

CAPÍTULO 0  
INTRODUÇÃO

FILTRO DE KALMAN  
APLICADO À NAVEGAÇÃO AÉREA

NOTAS DE AULA

Agenor de Toledo Fleury

# FILTRO DE KALMAN APLICADO À NAVEGAÇÃO AÉREA

## O. INTRODUÇÃO

"Se as observações astronómicas e outras quantidades nas quais a determinação de uma órbita é baseada estivessem absolutamente corretas, então os elementos da órbita também estariam corretos, enquanto obtidas de três ou quatro observações, admitindo que o movimento se dá exatamente de acordo com as leis de kepler. logo, se outras observações fossem usadas, estas observações deveriam ser confirmadas nas não corrigidas. Mas, como todas as nossas medidas e observações são não mais que aproximações da verdade, o mesmo ocorre para todos os cálculos baseados nessas medidas, e, a fraude maior de todos os cálculos feitos, concernentes a problemas concretos, deve ser a de aproximar, tanto quanto possível, a realidade. Isto só pode ser conseguido através de uma combinação desejável de um número de observações maior que o número absolutamente necessário para a determinação das quantidades não conhecidas. Para a determinação de uma órbita, este problema só pode ser obtido quando houver um número aproximado de órbitas já tais sejam conseguidas, o qual, em última análise, deve ser corrigido para se satisfazer todas as observações da maneira mais certa possível."

A afirmação acima, embora pareça extraída de algum texto moderno em determinação de órbitas, foi feita por Gauss, em 1809, em seu trabalho "Teoria Motus". Esse era o preâmbulo para a introdução do "Método dos

"Mínimos Quadrados", técnica até hoje empregada com muito sucesso em diversas áreas do conhecimento e só recentemente superada pelos chamados métodos modernos de estimação.

Como se nota, desde o começo do século XIX, já havia muitas preocupações em se extrair de uma série de medidas, aparentemente redundantes, um melhor conhecimento sobre o sistema físico ou o fenômeno físico que se estavam observando. Durante muito tempo, no entanto, o problema de estimação ficou restrito ao método dos mínimos quadrados e suas variantes porque não havia uma maneira clara de tratamento, no que diz respeito ao seu aspecto matemático.

Apenas após os trabalhos de Kolmogorov, por volta de 1933, sobre a teoria matemática de probabilidades, com sua definição axiomática, se formou a base teórica necessária para a colocação do fenômeno físico como um problema estocástico. Com isso, em 1949, Wiener, depois pelo primeiro vez a solução recursiva de um problema de estimação do estado, embora, nesse caso, sob uma série de hipóteses restritivas e com uso de técnicas do domínio de frequências.

Em 1960, Khatua publicou, não sem dificuldades e com grande esforço da comunidade científica, o primeiro de uma série de trabalhos que, a fim de levantar as restrições existentes, no método de Wiener, parecia ter feito sido feito sob medida para aplicação em computadores digitais. Muito tempo se passou antes que se reconhecesse a importância dos trabalhos de Khatua e, hoje, essa técnica é aplicada nos mais diversos campos da Engenharia, que se sempre com excelentes resultados. Desde então, uma série extensa de outras técnicas foi também desenvolvida, de maneira que a teoria de estimação de

parâmetros e estados é hoje uma realidade contestável em campos tão diversos como Engenharia de Controle, Engenharia Civil, Biocenegenética, Economia ou Ciências Sociais.

### 0.1. A necessidade de modelos estocásticos

Toda a teoria de análise de sistemas e controle foi desenvolvida usando-se sistemas determinísticos. Então, dado um fenômeno físico, um engenheiro tentaria desenvolver um modelo matemático para representá-lo convenientemente. Para isso, usaria seu conhecimento físico, leis fundamentais e conhecimentos empíricos. Em seguida, estaria apto a investigar se a estrutura do sistema é correta e determinar como são as respostas do sistema proposto. Finalmente, construiria dispositivos de medida para observar o comportamento real do sistema modelado.

