden conducir a oscilaciones de ciclo límite. Más tarde se presentan ejemplos de esto, junto con un análisis aproximado. También se han observado las oscilaciones de ciclo límite en casos más complejos.

Algunas propiedades de la cuantificación y el redondeo en un sistema con realimentación también pueden verse mediante un análisis lineal. Entonces la cuantificación y el redondeo se modelan como operaciones ideales con perturbaciones sumadas o multiplicadas. Las perturbaciones pueden ser determinísticas o estocásticas. Este tipo de análisis es especialmente útil para la estimación del orden de magnitud. Permite la investigación de sistemas complejos y es útil para comparar diferentes algoritmos.

También son útiles para determinar la sensibilidad de los algoritmos a los cambios de parámetros las técnicas del análisis de sensibilidad y el análisis numérico. Estos métodos se pueden emplear para comparar y visualizar diferentes algoritmos. Sin embargo, los métodos están limitados a la comparación del comportamiento de los algoritmos en lazo abierto. También es necesario comparar los efectos de la cuantificación y el redondeo con las otras perturbaciones en el sistema.

Análisis no lineal

El aspecto no lineal de los errores de cuantificación y redondeo se puede ilustrar mediante un sencillo ejemplo. Este muestra cómo la cuantificación y el redondeo pueden dar lugar a equilibrios múltiples y ciclos límites.

Ejemplo 15.7 – Ciclos límites debidos al redondeo

Considérese el siguiente modelo de un sistema sencillo con cuantificación:

$$y(k+1) = u(k) + Q[ay(k)]$$

donde la función Q representa el redondeo a enteros con las reglas normales. Supóngase que la entrada u es cero. La respuesta del sistema viene dada por la ecuación

$$y(k+1) = Q[ay(k+1)] \quad (15.18)$$

Si no hay cuantificación —es decir, $Q(y)=y$ — la ecuación tiene un punto de equilibrio, $y=0$. Este equilibrio es estable si $|a|<1$. Cuando hay cuantificación una gráfica ayuda a encontrar el equilibrio. La Figura 15.9a) muestra la gráfica de ay para $a>0$ y la imagen inversa de $Q[y]$. El equilibrio corresponde a la intersección de las gráficas. De aquí se tiene que un equilibrio es $y=0$. Si existen enteros $k>0$ tales que

$$\begin{cases} ak > k - 0,5 & k \text{ impar} \\ ak \geq k - 0,5 & k \text{ par} \end{cases}$$

también hay otros equilibrios.

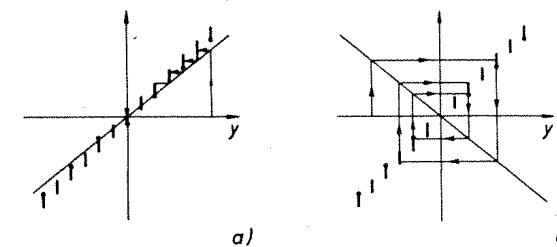


Figura 15.9 Gráfica para encontrar el equilibrio de la ecuación (15.18). El parámetro a tiene el valor 0,8 en a) y -0,8 en b).

Un análisis similar para $-1 < a \leq 0$ se presenta en la Figura 15.9b). La figura muestra que puede haber ciclos límites con un período de dos períodos de muestreo si

$$\begin{cases} ak > -k + 0,5 & k \text{ impar} \\ ak \geq -k + 0,5 & k \text{ par} \end{cases}$$

El ejemplo presenta algunos efectos del redondeo sobre un sistema. En general, puede haber múltiples equilibrios y soluciones periódicas con diferentes amplitudes y períodos. Es difícil analizar exactamente sistemas más complicados. Sin embargo, hay métodos aproximados que dan alguna idea de lo que ocurre.

Análisis con la función descriptiva

Si sólo hay en el lazo una no linealidad, se puede emplear el método de la función descriptiva para determinar de forma aproximada los ciclos límites.

Considérese el sistema de la Figura 15.10a). Se puede ver el método de las funciones descriptivas como una generalización del criterio de Nyquist. El punto crítico -1 se sustituye por $-1/Y_c(A)$, donde $Y_c(A)$ es la función que describe la no linealidad. La función descriptiva caracteriza la transmisión de una señal sinusoidal con amplitud A a través de la no linealidad. El método predice un ciclo límite si

$$H(e^{j\omega h}) = -1/Y_c(A)$$

[compárese con la Fig. 15.10b)]. La frecuencia, ω_1 , y la amplitud, A_1 , en la intersección son la frecuencia y la amplitud estimadas del ciclo límite.

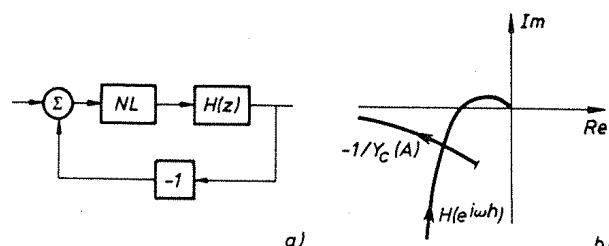


Figura 15.10 a) Sistema discreto con una no linealidad NL .
b) Empleo del método de la función descriptiva.

La función descriptiva de un cuantificador de redondeo es

$$Y_c(A) = \begin{cases} 0 & 0 < A < \frac{\delta}{2} \\ \frac{4\delta}{A} \sum_{i=1}^n \sqrt{1 - \left[\frac{2i-1}{2A}\delta \right]^2} & \frac{2n-1}{2}\delta < A < \frac{2n+1}{2}\delta \end{cases}$$

Esta función se ha dibujado en la Figura 15.11.

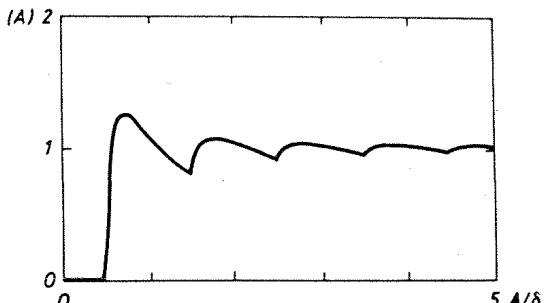
Ejemplo 15.8

Considérese el sistema del Ejemplo 5.3 con la función de transferencia discreta

$$H(z) = \frac{0.25K}{(z-1)(z-0.5)}$$

En el Ejemplo 5.3 se ha visto que el sistema en lazo cerrado sin cuantificación es asintóticamente estable si $K < 2$. Con una cuantificación de la señal de error, el método de

Figura 15.11 Función descriptiva del redondeo.



la función descriptiva predice que habrá un ciclo límite si K es mayor que 1.3 aproximadamente. La Figura 15.12 muestra el comportamiento del sistema para un nivel de cuantificación de $\delta=0.2$. Se aprecia claramente el ciclo límite para $K=1.6$.

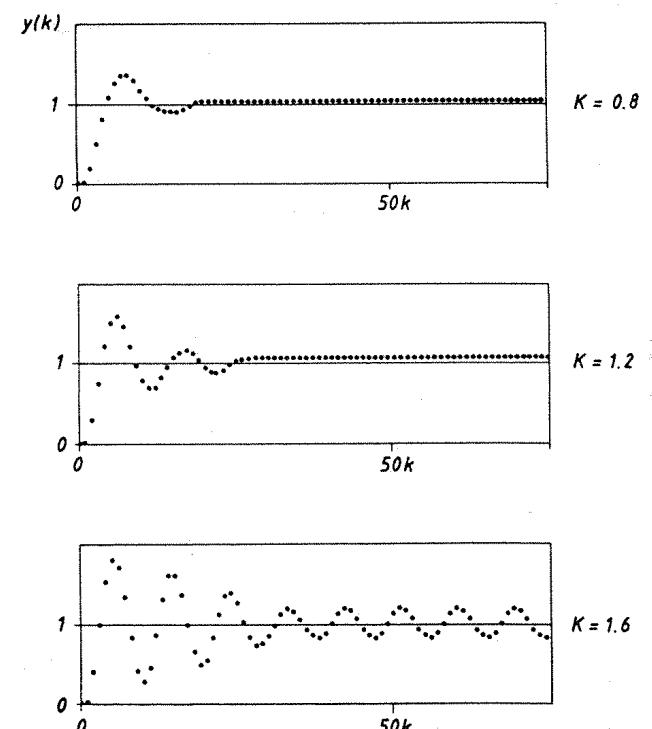


Figura 15.12 Salida del sistema del Ejemplo 15.8 cuando $\delta=0.2$ y $K=0.8, 1.2$ y 1.6 .

Análisis lineal

También se pueden estimar los efectos de la cuantificación y el redondeo mediante el análisis lineal. La idea es representar las operaciones por sus modelos ideales y por una perturbación aditiva e_a . Los convertidores A-D y D-A se representan simplemente por ganancias lineales con una perturbación que modela la cuantificación. Con cálculos en punto fijo, las sumas son exactas. Sin embargo, habrá errores en las multiplicaciones. Estas se representarán como multiplicaciones exactas, con un error aditivo, que representa el redondeo. En la Figura 15.13 se pueden ver ilustrados estos modelos.

Los errores se pueden modelar como señales determinísticas o estocásticas. En un modelo determinístico, el error se modela como constantes que tienen los valores de los errores de cuantificación y con la resolución de los cálculos aritméticos. En el modelo estocástico, el error introducido por el redondeo o la cuantificación se describe como un ruido blanco aditivo con distribución rectangular. Entonces se supone que los errores en diferentes instantes de muestreo no están correlacionados. Si la cuantificación se realiza como un redondeo, el error está distribuido de forma igual sobre el intervalo $(-\delta/2, \delta/2)$, donde δ es el paso de cuantificación. Si la cuantificación se hace como una truncación, el error se distribuye de forma igual sobre $(0, \delta)$.

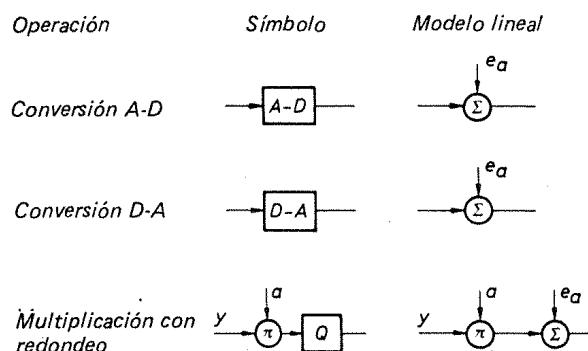


Figura 15.13 Modelos lineales de la cuantificación y el redondeo.

Un ruido rectangular distribuido sobre un intervalo de longitud δ tiene una varianza de $\delta^2/12$.

Utilizando los modelos lineales del redondeo y de la cuantificación, es posible reducir el problema de estimar el efecto de la cuantificación y del redondeo al problema de calcular las respuestas de un sistema lineal a entradas determinísticas o estocásticas. En la Sección 6.5 se han expuesto métodos para hacerlo. Empleando el modelo lineal es posible evaluar cuantitativamente los efectos de la cuantificación sin necesidad de cálculos detallados. También es fácil comparar el redondeo con otras

perturbaciones del sistema. En el modelo lineal el efecto del redondeo en el convertidor A-D es el mismo que el efecto del ruido de medida. El efecto sobre la señal de control puede ser sustancial para aquellas frecuencias en las que el regulador tiene una ganancia alta. El efecto del redondeo en el convertidor D-A es el mismo que una perturbación en la entrada del proceso. Debido a que normalmente el proceso atenúa las altas frecuencias, el efecto sobre la salida del proceso suele ser pequeño. Recuérdese, sin embargo, que el modelo lineal no abarca todos los aspectos del redondeo.

Análisis de sensibilidad

Tanto el análisis lineal como el no lineal tienen en cuenta los efectos del redondeo y de la cuantificación en el lazo de realimentación. Hay otros métodos más sencillos para estimar los errores en una instalación en lazo abierto. Aunque estos enfoques no tienen en cuenta la realimentación, permiten la comparación de diferentes algoritmos.

Se pueden obtener estimaciones groseras de los errores en los cálculos mediante la idea de los números de condición del análisis numérico. Esta noción se puede explicar de la siguiente forma. Considérese el problema de evaluar una función f . El número de condición, o la sensibilidad del problema, se define como el número C_p en la desigualdad

$$\frac{\|f(x + \delta x) - f(x)\|}{\|f(x)\|} \leq C_p \frac{\|\delta x\|}{\|x\|}$$

El número de condición permite relacionar el error relativo en el resultado con el error relativo en los datos.

Un algoritmo para resolver un problema es una secuencia finita de operaciones elementales especificadas (sumas, multiplicaciones, etc.) que generan una solución del problema. Un algoritmo a para evaluar la función f puede verse como una secuencia finita de aplicaciones

$$a = a_{(r)} \cdot a_{(r-1)} \cdots a_{(1)} \cdot a_{(0)}$$

donde cada $a_{(i)}$ representa una operación elemental. El número de condición para el algoritmo a se define como sigue:

$$C_a = \sup_x \frac{\|\delta \xi\|}{\|x\|}$$

donde

$$a(x) = f(x + \delta \xi)$$

El número de condición del algoritmo da entonces una estimación del error medido en unidades de error relativo en los datos de entrada. Los números de condición C_p y C_a se pueden emplear para estimar el error relativo mediante la fórmula

$$\frac{\|a(x + \delta x) - a(x)\|}{\|a(x)\|} \leq C_p \left\{ \frac{\|\delta x\|}{\|x\|} + C_a \right\}$$

cuando se evalúa la función f empleando el algoritmo a . Los números de condición son útiles para estimar qué sensibilidad tienen los problemas y los algoritmos a pequeños cambios en los datos. El error en un cálculo depende entonces tanto del problema como del algoritmo.

Cuando se realiza prácticamente una ley de control mediante un computador digital, el problema es la evaluación de la salida de un sistema dinámico sometido a una entrada. Los datos son la señal de entrada y los parámetros. El número de condición del problema puede verse influenciado en gran medida por la realización y por la parametrización escogidas. Entonces es importante considerar realizaciones, parametrizaciones y algoritmos.

Diferentes realizaciones

Una ley de control es un sistema dinámico. Se pueden obtener diferentes realizaciones transformando las coordenadas del espacio de estados. La elección de una realización adecuada es muy importante para la condición de la solución. En particular, las formas asociadas son muy malas desde un punto de vista numérico.

El análisis que sigue proporciona resultados cuantitativos para la sensibilidad de la forma asociada. Considérese un filtro lineal con polos en p_i distintos y el polinomio característico

$$A(z) = (z - p_1) \cdots (z - p_n) = z^n + a_1 z^{n-1} + \cdots + a_n$$

El polinomio característico A puede verse como una función de z y de a_i . Cuando el parámetro a_i se cambia al valor $a_i + \delta a_i$, los polos cambian de p_k a $p_k + \delta p_k$. Por tanto,

$$0 = A(p_k + \delta p_k, a_i + \delta a_i) \approx A(p_k, a_i) + \frac{\delta A}{\delta z} \Big|_{p_k} \delta p_k + \frac{\delta A}{\delta a_i} \Big|_{p_k} \delta a_i + \cdots$$

El primer término del miembro de la derecha es cero. Si se desprecian los términos de segundo orden y superiores, se tiene

$$\delta p_k \approx -\frac{\delta A / \delta a_i}{\delta A / \delta z} \Big|_{z=p_k} \cdot \delta a_i$$

Debido a que

$$\frac{\delta A}{\delta a_i} \Big|_{z=p_k} = p_k^{n-i}$$

y a

$$\frac{\delta A}{\delta z} \Big|_{z=p_k} = \prod_{j \neq k} (p_k - p_j)$$

se obtiene la siguiente estimación:

$$\delta p_k \approx -\frac{p_k^{n-i}}{\prod_{j \neq k} (p_k - p_j)} \delta a_i \quad (15.19)$$

Si el polinomio tiene una raíz p_k con multiplicidad m , la ecuación (15.19) se convierte en

$$\delta p_k \approx -\frac{p_k^{n-i}}{\prod_{j \neq k} (p_k - p_j)} (\delta a_i)^{1/m} \quad (15.20)$$

Si el filtro es estable, entonces $|p_k| < 1$ y el numerador de (15.19) será de módulo máximo para $i=n$. Entonces el coeficiente a_n será el parámetro más sensible. Además, el denominador será pequeño si los polos están próximos, lo que hace al sistema sensible a los cambios en los coeficientes. La ecuación (15.20) muestra que la sensibilidad es incluso mayor si el polinomio tiene raíces múltiples. Las ecuaciones (15.19) y (15.20) se pueden emplear para determinar los números de condición para la transformación de forma diagonal a forma asociada. De las ecuaciones se deduce que el cálculo de las formas asociadas es un problema de mala condición.

Ejemplo 15.9

Supóngase que se quiere realizar un sistema con la función de transferencia discreta

$$H(z) = \frac{10^{-8}}{(z - 0.99)^4} \quad (15.21)$$

Para obtener el programa del computador primero se determina una realización en el espacio de estados de la función de transferencia discreta. Entonces se obtiene el código para el computador como una realización directa de las ecuaciones en diferencias. Existen muchas elecciones posibles del sistema de coordenadas en la realización en el espacio de estados; se escogen una forma canónica controlable y una forma canónica de Jordan para

LISTADO 15.3 Programa para la realización de (15.21) basada en la forma canónica controlable.

```

begin
  y:=x4
  s:=-a1*x1-a2*x2-a3*x3-a4*x4+b4*u
  x4:=x3
  x3:=x2
  x2:=x1
  x1:=s
end

```

demonstrar que puede haber diferencias considerables en las propiedades numéricas. En los Listados 15.3 y 15.4 se dan los programas para las dos representaciones. La realización del sistema discreto incluye también un sistema monitor que hace funcionar el programa cada período de muestreo. Nótese que los programas sólo contienen sumas, multiplicaciones e instrucciones de asignación; entonces pueden realizarse fácilmente mediante una gran variedad de lenguajes de programación. Como las instrucciones de asignación sólo transfieren datos, no introducen errores numéricos. Esto significa que en una representación normal de una matriz los valores 0 y 1 se representan exactamente.

LISTADO 15.4 Programa para realizar (15.21) basada en la forma canónica de Jordan.

```

begin
    y:=x4
    x4:=-a*x4+x3
    x3:=-a*x3+x2
    x2:=-ax*x2+x1
    x1:=-a*x1+b*u
end

```

Los valores numéricos de los parámetros de la forma canónica controlable son

$$\begin{aligned} a_1 &= -3.96 & a_2 &= 5.8806 & b_1 = b_2 = b_3 &= 0 \\ a_3 &= -3.881196 & a_4 &= 0.96059601 & b_4 &= 10^{-8} \end{aligned}$$

La Figura 15.14 muestra la solución exacta y la solución obtenida con un VAX 11/780 cuando se emplea la forma controlable, (3.30). Nótese la extremada sensibilidad de los resultados. Con simple precisión se obtiene la curva *b*, que indica una realización inestable; sin embargo, con la doble precisión se consigue la misma curva que la solución analítica exacta. El sistema es muy sensible a los errores en los coeficientes, incluso empleando doble precisión (véanse las curvas *c* y *d*). Con la realización basada en la forma canónica de Jordan no se presentan problemas de sensibilidad.

La ecuación (15.20) da la sensibilidad de la ecuación característica respecto a los cambios en los parámetros. Si se perturba la ecuación característica con un término constante ϵ ,

$$(z - 0.99)^4 + \epsilon = 0$$

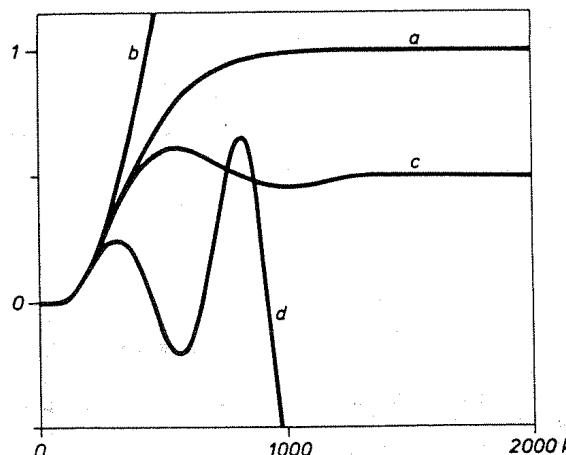


Figura 15.14 a) Respuesta exacta de (15.21) a un salto.
b) Simulación de (15.21) empleando la realización controlable y cálculos en simple precisión en un VAX 11/780.
c) Lo mismo que en b), pero con doble precisión y $a_4 = 0.96059602$.
d) Lo mismo que en b), pero con doble precisión y $a_3 = -3.8811959$.

que tiene las raíces

$$z = 0.99 + (-\epsilon)^{1/4}$$

Las raíces se mueven desde 0.99 a un círculo con centro en 0.99 y radio $r = |\epsilon|^{1/4}$. Si $\epsilon = 10^{-8}$ entonces $r = 10^{-2}$; es decir, el sistema puede ser inestable incluso aunque la perturbación sea muy pequeña.

En este caso el problema es serio porque hay valores propios múltiples cerca del círculo unidad. Esto ocurre, por ejemplo, si se emplea una frecuencia de muestreo rápida.

Realizaciones de buena condición

La dificultad que aparece con la forma asociada se puede evitar simplemente representando el sistema como una combinación de sistemas de primer y segundo orden (véase la Fig. 15.15).

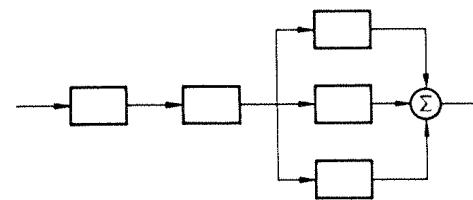


Figura 15.15 Realización de una ley de control mediante conexiones en serie o en paralelo de bloques de primer y segundo orden; es decir, cada bloque de la figura es un sistema de primer o segundo orden.

Si el sistema dinámico que representa al regulador tiene nr polos reales distintos y nc pares de polos complejos, el algoritmo de control se puede transformar a la forma del modelo

$$\begin{cases} z_i(k+1) = \lambda_i z_i(k) + \beta_{i1} y(k) & i = 1, \dots, nr \\ v_i(k+1) = \begin{bmatrix} \sigma_i & \omega_i \\ -\omega_i & \sigma_i \end{bmatrix} v_i(k) + \begin{bmatrix} \gamma_{i1} \\ \gamma_{i2} \end{bmatrix} y(k) & i = 1, \dots, nc \\ u(k) = D y(k) + \sum_{i=1}^{nr} \gamma_i z_i(k) + \sum_{i=1}^{nc} \delta_i v_i(k) \end{cases} \quad (15.22)$$

donde los polos complejos se representan mediante variables reales. Nótese que z_i son escalares y v_i son vectores con dos elementos. Para evitar las dificultades numéricas, la ley de control debe transformarse a la forma (15.22), que se realiza prácticamente en el computador de control. La transformación se puede hacer fácilmente mediante un paquete de programas de diseño con ayuda de computador. Con ecuaciones de la forma de (15.22) es fácil emplear cálculos en punto fijo y escalados.

Si la ley de control tiene valores propios múltiples, una forma canónica de Jordan sustituye a (15.22). A un valor propio λ de multiplicidad 3 le corresponde un bloque

$$z(k+1) = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} z(k) + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} y(k)$$

Efectos del período de muestreo

El período de muestreo también tiene una considerable influencia sobre la condición, como se ve en los ejemplos siguientes.

Ejemplo 15.10 – Efecto del período de muestreo sobre la precisión de los coeficientes

Considérese un sistema de primer orden con una constante de tiempo T . El equivalente discreto de este sistema es

$$x(kh + h) = ax(kh) + bu(kh)$$

donde

$$a = e^{-h/T}$$

Mediante cálculos sencillos se ve que

$$\frac{dT}{T} = -\frac{T}{h} \frac{da}{a}$$

Para una precisión relativa dada, la constante de tiempo equivalente es inversamente proporcional al período de muestreo.

Ejemplo 15.11 – Precisión numérica requerida para el control PI

Considérese la fórmula de actualización de la integral en un regulador PI:

$$i(kh + h) = i(kh) + e(kh) * h/ti$$

Si el período de muestreo es 0,03 s y la constante de integración es 15 m=900 s, la relación h/ti resulta $3 \cdot 10^{-5}$, lo que equivale a 15 bits. Para evitar esto se redondea la cantidad eh/ti . Por tanto, es preciso realizar los cálculos con una longitud de palabra mayor. Esta es la razón por la que en los reguladores PI específicos la integración se suele realizar con una representación de 24 bits.

Los ejemplos muestran que un muestreo rápido exige una elevada precisión en los coeficientes.

15.7. Programación

La programación es un aspecto importante de la realización práctica de un sistema de control, tanto respecto a la eficacia del sistema como al tiempo requerido para la realización.

El esfuerzo requerido y el enfoque utilizado dependen del software disponible y de la naturaleza del problema de control. Cuando se introdujo por primera vez el control con computador no se disponía de lenguajes de alto nivel para computadores de control. La potencia de cálculo era tan reducida que era necesario programar en ensamblador. Hoy día la situación ha cambiado de forma drástica, y existen muchos lenguajes de alto nivel utilizables, como Forth, BASIC, FORTRAN, C, Jovial, Pascal, Pascal Concurrente y Ada. Ada, que ha sido desarrollado por el Departamento de

Defensa de Estados Unidos para aplicaciones de control con computador, es el primer lenguaje diseñado y desarrollado para aplicaciones en tiempo real. La utilización de un lenguaje de alto nivel tiene muchas ventajas porque lleva mucho menos tiempo el desarrollar un sistema. El programa es mucho más legible, lo que lo hace mucho más fácil de modificar. El carácter y la dificultad de la programación dependen en gran medida de la aplicación. Los requisitos de la comunicación con el operador son críticos. A menudo el programa necesario para realizar la comunicación con el operador es mucho más largo que el programa exclusivamente de control. Esto se puede ver en algunos ejemplos.

Un sencillo sistema de control específico

Considérese un lazo sencillo de control que tiene unas pocas señales medidas, unas pocas salidas y una comunicación con el operador limitada. Al operador se le puede presentar la información y éste puede tener a su alcance unos cuantos pulsadores y ruedas selectoras.

La programación de un sistema de este tipo es muy sencilla. Si se dispone de un reloj de tiempo real, el Listado 15.5 presenta un esquema de programa.

LISTADO 15.5 Esbozo de un programa para una ley de control sencilla.

```
repeat
  Wait for clock interrupt
  Regulate
  Display
  forever
```

La primera línea es solamente un procedimiento que detiene la ejecución hasta que se produce una interrupción del reloj. El procedimiento *Regulare* es el programa necesario para realizar el control deseado. En el Listado 15.6 se puede ver un ejemplo típico.

LISTADO 15.6 Esbozo de un programa para el algoritmo de control con las señales de mando incluidas.

```
Procedure Regulate
begin
  1 Adin y uc umanual
  2 Din manual
  3 If manual then
  4   u := ulow:=uhigh:=umanual
  5 else
  6   u := u1 + D * y + Dc * uc
  7   u := sat(u, ulow, uhigh)
  8 Daout u
  9   x := Fo * x + Go * y + Gc * uc + K * u
 10  u1 := C * x
end
```

La instrucción *Din* en la línea 2 lee una señal digital cuyo valor sitúa en la variable Booleana *manual*. Si la variable es cierta, el regulador opera en modo manual y la variable de control es igual a la variable *umaniual*, que es una entrada analógica. Nótese que el programa es similar al Listado 15.2, excepto en las entradas provenientes del operador y en la función de saturación que se ha añadido para evitar la saturación del término integral.

El procedimiento *Display* del Listado 5.7 calcula algunas variables y las presenta en forma analógica o digital. Nótese que es inmediato la introducción de dispositivos para que el operador cambie los parámetros con sólo introducirlos como entradas analógicas.

El programa del Listado 15.5 es claramente fácil de depurar. Los procedimientos *Regulare* y *Display* son procedimientos secuenciales sencillos que se pueden comprobar fuera de línea. También es fácil de comprobar que el procedimiento de espera proporciona una interrupción todos los períodos de muestreo. Sólo fallará si el tiempo necesario para ejecutar los procedimientos es mayor que un período de muestreo. Esto se comprueba fácilmente midiendo los tiempos.

Lazos de control más complicados

Los principios empleados en el programa del Listado 15.5 se pueden extender a sistemas de control más complejos con varios lazos que tienen diferentes períodos de muestreo. Se obtiene entonces un programa, que puede representarse mediante el diagrama de bloques de la Figura 15.16. El programa P0 de la Figura 15.16 funciona a la frecuencia de muestreo dada por el reloj. Los programas P1, P2 y P3 funcionan cada tercer pulso de reloj. Para conseguir la representación de la Figura 15.16 es necesario que el tiempo requerido para ejecutar cada una de las ramas sea menor que el menor de los períodos de muestreo. Esto se puede lograr fácilmente en los sistemas con períodos de muestreo largos. Para sistemas con muestreo rápido, puede que sea

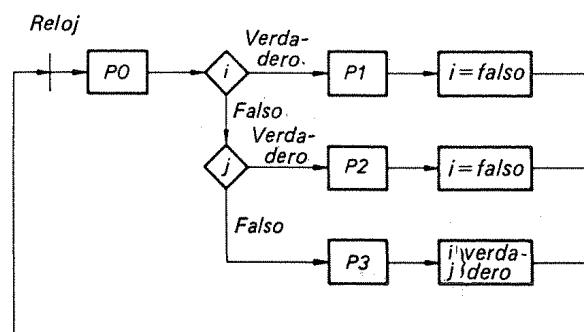


Figura 15.16 Diagrama de flujo de una ley de control multilazo con dos frecuencias de muestreo.

preciso descomponer los cálculos de una forma tediosa, antinatural y propensa a presentar errores.

Es fácil comparativamente depurar el programa presentado en la Figura 15.16 si tiene pocas ramas y si los procedimientos son sencillos. La dificultad de la depuración crece rápidamente al crecer la complejidad del sistema. Se necesitan nuevas ideas y conceptos para tratar estos problemas de forma conveniente.

Programación concurrente, o en tiempo real

Es normal considerar los lazos de control como actividades concurrentes que están funcionando en paralelo. Sin embargo, el computador digital opera secuencialmente en el tiempo. Los lenguajes de programación ordinarios sólo pueden representar actividades secuenciales. Entonces un problema clave es como convertir un número de actividades paralelas concurrentes en un programa secuencial. Esto se puede hacer manualmente, como se ha visto en la Figura 15.16. También hay software de propósito especial —sistemas operativos de tiempo real— que ayudan en la tarea. Sería salirse del propósito de este libro tratar detalladamente la programación concurrente. Se darán, sin embargo, las ideas básicas junto con algunos ejemplos.

Para la programación en tiempo real son conceptos fundamentales las nociones de *proceso* y de *tarea*. Representan actividades que pueden concebirse como funcionando en paralelo en el tiempo. Empleando estas nociones es posible pensar en el computador ejecutando varias actividades en paralelo. Por tanto, una actividad en tiempo real se puede estructurar de la misma forma que se estructura una actividad secuencial, mediante las nociones de procedimiento, o de subrutina. El sistema operativo en tiempo real organizará la ejecución de los procesos de forma que se obtenga el resultado deseado. Para poder hacer esto, a cada proceso se le asocia una prioridad. Los procesos se pueden planificar para ejecutarse periódicamente o como respuesta a sucesos tales como interrupciones o terminación de otras tareas.

Uno de los problemas clave de la programación en tiempo real es el de las *variables compartidas* y el de los recursos. Si procesos diferentes están usando los mismos datos hay que estar seguros de que un proceso no está intentando utilizar datos que están siendo modificados por otro proceso. Si dos procesos pueden utilizar el mismo recurso, hay que asegurarse de que el sistema no se *bloquea* en una situación en la que ambos procesos están esperando una acción del otro.

Un tercer problema es el del tiempo. La potencia de cálculo debe ser suficiente para que todas las actividades se completen en el tiempo requerido.