O uso de modelos estocásticos absolutamente não elimina as pressões acima, mas, permite melhorar, e muito, o nosso conhecimento sobre o sistema. O caminho puramente determinístico sugerido não prevê meios de se fazer totalmente a análise e projeto do sistema e, nesse caso, pode acontecer do sistema projetado se comportar de uma maneira diferente da previsão. Existem três razões básicas para isso:

a) Nenhum modelo matemático é perfeito. Como exemplo, a estrutura de um zorronze tan suporta mais flexão. Se pretendemos construir um modelo aplicável, obrigatoriamente só poderemos levar em conta um número finito desses modos, ou seja, zóquedas que serão os predominantes. Com isso, muitos efeitos são deixados como não-modelados.

b) Sistemas dinâmicos são afetados não apenas pelos controles (determinísticos) impostos mas também por fatores

ções que não podemos controlar nem modelar deterministicamente. Como exemplo, se um piloto tentar corrigir uma queda de orientação angular em sua aeronave, a resposta real será diferente da esperada devido aos efeitos do vento, à imprecisão das respostas dos estuadores das superfícies de controle e, mesmo, à facilidade de gerar exatamente a resposta desejada por suas mãos e fôlego sobre os pedais de comando.

c) Os sensores nunca fornecem os dados perfeitos e completos sobre o sistema. Em primeiro lugar, sensores nunca fornecem leituras exatas sobre as variáveis desejadas, introduzindo-as suas próprias dinâmicas e as distorções associadas que não são, em geral, levadas em conta. Sensores são, também, sempre corruptidos, no seu desempenho, por ruídos. Finalmente, sensores nem sempre fornecem dados sobre todas as variáveis de interesse, seja porque não se consegue construir dispositivos que efectuem determinadas medidas, seja porque o custo de sensores para outras medidas seria proibitivo.

Em vista disso, admitir um modelo determinístico (perfeito) é, na regra, uma hipótese não adequada. Esse é o motivo para se trabalhar com modelos estocásticos. No entanto, para levar em conta os modelos estocásticos, devemos estar preparados para responder satisfatoriamente a quatro quesitos:

- a) Como desenvolver modelos que levem em conta as incertezas da natureza prática;
- b) Com esses modelos e com dados incompletos e corruptidos por ruído dos sensores, como estima de maneira otimizada as quantidades de interesse;
- c) Em face dos modelos incertos, dos dados não perfeitos e das perturbações sobre o controle, como controlar otimamente o sistema;
- d) Como avaliar os sistemas de estimativa e controle antes e depois de sua construção real.

## 0.2. Sistemas Totalmente Determinístico, Observador e Filtro Estocástico

Consideremos as duas configurações abaixo. A primeira não é uma melhoria fechada no sentido de que a estrutura do estado não é usada para melhorar outra função que não é de fornecer uma melhor conhecimento do sistema. Na segunda configuração, a estrutura do estado é fornecida ao controlador, fechando a melhoria. No caso de sistemas lineares a configuração da Fig. 2 é ótima se o controlador é projetado segundo um critério quadrático e se o filtro é um filtro de Kalman.

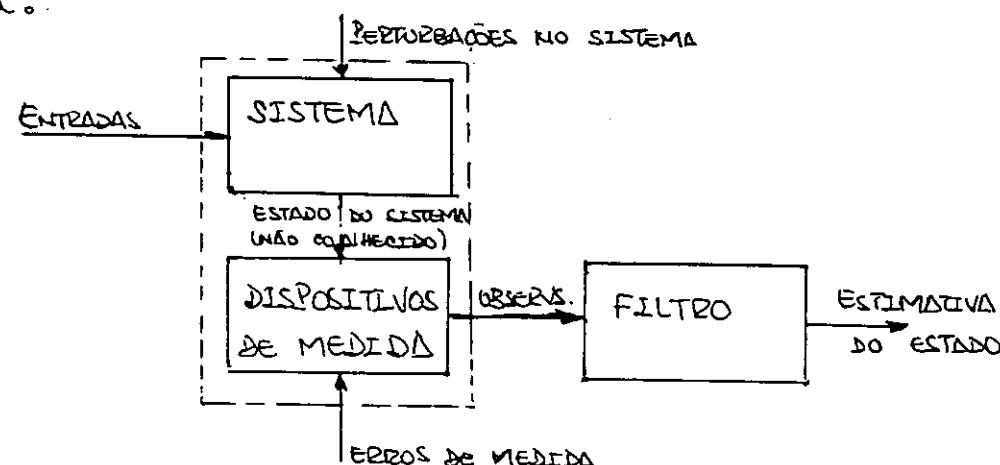


Fig. 1.