Un regulador con interacción con el operador

Uno de los ejemplos más sencillos de programación en tiempo real es un único lazo de control con interacción con el operador. La tarea puede ser hacer funcionar un lazo de control como el del Listado 15.4 con una frecuencia de muestreo de 20 ms y con una interfase con el operador tal que el operador pueda cambiar los parámetros desde un teclado o un terminal. Supóngase que el operador cambia los parámetros

escribiendo una cadena de caracteres en el terminal. Como el tiempo requerido para esto es considerablemente mayor que un período de muestreo, es preciso descomponer la operación en muchas partes pequeñas para emplear la solución que se presentó en la Figura 15.16. Esto es a la vez tedioso y antinatural. Es mucho más natural plantear el problema como dos procesos concurrentes. Un proceso, la *regulación*, debe ejecutarse una vez cada período de muestreo. El otro, la *comunicación con el operador*, puede ejecutarse mientras el proceso *regulación* está parado. Para garantizar que las acciones de control se llevan a cabo en períodos de muestreo regulares es preciso imponer la condición de que el proceso *regulación* tenga prioridad sobre la *comunicación con el operador* y que pueda interrumpir a la comunicación con el operador en cualquier momento. También se introduce, porque es conveniente, la norma de que el proceso *regulación* una vez que ha comenzado se ejecute hasta el final. En un caso como éste, al proceso *regulación* se le llama *tarea de primer plano*, o proceso de primer plano, y a la *comunicación con el operador* se le llama *tarea de fondo*, o proceso de fondo.

En el Listado 15.7 se da un programa, escrito en Basic, de Procesos que realiza las tareas que se han expuesto.

El Basic de Procesos dispone de una serie de entidades para entradas y salidas analógicas. También se pueden definir tareas o procesos, la forma de activarlos, y asignarles prioridades. En la línea 140 se define que la tarea 1 comienza en la línea 200. En la línea 170 se activa la tarea 1 con período *H* y prioridad 1. La línea 180 indica que la tarea 2 se activa cuando se pulsa una tecla cualquiera del teclado.

El Basic de Procesos tiene un programa monitor que supervisa la ejecución del programa. Después de la ejecución de cada línea del programa el control de la ejecución se transfiere al monitor. Entonces éste investiga si la ejecución debe proseguir secuencialmente según los números de línea o si debe ejecutarse una tarea con mayor prioridad. Como la tarea 1 tiene mayor prioridad que la tarea 2, la tarea 1 siempre se ejecutará hasta el final una vez comenzada su ejecución. Se activará la ejecución de la tarea 1 siempre que hayan pasado más de *H* segundos desde la activación anterior. Supóngase que el programa ejecuta la línea 330 cuando se ha realizado la tarea 1. Cuando se ha completado la ejecución de la línea, el monitor transfiere el control a la línea 200 y entonces se ejecutan las líneas de la 200 a la 280 antes de transferir el control a la línea 340. Entonces es posible que haya un retraso en la ejecución equivalente al tiempo de ejecución de una línea.

Nótese que las variables *H*, *S0* y *S1* se comparten entre las dos tareas. Estas variables las pone la *comunicación con el operador* (tarea 1) y las usa la *regulación* (tarea 2). Cuando se cambian las variables es importante que se modifiquen los valores de todas las variables antes de que se usen en la regulación.

Como los cambios de todas las variables se hacen en la línea 370 y se completa la ejecución de cada línea antes de que se devuelva el control al monitor, está garantizado que todas las variables se cambian antes de que sean usadas en la regulación. Nótese que pudiera no ser así si las asignaciones de variables se hicieran en líneas separadas.

El convenio de que la ejecución de cada línea se completa antes de devolver el control al monitor facilita el control de las variables compartidas. Sin embargo, se

LISTADO 15.7 Código en Basic de Procesos de un regulador PI con comunicación con el operador.

```

100 REM *****
110 REM *** REGULADOR PI ***
120 REM *****
130
140 TASK 1, 200:REM REGULADOR
150 TASK 2, 300:REM COMUNICACION CON EL OPERADOR
160 GOSUB 400:REM INICIALIZAR
170 ACTIVATE TASK 1 WITH PERIOD H AND PRIORITY 2
180 ACTIVATE TASK 2 ON KEYPRESS WITH PRIORITY 5
190
200 REM ** REGULACION **
210 ADIN Y UC
220 E = UC-Y
230 U = S0+E*I
240 IF U < ULOW THEN U = ULOW
250 IF U > UHIGH THEN U = UHIGH
260 DAOUT U
270 I = U + S1*E
280 DISMISS
300 REM***COMUNICACION CON EL OPERADOR ***
310 PRINT "GANANCIA = ";K;"NUEVO VALOR"
320 INPUT K
330 PRINT "TIEMPO DE INTEGRACION = ";TI;
      "NUEVO VALOR = "
340 INPUT TI
350 PRINT "PERIODO DE MUESTREO = ";H;
      "NUEVO VALOR = "
360 INPUT H1
370 H = H1 :S0=K:S1 = -K*(TI-H)/TI
380 DISMISS
400 REM ** INICIALIZACION **
410 H = 1
420 S0 = 10
430 S1 = -10
440 RETURN

```

producirán variaciones en el período de muestreo, ya que los tiempos de ejecución de las diferentes líneas puede variar.

Sistemas operativos de tiempo real

Para problemas con sólo dos procesos, no es difícil escribir un programa monitor que controle la ejecución de los procesos. Un programa monitor típico para esto puede escribirse en menos de 100 líneas de código ensamblador.

Este monitor sencillo se puede extender a varios procesos. Sin embargo, es una tarea de mucha mayor envergadura construir un monitor capaz de tratar situaciones más complejas. Un monitor de este tipo, al que también se le llama *sistema operativo de tiempo real*, puede ocupar desde unos pocos hasta 20 kilobytes de código.

Los sistemas operativos de tiempo real permiten definir tareas o procesos en un lenguaje de alto nivel como el FORTRAN o el Pascal. También se pueden ejecutar procesos a intervalos regulares o con una relación dada respecto a otros procesos. Los procesos se pueden introducir, arrancar y retirar en línea. Se pueden introducir y modificar las prioridades entre las diferentes tareas. La introducción de los sistemas operativos de tiempo real fue una de las mayores innovaciones cuando se introdujeron los computadores de control de procesos a mitad de los años sesenta. Ejemplos de sistemas operativos de este tipo son TSX de IBM y RSX de Digital Equipment Corporation. También se han introducido procesos, o tareas, en lenguajes más sencillos como BASIC, tal y como se vio en el Listado 15.7.

Los sistemas operativos de tiempo real son programas grandes de propósito general, que se suelen escribir en ensamblador. Son difíciles de mantener y modificar. Hay necesidad de disponer de sistemas operativos de tiempo real que se puedan adaptar a las necesidades de aplicaciones específicas. Por tanto, se han desarrollado lenguajes de programación con entidades para programación en tiempo real. Lenguajes de este tipo son Pascal Concurrente y Modula 2. Se espera que el lenguaje Ada se convierta en una herramienta típica para la realización de sistemas de control con computador.

Paquetes DDC

Existen técnicas especiales para programar sistemas de control consistentes en lazos de control idénticos. El programa se suele estructurar de la siguiente forma.

Leer todas las entradas analógicas y almacenarlas en una tabla.

Convertir todas las señales a unidades de ingeniería y almacenar los resultados en una tabla.

Aplicar el algoritmo de control de forma secuencial a todos los valores de la tabla empleando los parámetros del regulador almacenados en una tabla de parámetros.

Realizar la conversión D-A de todas las variables almacenadas en la tabla de salida.

Los programas de este tipo se suelen llamar paquetes DDC. La elección del algoritmo de control suele estar limitada a diferentes algoritmos PID. Los paquetes son fáciles de usar porque toda la programación se reduce a introducir los datos adecuados en las tablas. A los programas de este tipo se les llama *mandados por tabla*.

15.8. Conclusiones

En este capítulo se ha expuesto la realización práctica de leyes de control mediante un computador. El problema clave es la realización de un sistema discreto. Se han tratado en detalle los principios para hacerlo. Es inmediato generar el programa a partir del algoritmo de control. Se ha mencionado la importancia del

prefiltrado para evitar el enmascaramiento. Se han expuesto sofisticados filtrados digitales no lineales para eliminar los datos erróneos. La organización de las instrucciones del programa influye de forma considerable sobre el retardo de cálculo. Se han expuesto las dificultades que surgen por la saturación de los actuadores y las formas de evitarlas. Esto conduce inmediatamente a una solución al cambio de modo y a la inicialización.

También se han tratado los problemas numéricos y las consecuencias de la longitud finita de palabra. Se ha visto que es muy beneficioso transformar las ecuaciones que describen la ley de control a una forma que numéricamente sea de buena condición. Se han expuesto temas de operación como el cambio de modo y la interacción operador-máquina. Existen muchas posibilidades nuevas en este área. Finalmente, se ha expuesto también la programación de algoritmos de control. Aunque la presentación es un poco corta, la información que se da debe ser suficiente para realizar algoritmos de control mediante mini- y microcomputadores empleando lenguajes de alto nivel.

15.9. Problemas

- 15.1. Considérese el control de un integrador doble con un período de muestreo de 1 s. Calcúlese el control de tiempo finito para el sistema obtenido empleando un filtro antienmascaramiento que tiene como función de transferencia

$$G(s) = \frac{1}{s^2 + 1.4s + 1}$$

Compárese con la estrategia de tiempo finito obtenida para un integrador doble puro empleando la simulación.

- 15.2. Escribase un programa para calcular el producto escalar de dos vectores

```
begin
  s:=0
  for i:=1 to n do
    s:=s + a[i] * b[i]
end
```

en cada uno de los siguientes casos.

- a) s: entero; a, b: vectores de componentes enteras.
 - b) s: entero doble precisión; a, b: vectores de componentes enteras.
 - c) s: real; a, b: vectores de componentes reales.
 - d) s: real doble precisión; a, b: vectores de componentes reales.
- Compárense los tiempos de cálculo y las precisiones. Inténtese hacerlo con computadores que tengan punto flotante por software, y también por hardware.

- 15.3. Considérese el Ejemplo 15.11. Discútanse las posibilidades de emplear dos lazos con períodos de muestreo diferentes para mejorar la precisión de los cálculos.

- 15.4. Escribase un programa para un regulador PI digital que contenga la compensación de saturación del término integral realizada como un observador con constante de tiempo T_o .

- 15.5. Escribase un programa para un regulador PI digital que contenga la compensación de la saturación del término integral realizada como un observador de tiempo finito.
- 15.6. Escribase un programa en el lenguaje de alto nivel que se prefiera para un regulador PID digital que contenga la compensación de la saturación del término integral realizada como un observador con constante de tiempo T_o . Determinese el número de operaciones requeridas para una iteración. Compílese el programa. Determinese cuántos accesos a memoria necesita. Evalúese la duración del programa. ¿Cómo se relacionan los tiempos de cálculo medidos con el número de operaciones y los tiempos de cálculo dados en el manual del computador?
- 15.7. Considérese el algoritmo de control dado en (15.1), donde x_m se considera como entrada. Supóngase que el estado, la variable de control y la salida del proceso tienen dimensiones nx , nu y ny y que las matrices son completas. Determinese el número de sumas, multiplicaciones y divisiones que se requieren en una iteración.
- 15.8. Considérese el algoritmo de control de (15.2). Escribase un programa en el lenguaje de alto nivel que se prefiera que ejecute el algoritmo de control. Compílese, determinese cuántas memoria ocupa, y su tiempo de ejecución. Inténtese encontrar una fórmula buena y sencilla para determinar el tiempo de ejecución.
- 15.9. Repítase el Problema 15.8 pero usando una subrutina para calcular el producto escalar. Expóngase cómo ha influido la reestructuración del programa sobre el tiempo de cálculo y las necesidades de memoria.
- 15.10. Considérese el algoritmo de control con rechazo de datos extraños dado por las ecuaciones (15.1) y (15.5). Estímese el número de cálculos necesarios para una iteración. (Sugerencia: Una multiplicación de matrices de una matriz $n \times p$ por otra $p \times r$ necesita $N = pnr$ operaciones, donde una operación equivale a una suma y una multiplicación. La solución de la ecuación

$$Ax = B$$

donde A es $n \times n$ y B es $n \times p$, requiere aproximadamente

$$N = \frac{1}{2}n^3 + \frac{1}{2}n^2p$$

operaciones, donde la mayor parte de los cálculos se dedican a la triangulación de la matriz A .)

- 15.11. El programa del listado que sigue corresponde a un filtro de Kalman que puede rechazar de forma selectiva errores de medida grandes. Estúdiese el algoritmo y expóngase cómo se reconfigura el filtro después de una falsa medida. Nótese que el programa está escrito para el caso en que los cálculos duran aproximadamente un período de muestreo.

```

begin
  e = y - C*x
  R = R2 + C*C'P*CT
  {Comprobar qué medidas se aceptan}
  na := 0
  for i := to ny do
    begin
      t = e[i] * e[i]/R[i,i]
      if t < test then

```

```

begin
  na := na + 1
  ya[na] := y[i]
  ia[na] := i
end
end
{El número de medidas aceptadas es na.
EL vector ia[1..na] contiene los índices
de las medidas aceptadas. El vector ya[1..na]
contiene las medidas aceptadas.}
for i := 1 to do na do
begin
  for j := 1 to na do
    Ra[i,j] := R[ia[i], ia[j]]
  for j := 1 to nx do
    Ca[i,j] := C[ia[i],j]
    ea[i] := e[ia[i]]
  end
  K := A * P * CaT * Inv(Ra)
  x := A * x + B * u + K * ea
  P := A * P * AT + R1 - K * Ca * P * AT
end

```

- 15.12. Considérese un sistema discreto caracterizado por la función de transferencia discreta

$$H(z) = \frac{1}{(z - a)^n}$$

Calcúlese la sensibilidad de los polos respecto a los parámetros empleando la ecuación (15.2) en cada uno de los casos.

- a) El filtro está en forma asociada.
 b) El filtro está en la forma canónica de Jordan.
 15.13. Hágase un diagrama de flujo similar al de la Figura 15.16 para un sistema con lazos que tienen períodos de muestreo de 1, 2, 5 y 60 s.
 15.14. Considérese un sistema con la función de transferencia

$$H(z) = \frac{1}{z^n + a_1 z^{n-1} + \dots + a_n}$$

Supóngase que se realiza el sistema con aritmética de punto fijo. Describáse el redondeo como un redondeo normal de enteros. Demuéstrese que la condición para tener un error estacionario k sin entradas es

$$k + \sum_{i=1}^n Q[a_i k] = 0$$

Además, demuéstrese que la condición para que se produzca una oscilación de ciclo límite con un período igual a dos períodos de muestreo es

$$k + \sum_{i=1}^n (-1)^i Q[a_i k] = 0$$

- 15.15. En la Figura 15.17a) puede verse un regulador PI con convertidores A-D y D-A. En la Figura 15.17b) se presenta el modelo lineal obtenido con una realización del regulador PI. Demuéstrese que el error equivalente reducido a la entrada del convertidor D-A se puede representar como

$$\epsilon = e_{ad} + \frac{1}{a+b}(e_a + e_b)$$

En la Figura 15.17c) se puede ver el modelo lineal obtenido con otra realización del regulador PI. Dar el error correspondiente para esta representación y analicense las diferencias existentes entre las dos realizaciones de los modelos.

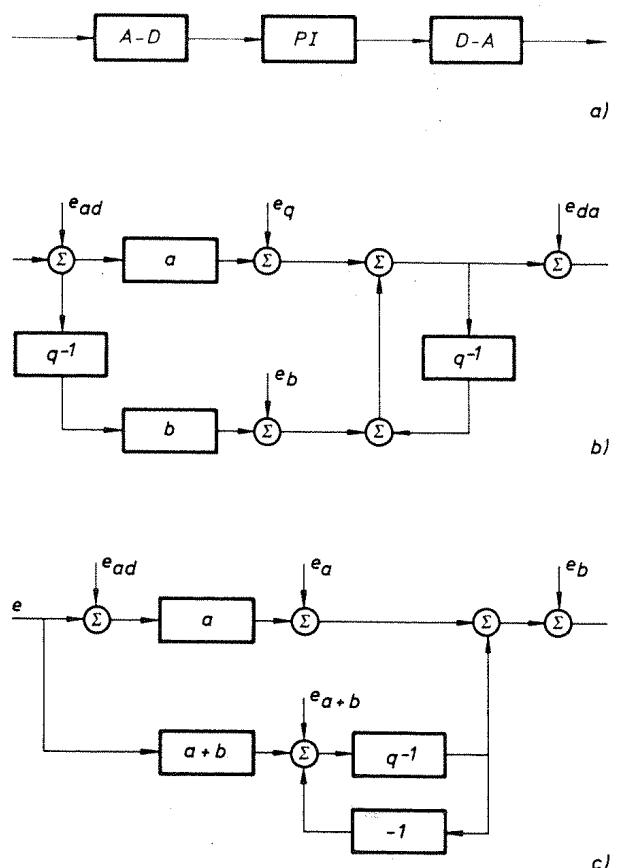


Figura 15.17

- 15.16. Considérese el siguiente algoritmo de un regulador PI:

```

    Adin uc y
    e := uc - y
    v := k*e + i
    u := max(min(512,v),0)
    Daout u
    i := u + k*h*e/ti
  
```

Supóngase que los convertidores A-D y D-A tienen una resolución de 8 bits y que todos los cálculos se realizan con números enteros. ¿Cuál es la longitud de palabra necesaria para representar la variable i evitando el desbordamiento? Usese $k=50$ y a) $h=1$, $t_i=300$ o b) $h=0,01$, $t_i=1.500$. Expóngase cómo influye el periodo de muestreo sobre el resultado.

- 15.17. Más abajo se listan tres algoritmos diferentes de un regulador PI. Empléese el modelo lineal del redondeo para analizar la sensibilidad de los algoritmos a la cuantificación en los convertidores A-D y D-A y al redondeo en las multiplicaciones. Supóngase que los cálculos se hacen en coma fija. Expóngase también la longitud de palabra necesaria para los algoritmos.

Algoritmo 1

```

    e := uc - y
    u := k*(e + h*i/ti)
    i := i + e*h
  
```

Algoritmo 2

```

    e := uc - y
    u := k*(e+i)
    i := i + e*h/ti
  
```

Algoritmo 3

```

    e := uc - y
    u := i + k*e
    i := i + k*h*e/ti
  
```

15.10. Referencias

El diseño de filtros se trata en los libros clásicos sobre circuitos. Los libros KUO, F. F. (1962): *Network Analysis and Synthesis*. New York: John Wiley. WILLIAMS, A. B. (1981): *Electronic Filter Design Handbook*. New York: McGraw-Hill.

son buenas fuentes. También se pueden encontrar consejos prácticos útiles en los manuales publicados por los fabricantes de amplificadores operacionales. También son útiles estos manuales para obtener información acerca de los convertidores A-D y D-A. Asegúrese de disponer de una nueva versión de cualquier manual que use porque la tecnología cambia rápidamente.

En las revistas comerciales sobre la industria de procesos se tratan los problemas relacionados con la saturación del término integral de los reguladores PID. Creemos que el planteamiento general expuesto en la Sección 15.4 es nuevo.

Las posibilidades de detección y rechazo de datos extraños se tratan en profundidad en WILLSKY, A. (1976): «A Survey of Design Methods for Failure Detection in Dynamic Systems,» *Automatica*, 12, 601-11.

que contiene también muchas referencias.

Un texto completo sobre los efectos de la cuantificación y el redondeo en los sistemas digitales de control es

MORONEY, P. (1983): «Issues in the Implementation of Digital Feedback Compensators,» Cambridge, Mass.: MIT Press.

que contiene muchas referencias. Los siguientes trabajos son clásicos en el área.

KNOWLES, J. B. y R. EDWARDS (1965): «Effect of a Finite-Word-Length Computer in a Sampled-Data Feedback System,» *Proc. IEE*, 112, 1197-1207.

CURRY, E. E. (1967): «The Analysis of Round-Off and Truncation Errors in a Hybrid Control System,» *IEEE Trans. Autom. Control*, AC-13, 601-4.

BERTRAM, J. E. (1958): «The Effect of Quantization in Sampled-Feedback Systems,» *Trans. Amer. Inst. Elec. Engrs.*, 77, 177-82.

SLAUGHTER, J. B. (1964): «Quantization Errors Digital Control Systems,» *IEEE Trans. Autom. Control*, AC-9, 70-74.

RINK, R. E. y H. Y. CHONG (1979): «Performance of State Regulator Systems with Floating-Point Computation,» *IEEE Trans. Autom. Control*, AC-24, 411-21.

También se tratan los problemas asociados con la cuantificación, el redondeo y el desbordamiento en la literatura sobre proceso de señales. Pueden encontrarse visiones de conjunto en

OPPENHEIM, A. V. y R. W. SCHAFER (1975): *Digital Signal Processing*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.

RABINER, L. R. y B. GOLD (1975): *Theory and Application of Digital Signal Processing*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.

Temas especializados se exponen en

JACKSON, L. B. (1970): «Roundoff Noise Analysis for Fixed-Point Digital Filters Realized in Cascade or Parallel Form,» *IEEE Trans. Audio & Electroacoustics*, AU-18, 107-22.

JACKSON, L. B. (1970): «On the Interaction of Roundoff Noise and Dynamic Range in Digital Filters,» *Bell Syst. Tech. Journal*, 49, 159-84.

JACKSON, L. B. (1979): «Limit Cycles in State-Space Structures for Digital Filters,» *IEEE Trans. Circuits & Systems*, CAS-26, 67-68.

PARKER, S. R. y S. F. HESS (1971): «Limit Cycle Oscillations in Digital Filters,» *IEEE Trans. Circuit Theory*, CT-18, 687-97.

WILLSON, JR., A. N. (1972): «Some Effects of Quantization and Adder Overflow on the Forced Response of Digital Filters,» *Bell Syst. Tech. Journal*, 51, 863-87.

BUTTNER, M. (1977): «Elimination of Limit Cycles in Digital Filters with Very Low Increase in Quantization Noise,» *IEEE Trans. Circuits & Systems*, CAS-24, 300-304.

WILLSON, JR., A. N. (1972): «Limit Cycles Due to Adder Overflow in Digital Filters,» *IEEE Trans. Circuit Theory*, CT-19, 342-46.



Existen muchos textos típicos sobre análisis numérico:

BJÖRK, G., Å. DAHLQVIST, y N. ANDERSSON (1974): *Numerical Methods*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.

es una buena fuente.

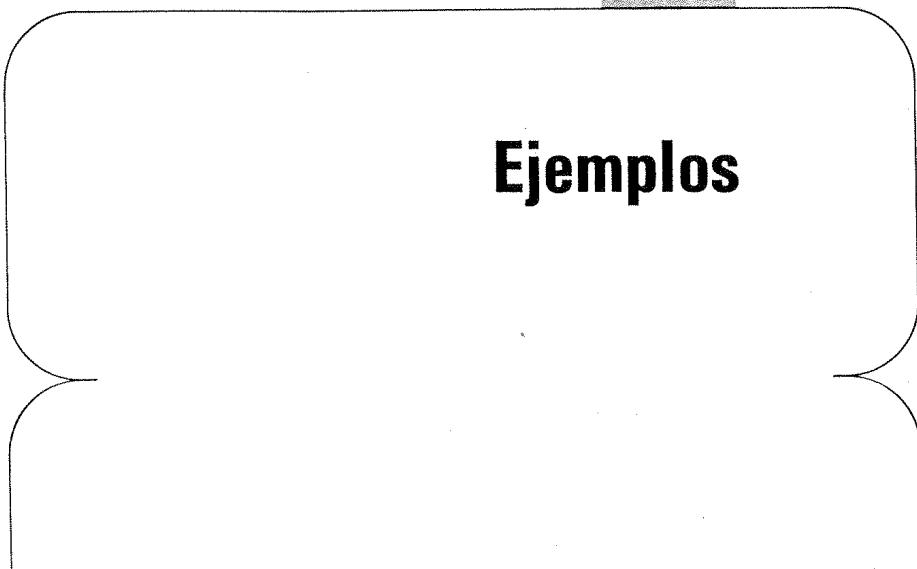
La programación concurrente se trata en

BRINCH-HANSEN, P. (1973): *Operating System Principles*. Englewood Cliffs, N. J.: Prentice-Hall, Inc.

BARNES, J. G. P. (1982): *Programming in ADA*. New York: Addison-Wesley.

También se da información muy útil en el material de los vendedores de sistemas de control con computador.

Ejemplos



En este apéndice se presentan los «procesos típicos» utilizados como ejemplos en este libro.

Ejemplo A.1 – Integrador doble

El integrador doble se utiliza en todo el libro como ejemplo principal para ilustrar las teorías presentadas. El proceso está descrito por la ecuación diferencial

$$\frac{d^2y}{dt^2} = u \quad (\text{A.1})$$

La función de transferencia es $G(s) = 1/s^2$. Se definen y e \dot{y} como los estados x_1 y x_2 del sistema, respectivamente. Entonces la representación de estados es

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}x + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u \\ y &= [1 \ 0]x\end{aligned} \quad (\text{A.2})$$

Muestreando (A.2) con un mantenedor de orden cero y período de muestreo h se obtiene el sistema discreto (véase el Ejemplo 3.1)

$$\begin{aligned}x(kh + h) &= \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}x(kh) + \begin{bmatrix} h^2/2 \\ h \end{bmatrix}u(kh) \\ y(kh) &= [1 \ 0]x(kh)\end{aligned} \quad (\text{A.3})$$

El operador de transferencia discreta de (A.3) es

$$H(q) = \frac{h^2(q+1)}{2(q-1)^2} \quad (\text{A.4})$$

Existen varios procesos físicos que pueden describirse como integradores dobles. Uno de ellos es la posición de un satélite, que puede describirse por la ecuación

$$J \frac{d^2\theta}{dt^2} = M_c + M_d$$

donde θ es el ángulo de la posición, M_c es el par de control, M_d es el par de perturbación y J es el momento de inercia.

Otro ejemplo que puede describirse con un integrador doble es la bola rodante sobre el carril inclinado (véase la Fig. A.1). La ecuación de la bola y el carril puede describirse por

$$\begin{aligned}J \frac{d^2\theta}{dt^2} &= mgr \sin \varphi \approx mgr \varphi \\ x &= r\theta\end{aligned}$$

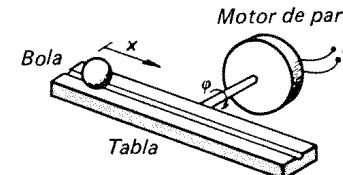


Figura A.1 Ilustración esquemática de la bola y la tabla.

o

$$\frac{d^2x}{dt^2} = mgr^2\varphi/J$$

donde θ es el ángulo de la bola, g es la aceleración normal, x es la posición de la bola y φ es el ángulo de inclinación del carril.

Ejemplo A.2 – Motor

Un motor CC puede describirse por un modelo de segundo orden con un integrador y una constante de tiempo (véase la Fig. A.2). La entrada es la tensión de alimentación del motor y la salida es la posición del eje. La constante de tiempo se debe a las partes mecánicas del sistema, mientras que la dinámica de las partes eléctricas es despreciable. Un modelo normalizado del proceso es

$$Y(s) = \frac{1}{s(s+1)}U(s)$$

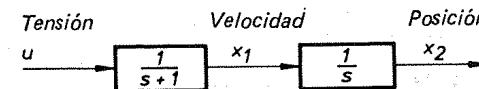


Figura A.2 Modelo normalizado de un motor CC.

Introducimos como estados la velocidad y la posición del eje (véase la Fig. A.2). Entonces el modelo en el espacio de estados es

$$\begin{aligned}\dot{x} &= \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}x + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u \\ y &= [0 \ 1]x(t)\end{aligned}\quad (\text{A.5})$$

Muestreando (A.5) con un mantenedor de orden cero se obtiene el modelo discreto

$$\begin{aligned}x(kh+h) &= \begin{bmatrix} e^{-h} & 0 \\ 1 - e^{-h} & 1 \end{bmatrix}x(kh) + \begin{bmatrix} 1 - e^{-h} \\ h - 1 + e^{-h} \end{bmatrix}u(kh) \\ y(kh) &= [0 \ 1]x(kh)\end{aligned}\quad (\text{A.6})$$

(véase el Ejemplo 3.2). También puede describirse con el modelo de (A.5) un motor CC controlado por corriente con la velocidad del eje como salida. Otro ejemplo que puede caracterizarse como un integrador y un polo simple es un barco. Consideremos como entrada al sistema el ángulo del timón y como salida la dirección de la proa. Entonces el movimiento del barco puede describirse por la función de transferencia

$$G(s) = \frac{K}{s(1+Ts)}$$

donde la constante de tiempo puede ser positiva o negativa dependiendo del tipo de barco. Por ejemplo, los grandes buques cisterna son inestables.

Ejemplo A.3 – Oscilador armónico

Consideremos un péndulo (véase la Fig. A.3). La entrada es la aceleración del pivote y la salida es el ángulo y . Entonces, el sistema se describe por la ecuación no lineal normalizada

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\sin x_1 + u \cos x_1 \\ y &= x_1\end{aligned}$$

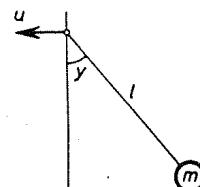


Figura A.3 Péndulo.

donde x_1 es el ángulo y x_2 es la velocidad angular. Linealizando alrededor de $u=x_1=0$ se obtiene

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}x + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u \\ y &= [1 \ 0]x\end{aligned}\quad (\text{A.7})$$

La función de transferencia de (A.7) es

$$G(s) = \frac{1}{s^2 + 1}$$

Esta función de transferencia puede generalizarse a

$$G(s) = \frac{\omega^2}{s^2 + \omega^2}$$

Una representación en el espacio de estados para esta función de transferencia es

$$\begin{aligned}\dot{x} &= \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix}x + \begin{bmatrix} 0 \\ \omega \end{bmatrix}u \\ y &= [1 \ 0]x\end{aligned}\quad (\text{A.8})$$

Muestreando (A.8) con un mantenedor de orden cero se obtiene el sistema discreto

$$\begin{aligned}x(kh+h) &= \begin{bmatrix} \cos \omega h & \sin \omega h \\ -\sin \omega h & \cos \omega h \end{bmatrix}x(kh) + \begin{bmatrix} 1 - \cos \omega h \\ \sin \omega h \end{bmatrix}u(kh) \\ y(kh) &= [1 \ 0]x(kh)\end{aligned}$$

También puede modelarse por (A.8) una grúa aérea.

Ejemplo A.4 – Proceso con un retardo de tiempo

Algunos procesos industriales pueden modelarse como sistemas dinámicos de primer orden con un retardo de tiempo. Un ejemplo es la máquina de papel (véase la Fig. A.4). La

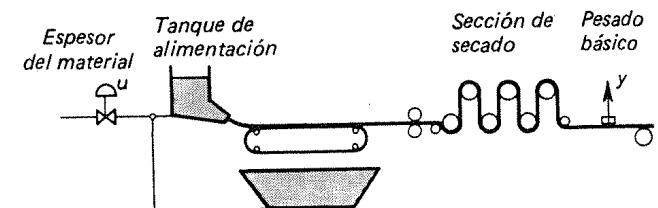


Figura A.4 Dibujo esquemático de una máquina de papel.

entrada es el espesor del material, es decir, la cantidad de pulpa. La salida es el peso básico, es decir, el espesor del papel. Las ecuaciones que describen el sistema pueden normalizarse a la función de transferencia

$$G(s) = \frac{1}{s + 1} e^{-\tau s} \quad (\text{A.9})$$

Otro proceso físico que puede describirse por (A.9) es un sistema de mezclas con largas conducciones de tuberías. El Ejemplo 3.6 da el muestreo de (A.9) con mantenedor de orden cero.