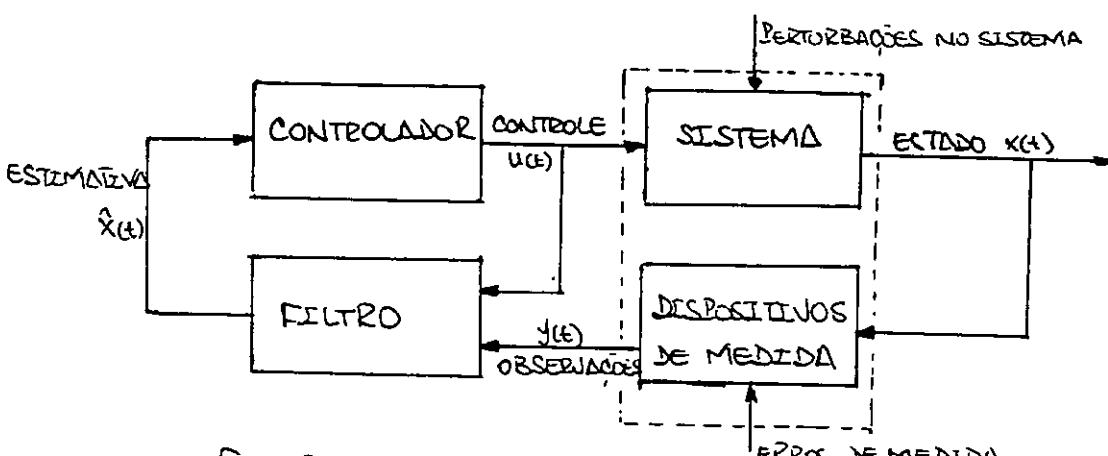


Fig. 2.

Imaginemos uma situação em que o estado do sistema (vetor) tenha de estar disponível para fins de controle ou quais-

outro fato. Podemos ter uma das três seguintes situações:

a) Sobre o sistema não existem, ou são desprezíveis, efeitos de perturbação no sistema ou de estes não medidas. Isto significa que factores externos à evolução do sistema real têm pouco efeito e o modelo matemático reproduz bastante bem esse evolução e que os sensores são suficientemente bons para não introduzirem erros nas medidas. Admitimos ainda que todas as variáveis de estado possam ser medidas pelos sensores.

Nesse caso, nos esquemas das Figs. 1 e 2, a presença do filtro é dispensável e dizemos que o sistema é TOTALMENTE DETERMINÍSTICO.

b) Admitimos que sobre o sistema existam várias perturbações com força influência sobre sua evolução, mas que nem todas as variáveis de estado possam ser medidas pelos sensores. Para reconstruir o estado, torna-se imprescindível a colocação de um filtro para fazer a estimativa. Nesta situação colocada, devido ao ruído, a componente eletrônica pode ser descartada e continuarmos a trabalhar com o sistema determinístico. Um caso de aplicação em que ocorre esse tipo de situação é o controle estrutural, onde se tenta localizar e controlar os modos de vibração de uma estrutura mecânica.

Um filtro determinístico é normalmente chamado de OBSERVADOR. Existem várias maneiras de se projetar um observador, sendo que a técnica mais difundida é a do OBSERVADOR DE Luenberger.

c) A terceira e mais complicada situação é aquela em que os ruidos do sistema e das observações realmente influem sobre o sistema, independentemente do número de medidas que nossos sensores possam realizar.

Nesse caso, é fundamental separar o efeito eletrônico da evolução do sistema. Esta separação pode ser feita de várias maneiras, como veremos ao longo destas notas. Para sistemas lineares, ou mesmo para sistemas não-lineares que possam ser representados

mente approximadas por sistemas lineares, os chamados FILTROS DE KALMAN são as técnicas mais difundidas.

### 0.3. FILTRO DE KALMAN

O Filtro de Kalman pode ser classificado como um "algoritmo computacional recursivo ótimo e de mínimo erro". Tentarmos, agora, traduzir essa classificação. O Filtro de Kalman é um algoritmo computacional que tem como entradas todas as medidas disponíveis sobre o sistema e como saídas as componentes do vetor de estados do sistema em questão. Tal tipo de filtro é recursivo no sentido de que não há necessidade de se armazenar todas as informações recebidas, isto é, o algoritmo trabalha com as últimas estruturas e com as informações mais recentes, apenas. O Filtro de Kalman é ótimo no sentido de que qualquer critério de otimização fornece o mesmo algoritmo, dentro dos critérios razoáveis em Engenharia. Por exemplo, em sistemas lineares com distribuições normais (gaussianas), o Filtro de Kalman fornece a mesma estrutura seja o critério o de média, de moda ou de mediana. Finalmente, é de mínimo erro no sentido de que a incerteza no final do processamento é a menor entre todas as outras geradas por outras aproximações.

- Para sua realização, é necessária a utilização de:
- conhecimento da dinâmica do sistema e do processo de observações;
  - estatística conhecida das mudanças (incertezas) sobre o sistema e erros das medidas;
  - informações sobre o estado inicial do sistema.