Ejemplo A.5 – Un modelo de inventario

Un inventario es un ejemplo típico que puede describirse normalmente como un sistema discreto. Ordenes y entregas se obtienen a intervalos regulares vinculados al calendario; por ejemplo, cada día o semana.

Sea $y(k)$ el inventario en el instante k antes de comenzar una transacción; $u(k)$ son las entregas del inventario que se ordenan en el instante k . Se supone que existe un retardo de un período desde que se efectúa la orden hasta que se actualiza la mercancía dentro del inventario. Finalmente, las entregas del inventario son $v(k)$. Introducimos las variables de estado $x_1(k) = y(k)$ y $x_2(k) = u(k-1)$. El inventario puede describirse por las siguientes ecuaciones de estado discretas:

$$\begin{aligned}x_1(k+1) &= x_1(k) + x_2(k) - v(k) \\x_2(k+1) &= u(k)\end{aligned}$$

$$\begin{aligned}x(k+1) &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u(k) + \begin{bmatrix} -1 \\ 0 \end{bmatrix}v(k) \\y(k) &= [1 \quad 0]x(k)\end{aligned}\tag{A.10}$$

La relación entrada-salida viene dada por

$$y(k) - y(k-1) = u(k-2) - v(k)\tag{A.11}$$

Funciones matriciales

En relación con los sistemas muestrados, interesan las funciones como $\exp A$ y $\ln A$, donde A es una matriz. La exponencial y el algoritmo de una matriz son *funciones matriciales*. Este apéndice da algunas propiedades de las funciones matriciales y discute algunos métodos para calcularlas.

El Teorema B.1 da una propiedad útil de una matriz cuadrada.

Teorema B.1 – El teorema de Cayley-Hamilton. Sea

$$a(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n = 0$$

la ecuación característica de la matriz cuadrada A . Entonces A satisface la siguiente ecuación

$$a(A) = A^n + a_1A^{n-1} + \cdots + a_nI = 0$$

Esto es, la matriz satisface su propia ecuación característica. □

Sea A una matriz cuadrada $n \times n$ y $f(\lambda)$ una función escalar del argumento escalar λ . Deseamos extender la función $f(\lambda)$ a una función con argumento matricial, es decir, $f(A)$. Si $f(\lambda)$ es un polinomio

$$f(\lambda) = \alpha_0\lambda^n + \alpha_1\lambda^{n-1} + \cdots + \alpha_m$$

entonces la función matricial $f(A)$ se define como

$$f(A) = \alpha_0 A^m + \alpha_1 A^{m-1} + \cdots + \alpha_m I$$

Los valores propios de $f(A)$ pueden calcularse utilizando el siguiente teorema.

Teorema B.2 Si $f(A)$ es un polinomio en A y e_i son los vectores propios de A asociados a los valores propios λ_i , entonces

$$f(A)e_i = f(\lambda_i)e_i$$

por tanto, $f(\lambda_i)$ es un valor propio de $f(A)$ y e_i es el correspondiente vector propio. \square

Además, si $f(\lambda)$ puede definirse por la serie de potencias

$$f(\lambda) = \sum_{i=0}^{\infty} c_i \lambda^i$$

la cual se supone que converge para $|\lambda| < R$, entonces la función matricial

$$f(A) = \sum_{i=0}^{\infty} c_i A^i$$

es convergente para todos los valores propios de A , λ_i , satisfaciendo $|\lambda_i| < R$.

Utilizando el teorema de Cayley-Hamilton puede demostrarse que para cada función f hay un polinomio p de grado menor que n tal que

$$f(A) = p(A) = \alpha_0 A^{n-1} + \alpha_1 A^{n-2} + \cdots + \alpha_{n-1} I \quad (\text{B.1})$$

Del Teorema B.2 obtenemos

$$f(\lambda_i) = p(\lambda_i), \quad i = 1, \dots, n \quad (\text{B.2(a)})$$

Si los valores propios son diferentes, entonces estas condiciones son suficientes para determinar α_i , $i = 0, \dots, n-1$. Si hay valores propios múltiples de multiplicidad m , entonces se mantienen las condiciones adicionales

$$\begin{aligned} f^{(1)}(\lambda_i) &= p^{(1)}(\lambda_i) \\ &\vdots \\ f^{(m-1)}(\lambda_i) &= p^{(m-1)}(\lambda_i) \end{aligned} \quad (\text{B.2(b)})$$

donde $f^{(i)}$ es la i -ésima derivada con respecto a λ .

Utilizando (B.1) y las condiciones de (B.2) es posible calcular funciones de matrices. Para sistemas de orden bajo, éste es un método muy adecuado para realizar los cálculos a mano.

Ejemplo B.1

Sea

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

determinar

$$\exp Ah = \alpha_0 A + \alpha_1 I$$

Ah tiene los valores propios $\pm ih$; el sistema de ecuaciones

$$e^{ih} = \alpha_0 ih + \alpha_1$$

$$e^{-ih} = -\alpha_0 ih + \alpha_1$$

se mantiene, dando

$$\alpha_0 = \frac{1}{2i}(e^{ih} - e^{-ih}) = \sin h$$

$$\alpha_1 = \frac{1}{2}(e^{ih} + e^{-ih}) = \cos h$$

Finalmente,

$$\exp Ah = \sin h \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \cos h \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos h & \sin h \\ -\sin h & \cos h \end{bmatrix}$$

Ejemplo B.2

Sea

$$\Phi = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}$$

calcular $\ln \Phi$. Los valores propios están dados por $(\lambda - 1)^2 = 0$; es decir, existen valores propios múltiples. Entonces la matriz logarítmica puede escribirse como

$$\ln \Phi = \alpha_0 \Phi + \alpha_1 I$$

donde α_0 y α_1 están dados por

$$\begin{cases} \frac{\ln 1}{\partial \lambda} = \alpha_0 + \alpha_1 \\ \left. \frac{\partial}{\partial \lambda} (\ln \lambda) \right|_{\lambda=1} = \alpha_0 \end{cases} \quad \text{or} \quad \begin{cases} 0 = \alpha_0 + \alpha_1 \\ 1 = \alpha_0 \end{cases}$$

Finalmente,

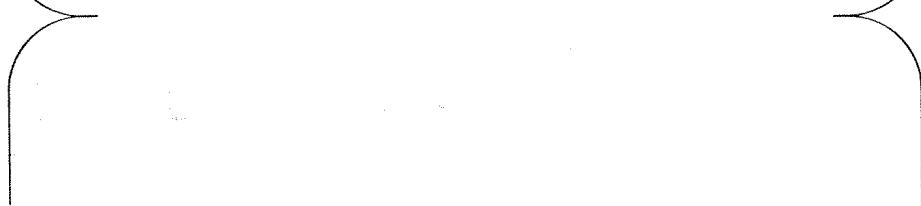
$$\ln \Phi = \begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & h \\ 0 & 0 \end{bmatrix}$$

Nota. En lugar de comenzar con el polinomio característico, es posible utilizar el polinomio minimal de la matriz. El grado de las series de (B.1) será entonces el grado del polinomio minimal menos uno. En general esto no reduce el tiempo de cálculo porque debe calcularse el polinomio mínimo, o alternativamente la forma de Jordan.

Otras propiedades de las funciones matriciales pueden encontrarse en

- GANTMACHER, F. R. (1960): *The Theory of Matrices*, vols. I y II. New York: Chelsea.
BELLMAN, R. (1970): *Introduction to Matrix Analysis*. New York: McGraw-Hill.

Simnon: Un programa de simulación interactivo



Un buen programa de simulación es una herramienta indispensable para investigar sobre sistemas controlados por computador. Existen muchos programas disponibles. Este apéndice describe el programa Simnon, que se ha utilizado en este libro.

Simnon es un programa interactivo para simulación de sistemas descritos por ecuaciones diferenciales ordinarias no lineales y ecuaciones en diferencias. El programa se desarrolló en el Departamento de Control Automático del Instituto de Tecnología de Lund. El trabajo ha sido financiado por el Instituto Sueco de Matemáticas Aplicadas (ITM) y por el Consejo Sueco de Desarrollos Técnicos (STU). Una descripción más detallada del programa se puede ver en Elmquist (1975) y Åström (1982). Este apéndice se basa en el trabajo de Elmquist (1977).

Introducción

El Simnon es un programa de simulación interactivo. Se ha utilizado ampliamente desde 1974 en investigación y docencia, y para resolver problemas de control industrial y de simulación. La interactividad hace posible realizar estudios de simulación en un tiempo muy corto. El lenguaje simple de modelos y la amplia posibilidad de detección de errores hacen de Simnon un programa adecuado para la enseñanza.

El usuario interactúa con el programa utilizando órdenes. Por ejemplo, hay órdenes para cambiar los parámetros del modelo, realizar la simulación, representar

los resultados de la simulación y modificar el modelo. El usuario puede construir un conjunto de órdenes personales de alto nivel por medio de macro-órdenes.

El modelo se puede describir en un lenguaje especial para modelos o en FORTRAN. El lenguaje de modelos es simple y fácil de aprender y permite detectar un gran número de errores. En el programa se incluye el compilador que trabaja en paralelo con el editor. Esto le permite al usuario corregir inmediatamente las líneas con errores.

El programa Simnon permite descomponer el sistema en subsistemas, que pueden describirse separadamente. Los subsistemas se conectan utilizando sus entradas y salidas. Esto permite disponer de una biblioteca de modelos de procesos. Además, es fácil cambiar la estructura del proceso o del controlador. Cuando se simulan sistemas controlados con computador, es normal describir los procesos físicos por ecuaciones diferenciales ordinarias y el computador y sus algoritmos de control por ecuaciones en diferencias. Por esta razón, el programa Simnon permite la descripción de subsistemas continuos y discretos.

Estructura del modelo

Hay tres tipos de sistemas en el programa Simnon: continuos, discretos y de interconexión.

Un subsistema continuo se define matemáticamente por

$$\begin{aligned}\dot{x}(t) &= f(x(t), t, u(t), p) \\ y(t) &= g(x(t), t, u(t), p) \\ x(t_0) &= x_0\end{aligned}$$

Se emplea la siguiente notación:

- t : tiempo (variable independiente).
- x : vector de variables de estado (variables dependientes).
- y : vector de variables de salida.
- p : vector paramétrico.
- t_0 : tiempo inicial para la simulación.
- x_0 : valores iniciales de los estados.

Un subsistema discreto se define matemáticamente por

$$\begin{aligned}x(t_{k+1}) &= f(x(t_k), t_k, u(t_k), p) \\ y(t_k) &= g(x(t_k), t_k, u(t_k), p) \\ x(t_0) &= x_0\end{aligned}$$

El k -ésimo instante de muestreo se designa por t_k . Los instantes de muestreo no tienen que ser equidistantes ni tampoco iguales en todos los subsistemas discretos.

Las ecuaciones de los subsistemas discretos no definen la salida entre los instantes de muestreo ni en instantes de muestreo anteriores al primero. Para permitir la conexión de sistemas continuos y discretos con diferentes instantes de muestreo debe extenderse la definición de las salidas. Esto se hace introduciendo circuitos mantenedores de orden cero en las salidas. Los estados también tienen circuitos mantenedores de orden cero puesto que también pueden utilizarse como salidas.

La conexión de subsistemas se hace en un *sistema de interconexión*. El sistema de interconexión define el modo en que se interconectan las entradas y salidas de los diferentes subsistemas. Este sistema puede representarse como un sistema estático. Puede ser variable en el tiempo, pero no tiene dinámica.

Si el sistema no contiene ningún lazo algebraico, entonces las ecuaciones del sistema pueden resolverse secuencialmente ordenándolas adecuadamente. Esto se hace de forma automática.

Los instantes de muestreo para cada subsistema discreto se especifican con una variable especial en la descripción del sistema. Esta variable se actualiza en cada instante de muestreo con el instante del siguiente muestreo. De este modo, después de la activación de un subsistema discreto, se conoce cuándo debe tomarse la siguiente muestra. Las ecuaciones diferenciales, que están parametrizadas por los estados discretos y las salidas, pueden resolverse sobre el intervalo de muestreo con una rutina de integración usual. Actualmente, hay un algoritmo predictor-corrector de Hamming, un algoritmo de Runge-Kutta y una rutina para ecuaciones diferenciales rígidas.

Lenguaje del modelo

El programa Simonon incluye un lenguaje especial para describir los subsistemas y sus conexiones. Las ecuaciones se introducen utilizando sentencias como las de Algol-60. Se ha demostrado que la construcción *if-then-else* es muy útil. El lenguaje del modelo es muy simple, lo cual ha permitido implantar una amplia comprobación de errores. Si se necesitan modelos complicados, es posible utilizar subrutinas FORTRAN o Pascal para describir el modelo.

En lo que sigue, las palabras claves del programa Simonon se escribirán en letras mayúsculas. En la descripción del modelo se pueden utilizar letras mayúsculas o minúsculas. La descripción de las instrucciones del lenguaje de modelos se obtiene a través de la orden HELP SIMNON.

Los sistemas continuos y discretos tienen la siguiente estructura:

```
CONTINUOUS SYSTEM <nombre>
declaraciones
asignaciones
END
```

```
DISCRETE SYSTEM <nombre>
declaraciones
asignaciones
END
```

El encabezamiento del sistema define su tipo y su nombre. Hay declaraciones para especificar tipos de variables. Los tipos son INPUT, OUTPUT, TIME, STATE, DER,

NEW y TSAMP. El tipo DER se utiliza en sistemas continuos para asociar a una variable la derivada de una variable de estado. La declaración NEW se utiliza de una forma equivalente en sistemas discretos para actualizar una variable de estado. En cada sistema discreto se utiliza una variable del tipo TSAMP para especificar el próximo instante de muestreo. La descripción del sistema también puede contener parámetros y variables auxiliares que no se declaran.

Los parámetros se asignan en una sentencia de la siguiente forma:

```
<parámetro> : <número>
```

Los valores iniciales de las variables de estado se pueden asignar del mismo modo. Las variables se asignan de la forma

```
<variable> = <expresión>
```

Las variables reciben valores en la sección de asignaciones. Por ejemplo, las derivadas, la actualización de estados y las salidas reciben valores. El compilador clasifica automáticamente las asignaciones de un modo adecuado para realizar los cálculos.

Los sistemas de interconexión tienen la siguiente estructura:

```
CONNECTING SYSTEM <nombre>
declaraciones
sección de conexiones
END
```

La sección de conexiones contiene sentencias para las variables INPUT de los subsistemas. Se puede utilizar el mismo identificador para variables en diferentes subsistemas. Por tanto, se utiliza la siguiente notación en un sistema de interconexión para referenciar las variables en los subsistemas.

```
<variable>[<subsistema>]
```

La parte del lado derecho de la asignación puede contener variables STATE o OUTPUT, las cuales se refieren del mismo modo.

Funciones y sistemas típicos

Se dispone de varias funciones típicas en el paquete. Estas funciones pueden utilizarse en asignaciones en subsistemas o en sistemas de interconexión. Se dispone de las siguientes funciones típicas:

ABS(x): Valor absoluto de x.

ATAN(x): Arco tangente de x. El resultado se da en radianes en el intervalo $(-\pi/2, \pi/2)$.

- ATAN2(x,y): Arco tangente de x/y . El resultado se da en radianes en $(-\pi, \pi)$.
- COS(x): Coseno de x , x en radianes.
- EXP(x): Función exponencial de x .
- INT(x): Parte entera de x .
- LN(x): Logaritmo natural de x , $x > 0$.
- LOG(x): Logaritmo en base 10 de x , $x > 0$.
- MAX(x,y): Valor máximo de x e y .
- MIN(x,y): Valor mínimo de x e y .
- MOD(x,y): Resto de dividir x entre y . Observar, sin embargo, que $\text{mod}(x, x) = x$.
- SIGN(x): Signo de x ; 1 si $x > 0$, 0 si $x = 0$ y -1 si $x < 0$.
- SIN(x): Seno de x , x en radianes.
- SQRT(x): Raíz cuadrada de x , $x \geq 0$.
- TAN(x): Tangente de x , x en radianes.

También se incluyen en el paquete algunos sistemas típicos de interés común. Estos son:

- DELAY (continuo): Retardo de tiempo puro.
- FUNC (continuo): Función de búsqueda en tablas.
- IFILE (discreto): Lee un valor de un fichero.
- LOGGER (discreto): Muestra y guarda variables Simnon en un fichero.
- NOISE1 (discreto): Generador de ruido aleatorio, gaussiano o rectangular.
- OPTA (discreto): Sistema para optimización.

Manipulación del modelo

La descripción del modelo se almacena en ficheros o almacenamientos de masas. La manipulación del modelo se realiza utilizando órdenes con argumentos. Las órdenes se introducen normalmente por un terminal, pero también pueden leerse de ficheros. Los resultados de la simulación se dibujan en un visualizador gráfico.

Más adelante se dará una descripción de las órdenes. Para una descripción completa conviene ver Elmquist (1975).

El compilador trabaja en paralelo con el editor. Cuando el compilador descubre un error da un mensaje de error y el usuario tiene la posibilidad de corregir la línea errónea mediante órdenes de edición. Las órdenes del editor se darán más adelante.

Facilidades de interacción

El núcleo del programa interactivo es un conjunto de subrutinas denominado Intrac. Estas rutinas manejan la interacción entre el usuario y el paquete de

aplicaciones. Por ejemplo, Intrac maneja la decodificación de órdenes. La estructura de cada orden es flexible. En algunos casos, pueden omitirse los argumentos. En este caso, se utilizan los valores por defecto.

Una situación que se presenta con frecuencia cuando se ejecuta el programa es que se repita varias veces la misma secuencia de órdenes. En este caso, el usuario puede definir un macro que contenga las órdenes. Entonces, este macro se utiliza como una nueva orden, admitiendo diferentes valores para los argumentos. Los marcros se definen utilizando órdenes para Intrac. En el macro se pueden introducir lazos de repetición e instrucciones de control.

Funciones de ayuda

El paquete de simulación incluye ayudas, que pueden utilizarse durante una sesión de simulación. La orden HELP proporciona un menú de las órdenes disponibles en Simnon. Se obtiene información adicional sobre cada orden escribiendo HELP <orden>. También es posible obtener información sobre el editor, Intrac y los subsistemas disponibles escribiendo HELP seguido de EDIT, INTRAC o SYSTEMS, respectivamente.

Ejemplo

Se ilustrarán con un ejemplo algunas características del programa Simnon. Consideraremos el sistema de tanque de la Figura C.1.

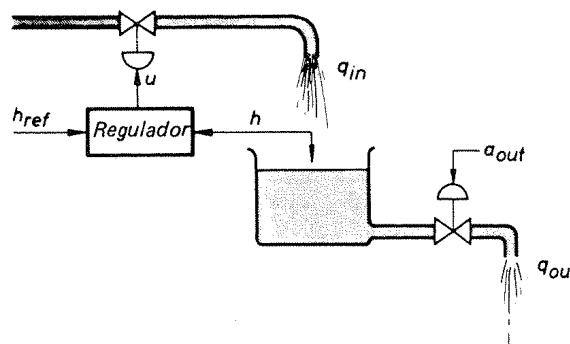


Figura C.1 Sistema de tanque con válvulas y un controlador.

La válvula de entrada está controlada por un regulador para mantener el nivel constante. La válvula de la salida se manipula exteriormente. Cuando se estudia este sistema la válvula de la salida se considera como una perturbación.

El estudio de la simulación se realiza para determinar el regulador adecuado y encontrar sus parámetros.

Las descripciones del modelo se muestran en el Listado C.1. El tanque y las válvulas se describen en el sistema denominado *tanque*. El sistema tiene dos entradas: la señal *v* a la válvula de entrada y el área de la válvula de salida *aout*. El nivel *h* es el estado que también se considera como salida. El tanque está controlado por un computador, que contiene un regulador PI. El subsistema que describe el regulador se denomina *dpi*.

LISTADO C.1 Los subsistemas *tank*, *dpi* y *regtank*.

```

CONTINUOUS SYSTEM tank
INPUT v aout
STATE h
DER dh
valve=IF v<0 THEN 0 ELSE IF v> 1 THEN 1 ELSE v
qin=qmax*valve
quot=aout*SQRT(2*g*MAX(h,0))
dh=(qin-qout)/area
qmax:1
g:9.81
area:10
END

DISCRETE SYSTEM dpi
INPUT y yref
OUTPUT u
STATE inte
NEW ninte
TIME t
TSAMP ts
e=yref-y
i1=inte+k*e*h/ti
u=k*e+i1
ninte=i1
ts=t+h
k:1
ti:1E10
h:1
END

CONNECTING SYSTEM regtank
TIME t
aout[tank]=IF t<100 THEN a1 ELSE a2
a1:0.01
a2:0.05
yref[dpi]=href
href:2
y[dpi]=h[tank]
v[tank]=u[dpi]
END

```

La interconexión de tanque y dpi se hace en *regtank*. La perturbación en el desagüe también se da en *regtank*. Para tiempo = 100, el área del desagüe se incrementa de 0,01 a 0,05. El tanque está inicialmente vacío. A continuación se escribe el programa de interacción con Simnon. Los comentarios están precedidos de comillas.

```

>SYST tank dpi regtank " Compilar modelo
>TURN S125 ON " Elegir escalas del dibujo.
>STORE h[tank] href qin " Preparación para almacenar los resultados.
>" 
>"Probar un regulador proporcional con k=1.
>" 
>PAR h:5 " Definir intervalo de muestreo.
>SIMU 0 200 " Simular 200 seg.
>ASHOW h href qin-MARK " Dibujar resultados.
>HCOPY " Realizar una copia (véase la Fig. C.2).
>" 
>" Eliminar el error estático con un regulador PI.
>" 
>PAR ti:25 " Cambiar el tiempo de integración.
>SIMU " Simular.
>ASHOW h href qin-MARK.

```

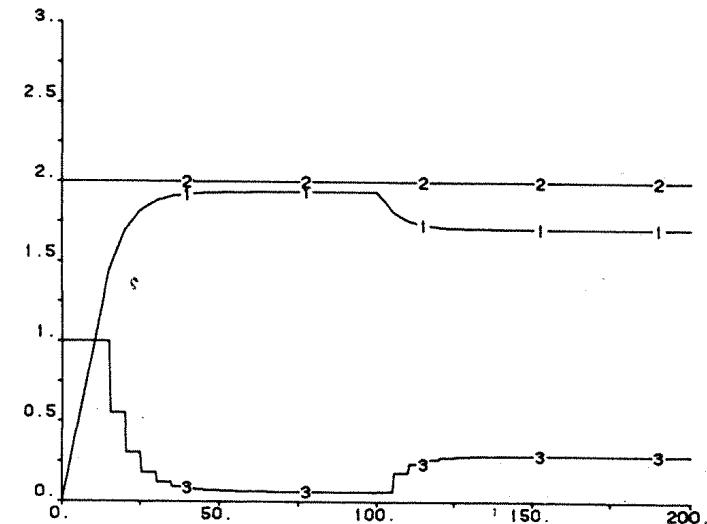


Figura C.2 Simulación con un regulador proporcional, $k=1$.

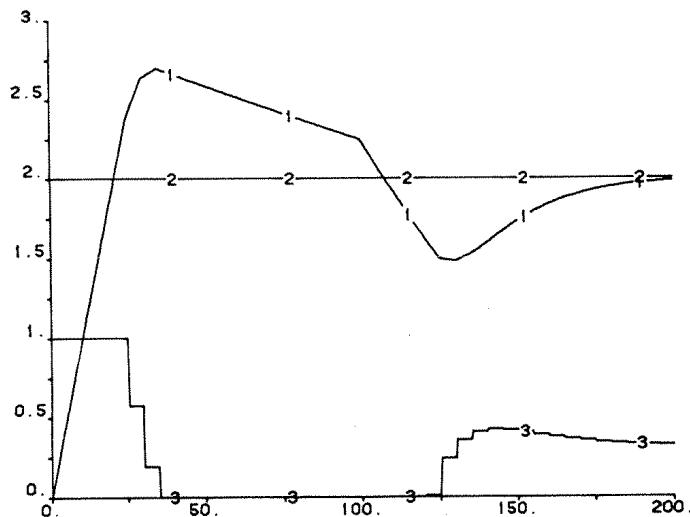


Figura C.3 Simulación con un controlador PI, $k=1$, $ti=25$.

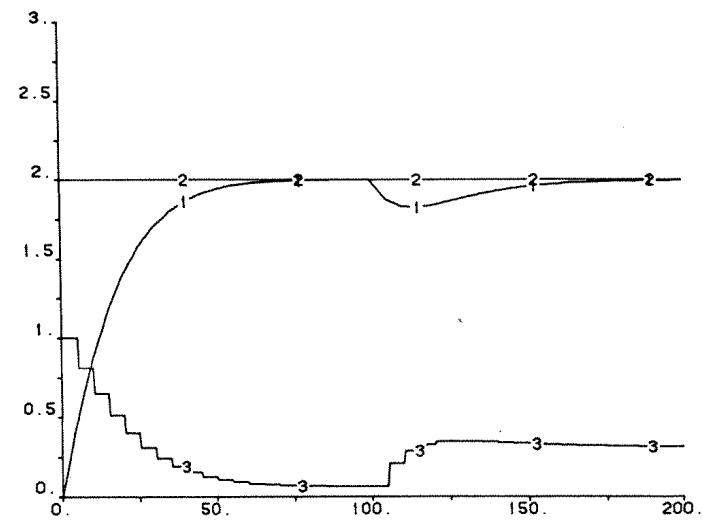


Figura C.4 Simulación con un controlador PI con compensación de saturación del término integral, $k=1$, $ti=25$.

```
>HCOPY " (véase la Fig. C.3).
>" Sobrelongación elevada, introducir antisaturación del término integral.
>" 
>SAVE pipar " Guardar los parámetros en el fichero denominado pipar.
>EDIT dpi " Editar el fichero dpi.
    EDIT
    >Lu= " Establecer una cadena de caracteres.
u=k*et1
  >C /u/v/ " Cambiar una cadena de caracteres.
v=k*et1
  >I u=IF v<0 THEN 0 ELSE IF v<1 THEN v ELSE 1 " Insertar Línea.
  >N " Siguiente Línea.
  >C /i1/i1+u-v/
ninte=i1+u-v
  >E " Abandona el editor.
  >SYST tank dpi regtank
  >GET pipar " Restaurar los parámetros.
  >STORE h[tank] href qin
  >SIMU
  >ASHOW h href qin-MARK
  >HCOPY " (véase la Fig. C.4).
```

Resumen de las órdenes del programa Simonon

A continuación se da una lista de órdenes, su sintaxis y una breve descripción. También se incluyen en esta lista algunas órdenes Intrac. Se pueden obtener más detalles sobre las órdenes utilizando la orden HELP. Esta lista es válida para la realización que funciona en un VAX 11/780. Se utiliza la siguiente notación.

- {op1|...|opn} Define diferentes alternativas, de las cuales debe elegirse una.
- [....] Los elementos dentro de corchetes son opcionales y pueden omitirse.
- {....}* Un asterisco indica que la parte anterior puede repetirse.
- <....> Define argumentos para las órdenes.

ALGOR {HAMPC|RK|RKFIX|DAS}. Selecciona la rutina de integración.

- HAMPC Predictor-corrector de Hamming (se utiliza por defecto).
- RK Runge-Kutta de intervalo variable.
- RKFIX Runge-Kutta de intervalo fijo.
- DAS Rutina de integración para sistemas rígidos.

AREA <filas> <columnas>. Define el área del dibujo que se usará posteriormente (véase SPLIT).

ASHOW. Representa variables almacenadas con escalado automático de los ejes. Véase SHOW para conocer su sintaxis.

AXES [<especificaciones de ejes> [<especificaciones de ejes>]]. Dibuja ejes.

<especif. de ejes>:={H|V} <valor min.> <valor máx.>
H, horizontal; V, vertical

DISP [(DIS|TP|LP){FF|LF})] [<variable>]*. Para visualizar variables.

DIS Visualizador (se utiliza por defecto).

TP Terminal.

LP Impresora.

FF Siguiente página (se utiliza por defecto si no se especifican variables).

LF Siguiente línea (se utiliza por defecto si se especifican variables).

Si no se especifican variables, se representan todas.

EDIT <nombre de un fichero>. Se utiliza para editar un fichero. El editor tiene dos modos INPUT y EDIT. El modo se cambia introduciendo una línea vacía. En el modo de edición es posible cambiar el fichero actual. Se dispone de las siguientes órdenes:

A[PPEND] <cadena> Añade una cadena a la línea.

B[OTT] Sitúa el puntero de líneas en el final.

C[HANG]/x/y/ Sustituye la cadena x por la cadena y.

D[EL] <entero> Borra n líneas.

DIS[ON|OFF] Permite o inhabíbe el eco del terminal.

E[XIT] Abandona el editor actualizando el fichero.

F[IND] <cadena> Busca una cadena de caracteres al principio de una línea.

I[NS] <cadena> Inserta una linea delante de la linea actual.

L[OC] <cadena> Localiza una cadena en cualquier posición en una línea.

LEAVE Abandona el editor sin actualización.

N[EXT] <entero> Mueve el puntero n líneas más adelante.

O[VERL] <entero> Sustituye n líneas con un nuevo texto.

P[RINT] <entero> Imprime n líneas.

R[ETYP] <cadena> Repite la línea actual.

T[OP] Sitúa el puntero de líneas al principio del fichero.

ERROR <intervalo de error>. Elige el intervalo de error para la rutina de integración. El valor por defecto es 0,001.

GET <nombre de fichero>. Obtiene el valor de los parámetros y los valores iniciales de un fichero que previamente se ha almacenado utilizando SAVE.

HCOPY [<factor de escala>] “<comentario>”. Hace una copia en papel de las curvas representadas en el visualizador o pantalla. La copia se escala con <factor de escala>, el cual puede estar en el intervalo (0,5, 1,6). El comentario también se representa en la copia.

HELP [<orden>|EDIT|INTRAC|SIMNON|SYSTEMS]. Obtiene más información sobre las órdenes, el editor, Intrac, Simnon y los sistemas típicos. Se obtiene un menú de órdenes escribiendo HELP.

HELPDEV={LP|TP|DIS}. Determina el dispositivo para la salida de la orden HELP (véase DIS). La salida por impresora se escribe en un fichero intermedio que se imprime con la orden LP.

INIT <variable de estado>: {<número>|<variable>}. Cambia el valor inicial de una variable de estado.

LET <variable>=<número>. Es una orden Intrac que también se utiliza para asignar valores a los parámetros del sistema.

LIST [(DIS|TP|LP){FF|LF})] {<nombre de fichero>}*. Lista ficheros. Es análoga a DISP pero para ficheros: por ejemplo, subsistemas.

LP. Lista en la impresora variables y ficheros creados con las órdenes DISP, LIST o PRINT.

MARK. Introduce texto en una gráfica. Para ver su sintaxis puede escribirse HELP INTRAC MARK.

NEWS. Obtiene características del paquete de simulación.

PAR <parámetro>: {<número>|<variable>}. Cambia el valor de un parámetro.

PLOT [<variable>]*[<variable>]]. Selecciona las variables para representar en el dibujo cuando se utiliza la orden SIMU. Ejemplo: PLOT X1 X2 proporciona X1 y X2 como funciones del tiempo, mientras que PLOT X1(X2) proporciona X1 como una función de X2.



PRINT [{DIS|TP|LP} [FF|LF]] <nombre de fichero> [<líneas>] [/<tiempo de comienzo>]. Lista los ficheros generados con las órdenes STORE+SIMU.

<líneas> se imprimen líneas empezando por <tiempo de comienzo>. Los otros parámetros son idénticos a los de DISP.

SAVE <nombre de fichero> [<nombre de sistema>] <-[PAR|INIT]>. Guarda valores de parámetros y valores iniciales en el fichero <nombre de fichero> que pueden ser utilizados con la orden GET. Si sólo se desean guardar los parámetros o los valores iniciales, deben utilizarse las opciones PAR e INIT, respectivamente.