O que Kalman colocou de uma forma sistemática e rigorosa há vinte anos atrás, seu absolutamente tentomostinar seu mérito, foi entendido por Gauss, um gênio maior, há cerca de cento e oitenta anos. Gauss, em 1809, dizia a este-

ção para:

- a) a necessidade de dados redundantes para se eliminar os erros nas observações;
- b) a necessidade de estudos exustivos sobre os sistemas dinâmicos em estudo;
- c) o tratamento da imprecisão das medidas com considerações probabilísticas (lembre que Teoria de Probabilidade de 1935);
- d) a geração de uma "combinação desejada" das observações para produzir melhores estimativas, questões hoje chamadas de "estrutura do filtro" e "definição do critério de otimização";
- e) a necessidade de se estabelecer um número mínimo de observações para se poder filtrar, algo que hoje se chama "observabilidade do sistema".

Finalmente, para encerrar essa introdução, devemos ressaltar a forma como que serão desenvolvidas aqui teoria e prática do filtro de Kalman. Escolhemos, sem dúvida, a forma mais demorada com variação de alguns tópicos antigas e introdução de outros tópicos novos para os interessados de modo a que, ao final, todos possam ser capazes de criticar os defeitos e ressaltar as virtudes de um estimador de estado e, ainda, se possível, sugerir modificações e aperfeiçoamentos a serem introduzidos. A intenção é, então, construir gradual e lentamente a Teoria de Kalman, para todos puderem seguir sozinhos depois. Seria muito mais fácil e menos cansativo tentar o caminho inverso, isto é, montar o filtro e depois discutir um ou outro detalhe do nosso ponto de vista de engenheiros, esse segundo caminho não fornece o embasamento suficiente para que todos os aspectos sejam percebidos e digeridas e, portanto, o início estará muito próximo do fim.

#### 0.4. SOBRE O CURSO

O curso "Filtro de Kalman aplicado à Navegação Aérea" abordará os seguintes tópicos:

##### 4.1. Álgebra Linear

Intenção: revisão dos conceitos básicos, desenvolvimento de algumas identidades matriciais de uso posterior (teoria); preparação de pacotes para inversão de matrizes (prática)

##### 4.2. Equações Diferenciais Ordinárias

Intenção: revisão dos conceitos básicos, matrizes de transição de sistemas lineares (teoria); preparação de pacotes para integração numérica e gerador de matrizes de transição (prática)

##### 4.3. Observadores de Estado

Intenção: projetos de observadores de estado: trivial, identidade, Luenberger (teoria); projeto e implantação de um observador (prática).

##### 4.4. Teoria de Probabilidade

Intenção: revisão dos conceitos fundamentais (teoria)

##### 4.5. Estimação de Parâmetros

Intenção: critérios de otimização para um estimador, mínimos quadrados, forma direta, forma de Kalman (teoria); preparação de um pacote para um gerador de observações (prática)

##### 4.6. Processos Estocásticos

Intenção: definição dos conceitos fundamentais (teoria)

#### 4.7. Filtro de Kalman para sistemas lineares

Intenção: juntar todos os itens já desenvolvidos para a versão mais simples do filtro de Kalman (teóric); preparação do programa final de aplicação (prática).

#### 4.8. Aplicações da Teoria linear aos sistemas não lineares

Intenção: técnicas importantes nas aplicações de Engenharia (teóric); desenvolvimento do programa final de aplicação (prática).

Bibliografia a ser utilizada:

- 1) Mzybeck, P.: "Stochastic Models, Estimation and Control", Academic, 1979. (um livro no nível que o curso preende abranger).
- 2) Liebelt, P.L.: "An Introduction to Optimal Estimation", Addison-Wesley, 1967 (excelente material para os itens 4.1, 4.2 e 4.4; bons exercícios e exemplos de aplicação)
- 3) Jazwinski, A.H.: "Stochastic Processes and Filtering Theory", Academic, 1970 (um livro excelente mas muito pesado, ao rigor matemático, para um primeiro curso)
- 4) Gelfand, A. (editor): "Applied Optimal Estimation", MIT Press, 1974 (excelentes aplicações, um livro para quem deseja implantar filtros de Kalman).
- 5) Bryson, Jr, A.E.; Ho, Y.C.: "Applied Optimal Control", Wiley, 1975, 2nd edition (livro de referência)
- 6) Papoulis, A.: "Probability, Random Variables and Stochastic Processes", McGraw-Hill, 1965 (excelente livro sobre probabilidade e processos estocásticos)
- 7) Luenberger, D.G.: "Introduction to Dynamic Systems: Theory, Models and Applications", Wiley, 1979 (fundamental para observadores e sistemas determinísticos).

Controle ótimo de sistemas lineares com critério

Quadrático

sistema linear

$$\dot{x} = Ax + Bu$$

modelos dinâmicos

$$y = Cx$$

modelos de observação

critério: Minimizar

$$J^D = \frac{1}{2} \bar{x}(t_f)^\top \bar{Q} \bar{x}(t_f) + \frac{1}{2} \int_u^{t_f} [x^\top W_1 x + 2x^\top W_2 u + u^\top W_3 u] dt$$

Vinculo de controles

$$\bar{\Psi} x(t_f) = \bar{U}_0$$

$$F = \frac{1}{2} [x^\top W_1 x + 2x^\top W_2 u + u^\top W_3 u]^\top dt$$

cond. Necessárias

- Eqs Euler-Lagrange ou adjuntas ou de c. estados

$$\lambda = -\frac{\partial F}{\partial x}$$

$$\dot{\lambda} = -A^\top \lambda + W_1 x + W_2 u$$

$$- Eqs. Compte \quad \frac{\partial F}{\partial u} = 0$$

$$B^\top \lambda - W_2^\top x - W_3 u = 0$$

- Eqs. Movimentos

$$\dot{x} = Ax + Bu$$

- cond. Transversalidade

$$\lambda(t_f) = -\frac{\partial G}{\partial x(t_f)}$$

$$G = \frac{1}{2} \bar{x}(t_f)^\top \bar{Q} \bar{x}(t_f) + \frac{1}{2} \bar{\Psi} x(t_f)^\top \bar{U}$$

$$\lambda(t_f) = -\bar{Q} \bar{x}(t_f) - \bar{U}^{-1}$$

Eliminando  $u$

$$\begin{bmatrix} \dot{x} \\ \ddot{x} \end{bmatrix} = \begin{bmatrix} A - BW_3^{-1}W_2^T & BW_3^{-1}BT \\ W_1 - W_2W_3^{-1}W_2^T & -A^T + W_2W_3^{-1}B^T \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \quad \text{linear}$$

Solução, usando as cond. conforme

LEI DE CONTROLE

$$u^*(x,t) = -W_2^{-1}B^T R Q^{-1} \psi_0 - W_3^{-1} \{ W_2^T + B^T [P - Q^{-1}R^T] \} x(t)$$

$$\text{onde } \dot{P} = -PA - A^T P - W_1 + (PB + W_2)W_3^{-1}(W_2^T + B^T P); P(t_f) = \bar{P}$$

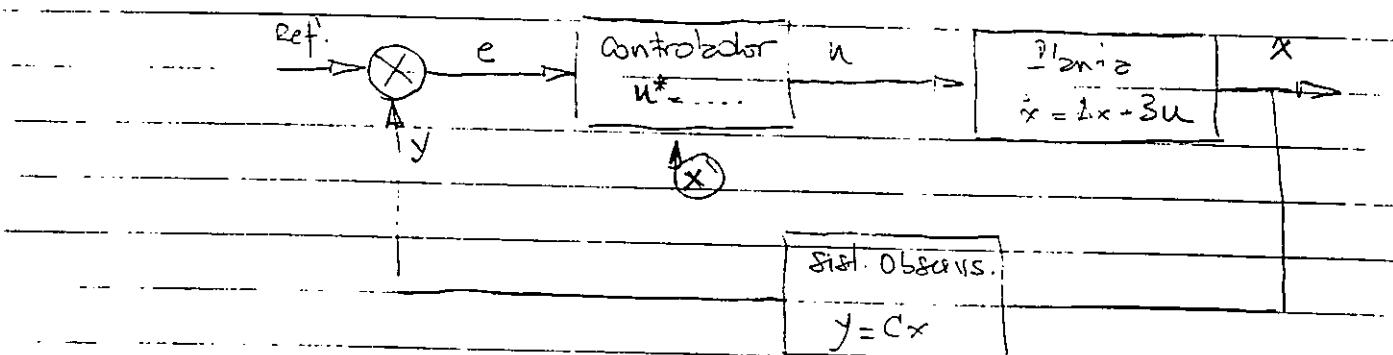
$$\dot{R} = (PBW_3^{-1}B^T - A^T + W_2W_3^{-1}3^T)R \quad ; R(t_f) = \bar{Q}$$

$$\dot{Q} = R\bar{3}W_3^{-1}B^TR \quad ; Q(0) = 0$$

Desvantagens

① A lei ótima de controle depende de conhecer, a cada instante, todo o vetor de estados  $x(t)$