SHOW {[<principio> <final>] {<variable>}* [<variable>] [-MARK]-LIST} [/<nombre de fichero>]. Representa variables almacenadas en el fichero <nombre de fichero>. Se utiliza con la orden STORE. Las variables especificadas se representan desde el instante <inicial> hasta <final>. Si se utiliza MARK las diferentes variables se numeran en el dibujo. La opción LIST lista los nombres de todas las variables.

SIMU [<tiempo inicial> <tiempo final> [<incremento>] [-<opción>]

[/<nombre de fichero> [<incremento>]]

<opción>:={MARK|CONT}

Simula el sistema desde <tiempo inicial> a <tiempo final> utilizando el máximo intervalo <incremento> [por defecto (tiempo inicial - tiempo final)/100]. Utilizando MARK se numeran las variables definidas por PLOT. Con CONT, la simulación continúa tomando como condiciones iniciales las variables de estado obtenidas anteriormente. Cuando se especifica <nombre de fichero>, las variables del dibujo se almacenan en <nombre de fichero> con <incremento> como intervalo de muestreo.

SPLIT <filas> <columnas>. Divide la pantalla en un máximo de seis áreas para dibujos.

<filas>:={1|2|3} por defecto 1.

<columnas>:={1|2} por defecto 1.

STATE. Se utiliza con la rutina de integración DAS. Para ver su sintaxis, escribir HELP STATE.

STOP. Abandona el programa Simnon.

STORE [<variable>]* [-ADD]. Se utiliza para seleccionar variables que se almacenan en cada simulación. Con ADD se pueden añadir nuevas variables a una lista definida anteriormente. Las variables se pueden visualizar utilizando ASHOW o SHOW e imprimir con PRINT.

SWITCH <opción> {ON|OFF}
 <opción>:={CLOCK|DATE|ECHO|EXEC|LOG|TRACE}
 Controla la ejecución de Intrac.

CLOCK	Añade tiempos a la salida gráfica. Se supone OFF por defecto.
DATE	Añade la fecha a la salida gráfica. Se supone OFF por defecto.
ECHO	Se representan las Macro-órdenes. Se supone OFF por defecto.
EXEC	Ejecuta órdenes mientras se genera un macro. Se supone OFF por defecto.
LOG	Las órdenes ejecutadas dentro de un macro se registran en la impresora. (Se supone OFF por defecto.)
TRACE	Afecta a ECHO y LOG.

SYST {<subsistema>}* [-<opción>] [/<nombre de fichero>]
 <opción>:={EDIT|EXIT}

Define el sistema. Las subrutinas se compilan. Si hay varios subsistemas, el último debe ser un sistema de interconexión. EDIT quiere decir que el compilador está dentro del editor en cada fichero. Si se especifica nombre de fichero, se escriben las ecuaciones ordenadas en el texto de un fichero.

TEXT '<cualquier texto que no contenga comillas>'. Incluye texto en el dibujo.

TURN <opción> {ON|OFF}

<opción>:={S125|DARK|DIS|OVFLO}

Permite seleccionar determinadas condiciones.

S125	Selecciona una serie limitada de factores de escala (1., 2., 5.) cuando se eligen las escalas de los ejes. Se supone OFF por defecto.
DARK	ON indica que no se conectan las curvas de las gráficas entre los instantes de muestreo. Se supone OFF por defecto.
DIS	Informa al Simnon si el usuario dispone de un visualizador gráfico. Se supone ON por defecto.
OVFLO	ON indica que la simulación debe parar si hay un rebasamiento en los cálculos. Se supone OFF por defecto.

Referencias

- ELMQVIST, H. (1975): Simnon user's guide. Department of Automatic Control, Lund Inst. of Techn., CODEN: LUTFD2(TFRT-3091)/(1975).
- ELMQVIST, H. (1977): An Interactive Simulation Program for Nonlinear Systems. Proceedings Simulation '77, Montreux 1977.
- ÅSTRÖM, K. H. (1982): A Simnon Tutorial. Department of Automatic Control, Lund Inst. of Tech., CODEN:LUTFD2(TFRT-3168)/(1982).

Demostración del teorema 10.1



Teorema 10.1. Sean A , B y C polinomios de coeficientes reales. Entonces, la ecuación

$$AX + BY = C \quad (\text{D.1})$$

tiene solución si y sólo si el máximo común divisor de A y B divide a C .

Demuestra. Sin pérdida de generalidad, puede suponerse que $\text{gra } A \geq \text{gra } B$. Introducimos $A_0 = A$ y $A_1 = B$. Sea Q_1 el cociente y A_2 el resto obtenido cuando A_0 se divide por A_1 ; es decir,

$$A_0 = A_1 Q_1 + A_2$$

donde $\text{gra } A_2 < \text{gra } A_1$. Dividimos A_1 por A_2 :

$$A_1 = A_2 Q_2 + A_3$$

Continuando del mismo modo se obtiene

$$\begin{aligned} A_0 &= A_1 Q_1 + A_2 \\ A_1 &= A_2 Q_2 + A_3 \\ &\vdots \\ A_{n-2} &= A_{n-1} Q_{n-1} + A_n \\ A_{n-1} &= A_n Q_n + A_{n+1} \\ A_{n+1} &= 0 \end{aligned} \quad (\text{D.2})$$

donde A_{n+1} es el primer polinomio nulo. Puesto que

$$\text{gra } A_0 > \text{gra } A_1 > \text{gra } A_2 > \dots > \text{gra } A_{n+1}$$

el algoritmo debe terminar en un número finito de pasos. El algoritmo se denomina algoritmo de Euclides.

Como $A_{n+1} = 0$, se deduce que A_{n-1} es divisible por A_n . Por tanto, el polinomio A_n es el máximo común divisor de A_n y A_{n-1} . Se deduce de (D.2) que

$$A_{n-2} = A_{n-1} Q_{n-1} + A_n = A_n Q_n Q_{n-1} + A_n$$

Por tanto, el polinomio A_n es un factor de A_{n-1} y A_{n-2} . También es el máximo común divisor porque si hubiese un factor Q mayor, entonces se tendría que

$$A_{n-2} = A_{n-1} Q$$

Sin embargo, esto implica que $A_n = 0$, lo cual contradice la suposición de que el algoritmo finaliza con $A_{n+1} = 0$. Procediendo recursivamente, se encuentra que A_n es el máximo común divisor de $A_0 = A$ y $B_0 = B$. Entonces, la ecuación (D.1) puede escribirse como

$$AX + BY = (A'X + B'Y)A_n = C \quad (\text{D.3})$$

Esto demuestra que la ecuación no tiene solución a menos que A_n sea un factor de C .

Si A_n es un factor de C , pueden dividirse ambos lados de (D.3) por C ; esto da la ecuación (D.3), donde A y B no tienen factores comunes. En este caso el algoritmo de Euclides da $A_n = 1$ como el máximo común divisor de A y B . El algoritmo (D.2) da

$$1 = A_n = A_{n-2} - Q_{n-1} A_{n-1}$$

De (D.2) se deduce que

$$A_{n-1} = A_{n-3} - A_{n-2} Q_{n-2}$$

Por tanto,

$$\begin{aligned} 1 &= A_{n-2} - Q_{n-1}(A_{n-3} - Q_{n-2} A_{n-2}) \\ &= -Q_{n-1} A_{n-3} + (1 + Q_{n-1} Q_{n-2}) A_{n-2} \\ &= U_{n-3} A_{n-3} + V_{n-2} A_{n-2} \end{aligned}$$

Procediendo recursivamente, se encuentra que

$$1 = U_0 A_0 + V_1 A_1 = U_0 A + V_1 B$$

Entonces, una solución a la ecuación (D.1) es

$$X = U_0 C$$

$$Y = V_1 C$$

Índice alfabético

- Acción integral, 261, 367
Ackermann, J., 34
Ackermann's formula, 242
Actuador, 92, 432
Actuadores no lineales, 432
Alcanzabilidad, 132
Algoritmo de Dahlin Higham, 288
Algoritmo de Euclides, 491
Algoritmo de la raíz cuadrada, 397
Algoritmo explícito, 406
Algoritmo rápido, 329
Amortiguamiento relativo, 244
Análisis de la función descriptiva, 446
Ancho de banda, 208, 284, 419, 441
Anderson, B.D.O., 191, 333
Anderson, N., 467
Antienmascaramiento, 45, 115
Antisaturación del término integral, 224, 433–437
 forma entrada salida, 436
 modelo general en el espacio de estados, 434
 observador explícito, 433
 regulador PI, 223
Antoniou, A., 236
Aproximación de diferencia en retroceso, 214
Aproximación de sistemas continuos, 213
 diferencia en avance, 214
 diferencia en retroceso, 214
 método de Euler, 214
 transformada Z, 213
 Tustin, 214
Aproximación de Tustin, 214
Aproximación por diferencias, (*véase* Aproximación de sistemas continuos)
Arbib, M.A., 35
Aritmética de punto flotante, 443
Aritmética de punto fijo, 443
Aspectos operacionales, 438
Aström, K.J., 33, 35, 89, 150, 151, 190–191, 303,
 333, 382, 403, 422–423, 489
Auslander, D.M., 237
Autoregresión, 168
Autosintonizado, 418

Banda lateral, 99
Bar-Shalom, Y., 423
Baradello, C.S., 89
Barker, R.H., 31, 34
Barnes, J.G.P., 467
Battin, R.H., 191
Bellman, R., 32, 35, 422, 475
Bertram, J.E., 31, 88, 150, 265, 466
Bierman, G., 334, 397, 403
Björk, G., 467
Blanchard, J., 150
Blomberg, H., 33, 35
Bolyanski, V.G., 35
Bongiorno, Jr., J.J., 382
Boudreau, J.A., 211
Box, G.E.P., 191, 382, 403
Brinch-Hansen, P., 467
Bristol, E.H., 210, 237
Brown, G.S., 190
Buckley, P.S., 210
Bucy, R.S., 191, 334
Burns, B.R.A., 211
Buttner, M., 466

Caines, P.E., 423
Cálculo de la varianza, 179
Cálculo del operador de desplazamiento, 67
Cálculo del producto escalar, 444
Cambio de parámetros sin salto, 442
Campbell, D.P., 73, 80, 258
Cero, 73, 80, 258
 correspondencia, 80
Chan, C.S., 237
Chestnut, H., 190
Chong, H.Y., 466
Circuito mantenedor, 97
 función de transferencia, 97
mantenedor de orden superior, 41
mantenedor de orden cero, 40
mantenedor de primer orden, 41
Clarke, D.W., 422
Cohn, A., 124, 150

Colocación de polos:
por realimentación de la salida, 255, 266, 277,
348, 354
por realimentación del estado, 238

Compensación de dipolo, 285

Constante de derivación, 219

Constante de integración, 219

Construcción del modelo, 384

Control adaptativo, 33, 404, 440
control adaptativo por modelo de referencia, 414, 415
control autosintonizado, 405–407
control dual, 417
control preventivo, 417
controlador de certidumbre, 417

Control adaptativo con modelo de referencia, 405, 414, 415

Control algorítmico del modelo, 288

Control de referencia, 18

Control de varianza mínima, 206, 338, 346

Control digital directo (DDC), 19, 460

Control dual, 405, 417

Control en tiempo finito, 145, 246

Control lineal cuadrático, 304, 309, 339
caso determinístico, 309
estabilidad, 315
información completa del estado, 313
margen de ganancia, 316
polos en lazo cerrado, 314
propiedades, 314
selección del período de muestreo, 329

Control lineal cuadrático gaussiano, 304, 326, 355
el problema del servo, 328
propiedades, 328

Control óptimo, 207, 304, 335

Controlabilidad, 132, 196

Controlador PID:
aspectos operacionales, 222
digital, 218
estructuras, 220
reglas de sintonizado, 225

Controlador PID digital, 218
forma de posición, 220
forma de velocidad, 220
forma incremental, 220

Controlador preventivo, 417

Convertidor A-D, 53, 92, 443

Convertidores D-A, 53, 91, 443

Criterio de Nyquist, 126

Cuantización, 443–449
análisis de la función descriptiva, 446
análisis de sensibilidad, 449
análisis lineal, 448
análisis no lineal, 445
ciclos límites, 445

Curry, E.E., 466

Curry, R.E., 118

Dahlin, E.B., 303

Dahlqvist, A., 467

Datos erróneos, 431

DDC, (*véase* Control digital directo)

Decaulne, P., 190

Densidad autoespectral, 165

Densidad espectral, 165
interpretación, 165

Densidad espectral cruzada, 165–166

Densidad espectral racional, 176

Desbordamiento, 443

Desbordamiento negativo, 443

Descomposición de interruptores, 117

Descomposición de Kalman, 138

Diseño ascendente, 198, 201

Diseño de controladores, 192

Diseño descendente, 198

Doetsch, G., 89

Dorf, R.C., 211

Doyle, J.C., 334

Dualidad, 327

Ecuación característica, 123

Ecuación de Lyapunov, 132

Ecuación de Riccati, 311
algebraica, 317
solución, 329

Ecuación del sistema, 54
inversa, 135

Ecuación diferencial estocástica, 183

Ecuación diofántica, 274, 356

Ecuación en diferencias estocástica, 169–170

Ecuación en diferencias, 30

Ecuación normal, 390

Edmunds, J.M., 422

Edwards, R., 466

Egardt, B., 423

Elgerd, O., 211

Elmqvist, H., 151, 489

Enmascaramiento, 37, 42, 58, 100, 428

Equivalente en tiempo discreto, (*véase* Aproximación de sistemas continuos)

Error de cuantificación de la integración, 224

Errores de medida, 157

Especificaciones, 206

Especificaciones de diseño, 206

Especro racional, 176

Estabilidad, 121
asintótica, 121
entrada acotada salida acotada (BIBO), 122
pruebas, 123
criterio de Nyquist, 126
de Jury, 125
método de Lyapunov, 131

Estabilidad asintótica, 121

Estabilidad BIBO, 122

Estabilidad entrada acotada–salida acotada, 122

Estimación de parámetros, 387

Estimación óptima, (*véase* Filtro de Kalman)

Estimador, (*véase* Observador)

Estimador del estado, (*véase* Observador)

Estructuración, 197

Excitación persistente, 391

Eykhoff, P., 33, 35, 403

Factorización espectral, 176, 186, 317, 355

Factorización, U-D 336

Falb, P.L., 35

Fase no mínima, 82

Feldbaum, A.A., 423

Filtrado, 320

Filtro de Bessel, 46

Filtro de Butterworth, 46

Filtro de eliminación de banda (Notch), 296

Filtro de Kalman, 305, 322–326
estimación actual, 322
propiedades en frecuencia, 324

Filtro digital, 213

Filtro ITAE, 46

Flower, J.O., 119

Forge, S.C., 119

Forma asociada, 65, (*véase también* Forma canónica)

Forma canónica:
controlable, 65
diagonal, 64
observable, 64

Forma controlable, 65

Forma cuadrática, valor medio, 308

Forma de Jordan, 64

Forma diagonal, 50

Forma observable, 64

Foss, A.S., 211

Franklin, G.F., 31, 34, 35, 51, 88, 273, 334

Frecuencia de cruce, 209

Frecuencia de Nyquist, 40

Frecuencia natural, 244

Frecuencia no amortiguada, (*véase* Frecuencia natural)

Frecuencia:
contenido de las señales muestreadas, 99
diseño de la respuesta, 232–234
predesvio, 215
respuesta de los sistemas muestreados, 98
solapamiento, 42, 44

Freeman, H., 31

Friedland, B., 31

Función de autocovarianza, 165, (*véase también* Función de covarianza)

Función de correlación, 166

Función de coste, 306
determinación, 319
muestreo, 307
transformación, 308

Función de covarianza cruzada, 165

Función de covarianza, 164, 171

Función de distribución, 164

Función de Lyapunov, 131, 171

Función de transferencia discreta, 76, 109
cálculo, 76

Función matricial, 473

Función ponderatriz, 66

Gamkrelidze, R.V., 35

Ganancia proporcional, 219

Gantmacher, F.R., 475

Gardenhire, L.W., 51

Gauss, K.F., 403

Gawthrop, P.J., 422

Generador de perturbaciones, 188

Gille, J.C., 190

Glicksberg, I., 35

Goff, K.W., 237

Gold, B., 236, 466

Goodwin, G.C., 35, 403, 423

Gordon, G., 151

Gould, L.A., 191, 382

Gross, O.A., 35

Guía del operador, 18

Gunkel, III, T.L., 334

Hagander, P., 89

Hanson, R.J., 403

Helmersson, A., 423

Hess, S.F., 466

Higham, J.D., 303

Historia:
tecnología de computadores, 16
teoría de control con computador, 30

Ho, Y.C., 151

Hurewicz, W., 31, 34

Identificación de sistemas, 33, 383, 385, 390

Identificación recursiva, 392

Impulso, 157

Información completa del estado, 306, 313

Información incompleta del estado, 306

Inicialización, 440

Innovaciones, 178, 343

Integración condicional, 224

Integrador doble, 56, 142–145, 240, 243, 246, 251, 254–257, 314, 316, 468
con retardo, 60

Interfase con el operador, 438

Inverso inestable, 82
sistemas con, 82, 349

Isermann, R., 34, 403, 422

Jabr, H.A., 382

Jackson, L.B., 466

James, H.M., 191

Jenkins, G.M., 191, 382, 403

Jerri, A.J., 51

Jury, E.I., 31–33, 51, 88–89, 119, 125, 150–151, 190

Kailath, T., 89, 265

Kaiser, J.F., 191, 382

Kalman, R.E., 31–35, 88, 150–151, 191, 333–334, 422

Karlin, S., 191

Kezer, A., 423

Knowles, J.B., 466

Kolmogorov, A.N., 191

Konar, A.F., 119

Kotelnikov, V.A., 34, 51

Kranc, G.M., 31, 119

Kucera, V., 33, 35, 89, 303, 382

Kuo, B.C., 35, 237, 334

Kuo, F.F., 465

Kwakernaak, H., 333

Landau, L.D., 423
Laning, J.H., 191
Laub, A.J., 334
Lawden, D.F., 34
Lawson, C.L., 403
Lefschetz, S., 32
 Lema de la inversión de la matriz, 393
Li, Y.T., 118
Lin, Y.H., 423
Lindorff, D.P., 119
Linville, W.K., 31, 34
Ljung, L., 403, 423
 Longitud de palabra, 443
Luenberger, D.G., 265
 Lugar de las raíces, 124, 140, 285
MacColl, L.A., 31, 33
Mahesh, J.K., 119
 Mantenedor de orden cero, 40
 equivalente, 55
 muestreo, (*véase* Muestreo, de un sistema)
 Mantenedor de orden superior, 41
 Mantenedor de primer orden, 41
 Margen de fase, 209
 Margen de ganancia, 209
 Matriz de controlabilidad, 133
 inversa, 135
 Matriz de observabilidad, 137
 Matriz de transformación, cálculo, 134
 Matriz exponencial, cálculo, 55
Mayer, R.W., 190
McGarty, T.P., 191
Medio móvil, 168
Meiry, J.L., 118
Melzer, S.M., 334
Método de Euler, 214
Método de la transformada w, 124
Método de Lyapunov, 131
Método de máxima probabilidad, 392
Método de mínimos cuadrados, 386–387
 cálculos recursivos, 392
 ecuación normal, 390
 extensiones, 391
 factorización U D, 396
 interpretación estadística, 391
 ponderación exponencial de los datos, 396
Métodos de las transformadas, 30
Miminis, G.S., 265
Minicomputador, 20
Mischenko, E.F., 35
Modelo de inventario, 472
Modelo de modulación de un tren de impulsos, 96–97
Modelo de modulación, 92–93
Modelo estroboscópico, 52, 90
Modelo matemático, 384
Modelos del espacio de estados, transformación, 63
Modelos entrada salida, 66
Modelos estocásticos de perturbaciones, 162
Modelos:
 orientados a computador, 52
 orientados a proceso, 90

Modo de estimación, 440
Modo de seguimiento, 439
Modulación, 93
Mónico, 268
Monopoli, R.V., 423
Monorey, P., 466
Moore, J.B., 191, 333
Morse, A.S., 423
Motor, 278–281, 469
Motor de combustión interna, 105
Muestreadores, 95
Muestreo:
 de un sistema, 54
 de una señal, 36–37, 100
 ecuación diferencial estocástica, 187
 frecuencia, 37
 instante, 23
 intervalo, (*véase* Muestreo, período)
 inverso de, 58
 período, 23, 37, 428
 período, selección, 47, 83, 216, 227, 244, 293,
 329, 371, 374–376
 retenedor de orden cero, 54–55, 74
 sistema con retardo temporal, 59
 teorema, 30, 37
 tiempo, 16, 37
Muestreo multiperiódico, 23, 37, 114
Muestreo periódico, 37
Muestreo y mantenimiento, 93, 98
 idealizado, 98
Narendra, K.S., 151, 423
Neuman, C.P., 89
Newton, G.C., 191, 382
Nichols, N.B., 191, 225, 237
Nyquist, H., 30, 50
Observabilidad, 132, 136, 196
Observador, 249, 425, 433
 cálculo directo, 249
 de orden reducido, 253
 empleando un sistema dinámico, 252
Observador de Luenberger, 253
Oldenburg, R.G., 30, 33
Operador de desplazamiento hacia atrás, 67
Operador de desplazamiento hacia delante, 68
Operador de transferencia discreta, 71
Oppenheim, A.V., 236, 446
Orden de un sistema, 73
Oscilaciones ocultas, 31, 146, 350
Oscilador armónico, 58, 122, 470
Paige, C.C., 265
Papon, J., 303
Papoulis, A., 191
Pappas, T., 334
Parametrización, 441
Parker, S.R., 466
Parks, P.C., 423
Parzen, E., 191
Payne, R.L., 35, 403
Pelegrin, M.J., 190

Pernebo, L., 303
Perturbaciones, 152, 157
 determinísticas a trozos, 159
 modelos clásicos, 156
 modelos estocásticos, 162
 reducción en la fuente, 153
 reducción por prealimentación, 155
 reducción por predicción, 155
 reducción por realimentación, 154, 368, 371
Perturbaciones de la carga, 156
Peterka, V., 382, 422
Philips, R.S., 191
Planificación de ganancia, 405, 413
Polinomio característico, 123
Polinomio reciproco, 69
Polo, 73, 78, 258
 correspondencia, 79
 exceso, 69
Polya, G., 210
Pontryagin, L.S., 32, 35
Powell, J.D., 35, 273
Prager, D., 422
Prealimentación, 155, 203
Predesvio, 215
Predicción, 155, 320, 340
Predicción óptima, 320–326, 340–346
Predictor de Smith, 203, 227–229, 286
Prefiltrado, 45, 428
Prefiltro analógico, (*véase* Prefiltrado)
Presentación de la información, 440
Principio de equivalencia de certeza, 305, 417
Principio de los argumentos, 127
Principio del modelo interno, 367
Principios de control, 199
Probabilidad de transición, 169
Probabilidad inicial, 169
Problema del servo, 208, 258, 365–366
Problemas del regulador, 206
Proceso aleatorio, 162
Proceso AR, 168
Proceso ARMA, 168, 177, 337
Proceso de Markov, 168
Proceso estacionario, 165, 174, 185
 filtrado, 174, 185
Proceso gaussiano, 164–165
Proceso MA, 168
Proceso normal, 164
Procesos débilmente estacionarios, 165
Procesos estocásticos, 163
 AR, 168
 ARMA, 168, 177
 completamente determinístico, 163
 continuo, 181
 discreto, 163
 MA, 168
 Markov, 168
Programación, 454
Programación dinámica, 309
Programación en tiempo real, 457
Pseudoinversa, 390
Pulso, 157

Rabiner, L.R., 236, 466
Ragazzini, J.R., 31, 34, 51, 88
Ramadge, P.J., 423
Rault, A., 303
Realimentación, 154, 202
Realimentación del estado, 238
 con integrador, 261
 rediseño, 229
Realización, 163, 450
Realización de un sistema discreto, 426
Rebasamiento, 208
Reconstrucción, 40, 252
Reconstrucción de la señal, 36, 40
 mantenedor de orden cero, 40
 mantenedor de orden superior, 41
 mantenedor de primer orden, 41
Shannon, N., 40
Rediseño digital, 212
Redondeo, 443–444, 448
Regulador PI, 437, 454, 459
Relación de amortiguamiento, (*véase* Amortiguamiento relativo)
Representación de las innovaciones, 178
Representación del estado con un observador explícito, 425
Representación del regulador, 425
 entrada-salida, 268, 426
 espacio de estados, 252, 259–260, 425
Representación muestreada, 96
Respuesta de pulso, 66
Retardo de cálculo, 429–430
Retardo temporal, (*véase* Retardo)
Retardo:
 de cálculo, 429
 muestreo de un sistema con retardo, 59
Richalet, J., 303
Rink, R.E., 466
Rissanen, J., 265
Rizaldo intermuestreos, (*véase* Oscilaciones ocultas)
Robustez, 129, 368, (*véase también* Sensibilidad)
Rosenbrock, 33, 35, 89, 265
Ruido blanco, 166, 182
Safonov, M.G., 281
Sandell, Jr., N.R., 334
Sarachik, P.E., 31
Saridis, G.N., 423
Sartorius, H., 30, 33
Saturación, (*véase* Saturación del término integral)
Saturación del integrador, (*véase* Saturación del término integral)
Saturación del término integral, 223–224, 433, 456
Schafer, R.W., 236, 466
Schur, J., 124, 150
Seguimiento del modelo, 287
Seguridad, 442
Señal determinística a trozos, 159–162
Señal PWD (*véase* Señal determinística a trozos)
Sensibilidad, 282, 290, 368 (*véase también* Robustez)
Shannon, C.E., 30, 34, 51
Shinskey, F.G., 210, 237

Otros libros de INFORMATICA publicados por



Simmon, 142, 476
Simon, H.A., 334
Simulación, 141, 476
Simulación digital, 142, 476
Sincronización, 92
Sintonizado por relación de funciones, 418
Sistema intrínsecamente muestrado, 27
Sistema muestrado, 16
 intrínsecamente muestrado, 27
Sistemas continuos:
 aproximación, 212
 en Simmon, 478
 muestreo, 53-54, 59
Sistemas controlados con computador, 15-16
 aspectos operacionales, 438
 guía del operador, 18
 interfase del operador, 438
 modelos, 52-53
 naturaleza periódica, 93
Sistemas de interconexión, 479
Sistemas discretos, 16, 120
 en Simmon, 478
 realización, 462
Sivan, R., 333
Sklansky, J., 31
Slaughter, J.B., 466
Smith, O.J.M., 237
Söderström, T., 403
Sorensen, H.W., 403
Stein, G., 334
Sternby, J., 98
Suavizado, 320

Takahashi, Y., 237
Tecnología de computadores, 16
Teorema de Cayley Hamilton, 473
Teorema de separación, 305, 326, 348
Teorema del valor final, 75
Teoría algebraica de sistemas, 33, 105
Teoría de control estocástico, 32
Teoría de control con computador, 23
 control de referencia, 18
 control digital directo, 19
 guía del operador, 18
 historia, 30
Teoría del espacio de estados, 32
Testud, J.L., 303
Thornton, C.L., 397

Tiempo de asentamiento, 208
Tiempo de subida, 208
Tou, J.T., 31, 34
Transferencia sin salto, 222, 439
Transformación bilineal, 214
Transformación de modelos del espacio de estados, 63
Transformada Z, 31, 73, 107
 aproximación, 213
 inversa, 75, 107
 modificada, 31, 114
 retardada, 31
Transformada Z inversa, 75, 107
Transformada Z modificada, 114
Transitorios de cambios, 441
Tse, E., 423
Tsien, H.S., 191
Tsypkin, Y.Z., 31, 34, 88

Unbehauen, H., 423

Valavani, L.S., 423
Valor en estado estacionario, 140
Valor medio, 164, 169, 171
Van Dooren, P., 334

Watts, D.G., 403
Wellstead, P.E., 422
Whitaker, H.P., 423
Whitbeck, R.F., 119
Whittle, P., 191
Wiener, N., 191, 381
Williams, A.B., 465
Willsky, A.S., 236, 466
Willson, Jr., A.N., 466
Windett, G.P., 119
Wirth, N., 210
Wittenmark, B., 35, 303, 422-423
Wolowich, W.A., 265, 303
Wonham, W.M., 33, 35

Yamron, J., 423
Ylinen, R., 33, 35
Youla, D.C., 382

Zadeh, L.A., 31, 34
Zanker, P., 422
Ziegler, J.G., 225, 237

Arquitectura de Computadores, Miguel, De/Angulo. 516 páginas.

Describe los principios fundamentales en los que se basa la organización y la estructura de los modernos ordenadores, así como las líneas maestras que siguen las máquinas inteligentes de la 5^a generación.

Auditoría informática, Thomás/Douglas. 216 páginas.

Trata la auditoría informática, desde la especificación del sistema hasta la realización de auditorías de datos reales y resultados de los sistemas que se están utilizando.

Control de procesos industriales por computador, Angulo/Nó. 176 páginas.

Sirve de guía para tomar contacto con el control de los procesos de tipo continuo por ordenador.

Curso teórico-práctico sobre microprocesadores, Angulo. 426 páginas.

De gran utilidad para alumnos universitarios de Informática y Electrónica.

Diseño de programas para sistemas, Gauthier. 290 páginas.

Presenta las técnicas del sistema en función de la lógica del diseño, del criterio de aplicación y de la metodología para su aplicación.

Estructura, lógica y diseño de programas, Cohen. 338 páginas.

Ofrece una gran variedad de ejemplos probados, que van desde los puramente comerciales a los sistemas científicos en tiempo real.

Informática para actividades profesionales, Grigorieff. 136 páginas.

Resuelve, sencilla y claramente, qué es la informática; cómo conocer su mercado; cuáles son sus características, sus ventajas e inconvenientes.

Ingeniería del Software, Thorin. 148 páginas.

Facilita el estudio y comprensión de todos los factores básicos de la programación y de las estructuras de datos.

Investigación operativa con el computador, Whitaker.

Programación lineal. Análisis de redes. Análisis del camino crítico. Control de inventarios. Valoración de inversiones. Colas y simulaciones.

Manual de los sistemas de información. Hartman. Dos tomos. Tomo I. 348 páginas. Tomo II. 384 páginas.

El objetivo fundamental es contribuir al progreso de la ingeniería de sistemas.

El microordenador en la pequeña empresa, Pannell. 148 páginas.

Ayuda a resolver los problemas que puedan presentarse a la hora de instalar un ordenador.

Microporcesadores de 16 bits, Angulo. 356 páginas.

Dirigido a lectores que se supone conocen ya los microporcesadores de 8 bits.

Microporcesadores de 32 bits, Angulo. 362 páginas.

Se comentan todos los microporcesadores de 32 bits, pero también se hace una descripción detallada de los modelos de la familia 32000 de National Semiconductor y de la 80386 de Intel.

Redes de Computadores, Menascé.

Conceptos básicos. Protocolos de acceso a la red. Protocolos de transporte. Protocolos de alto nivel. Redes locales. Banco de datos distribuidores.

Sistemas de explotación de computadores, Crocus. 424 páginas.

Responde a los deseos de quienes se interesan por seguir de cerca la evolución de los sistemas operativos.

Sistemas de información. Análisis. Diseño. Puesta a punto, Lucas. 540 páginas.

Describe los sistemas de informática de los departamentos de ordenadores de las empresas.

Sistemas operativos PC-MS/DOS, Politis/Vanryb. 230 páginas.

Trata todas las versiones del PC-DOS, desde la 1 hasta la 3.2, así como las correspondientes del MS-DOS.

Subrutinas para aumentar la precisión de cálculo, Barrio.

Describe un conjunto de subrutinas BASIC que permiten aumentar la precisión de cálculo de cualquier ordenador.

Teleinformática. Introducción, panorámica y perspectivas, Galán/Cordero. 132 páginas.

Presenta ideas sobre las futuras tendencias que constituirán el armazón de la teleinformática.