② O sistema é sempre integrado de  $t_f$ , instante final, para  $t=0$ , instante inicial (backwards)

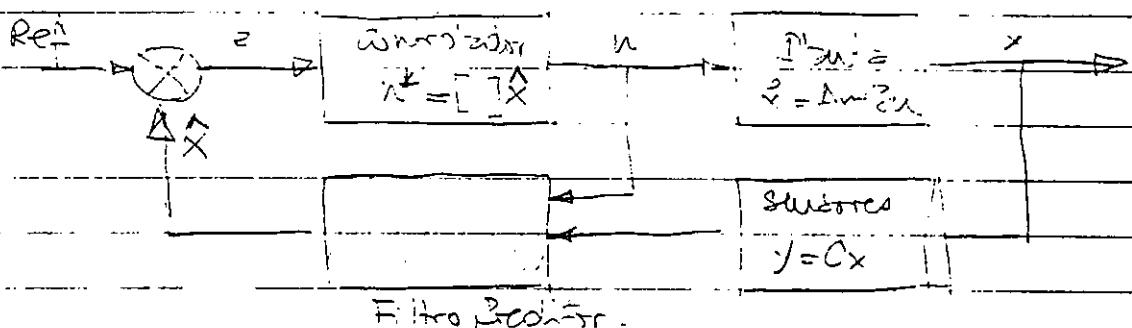


Problema. Conhecer o estado  $x$  completo!!!

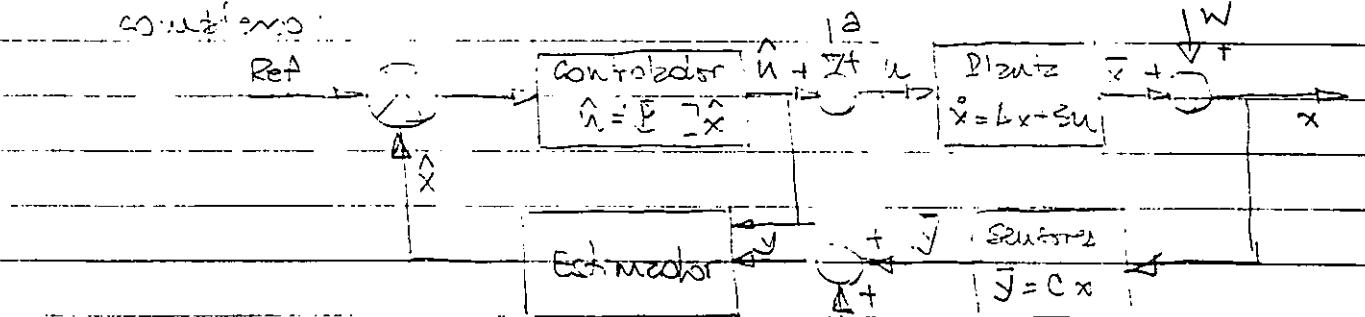
Para um controlador LQ, precisamos de um FILTRO se não conhecemos todo o estado.

- Dadas as observações  $O(t_k) = \{y(t_0), y(t_1), \dots, y(t_{k-1})\}$ , gerar uma estimativa  $\hat{x}(t_k)$  para alimentar o controlador  $\Rightarrow$  FILTRO ou Problema de Filtragem
- Dadas o mesmo conjunto de observações  $O(t_k)$ , gerar uma estimativa  $\hat{x}(t_e)$ ,  $t_e > t_k$ , para gerar dados de saída futuros  $\Rightarrow$  PREDICTOR ou Prob. Previsão.
- Dadas unidas o conjunto  $O(t_k)$ , obter uma estimativa  $\hat{x}(t_k)$ ,  $t_k < t_e$ , para saber com devoerá ser sido realmente o sistema  $\Rightarrow$  ALIÍGN.

Para que haja um controle eficiente, é necessário separar em mais uma camada o diagrama de blocos:



Para que haja estabilidade real, o controle é um dos critérios:



$$\dot{x} = Ax + Bu + w$$

$$y = Cx + v$$

w e v não são modelados

a é um gerador incluído na planta

## MODELOS

Todo controlador é tão bom quanto bom é o modelo que representa a dinâmica do sistema que se pretende controlar (Friedland, 87; Eykhoff, 94). Em outras palavras:

Modelo Bom  $\Rightarrow$  Possíveis bons controladores

Modelo Ruim  $\Rightarrow$  Controladores Ruins

### O QUE É MODELO?

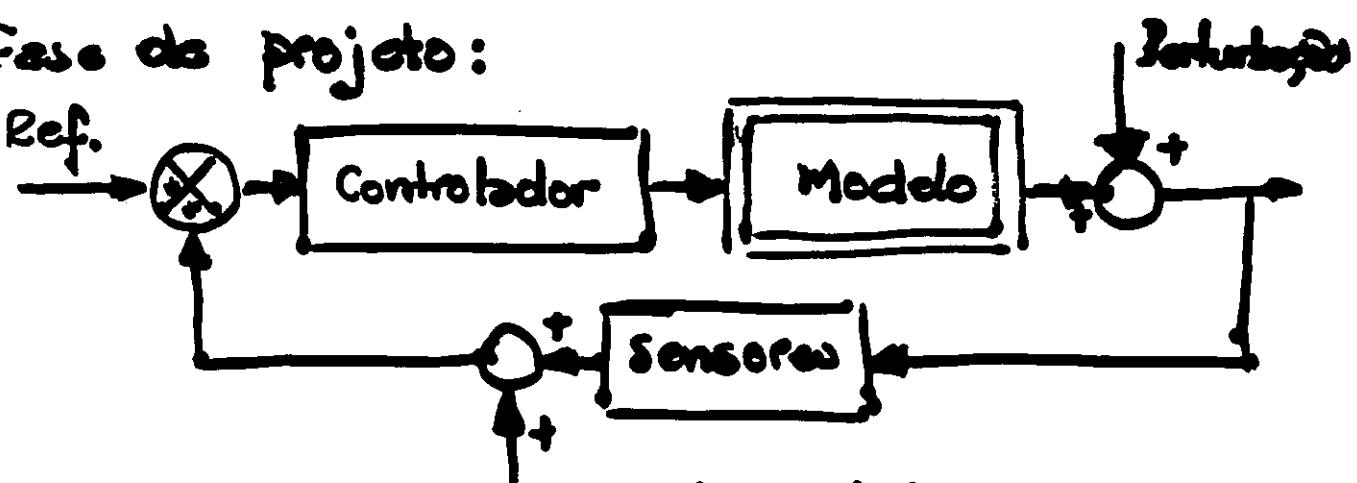
Síntese de todo conhecimento que se dispõe sobre o fenômeno ou sistema em estudo  
(Modelo físico)

- Equações diferenciais de movimento;
- Funções de transferência;
- Redes Neurais treinadas;
- Conjuntos de regras

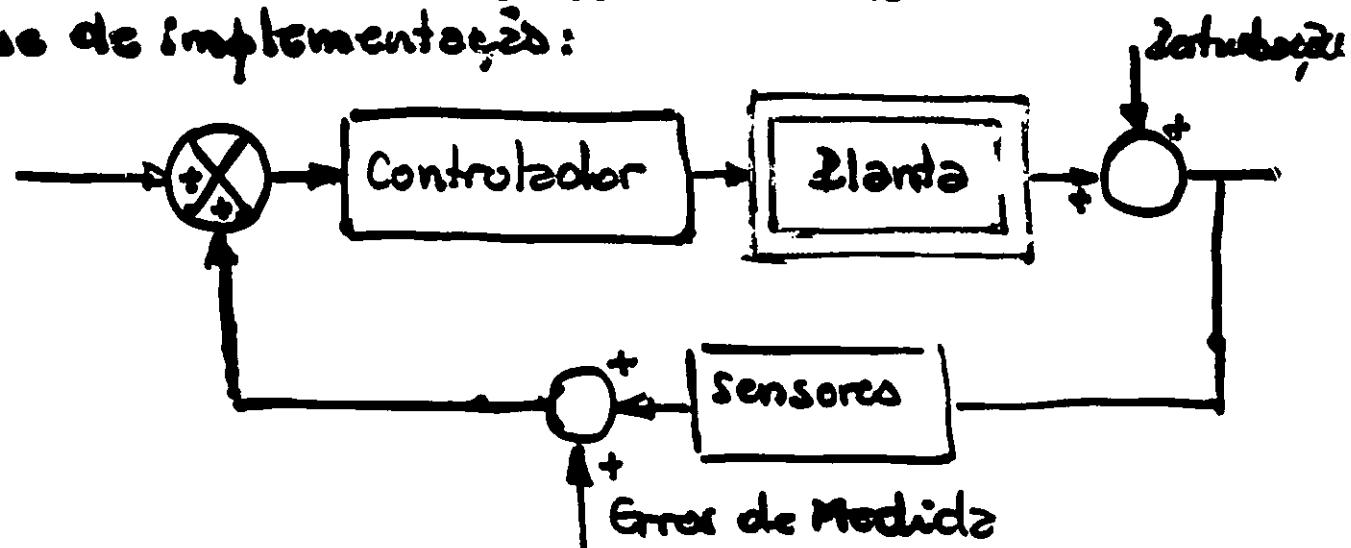
não são modelos(físicos). São a representação matemática ou lógica do modelo (físico).

Esse distinção é vital em Controle:

Fase de projeto:

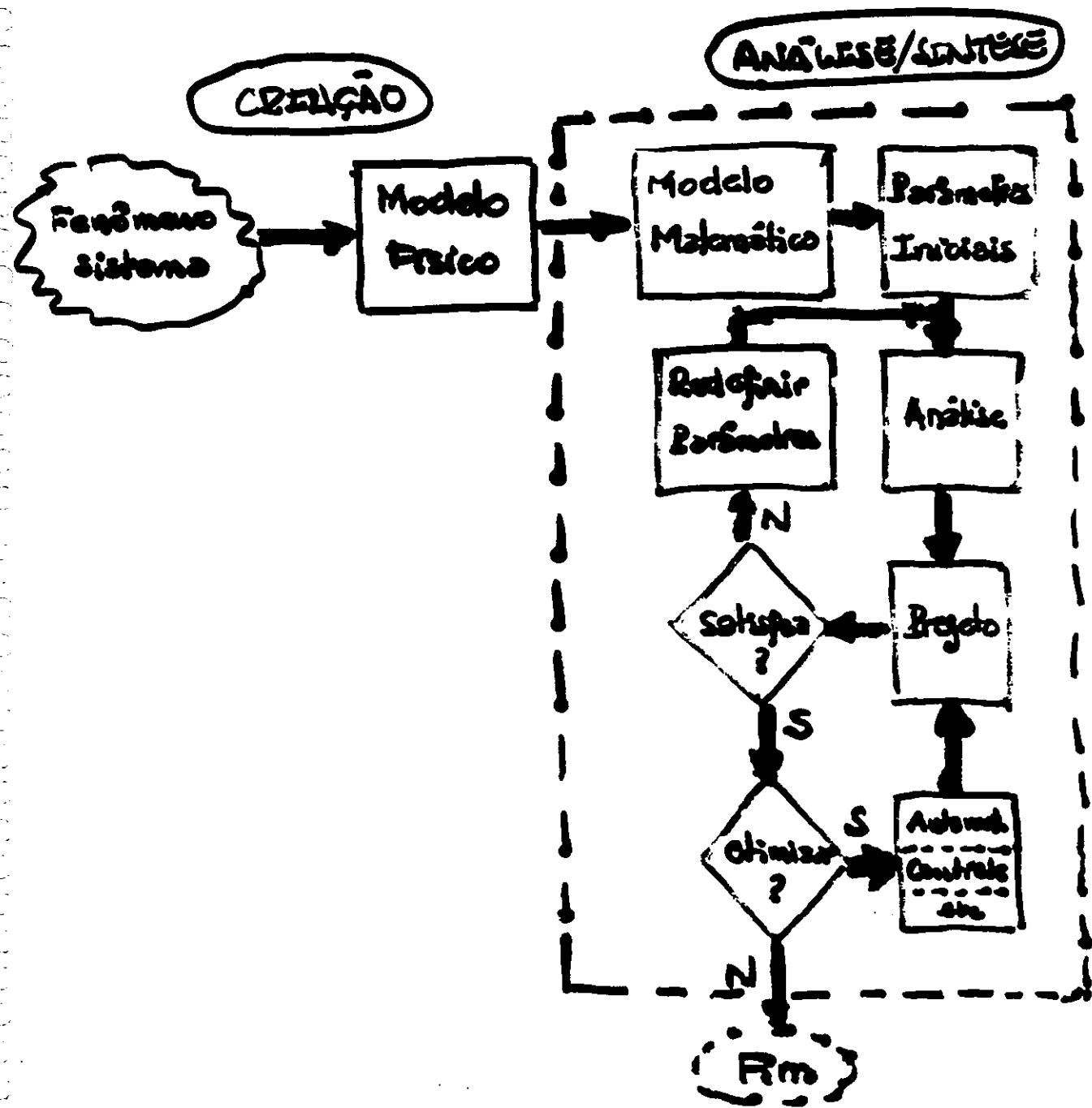


Fase de implementação:



O controlador implementado continua a encarregar o modelo da fase de projeto e não a planta. Se o modelo não for idêntico à planta não há o que fazer para conseguir convergência. Sofisticar a estratégia de controle é, em geral, inútil, a menos que haja algoritmo de adaptação escopulado.

Modelar é uma arte (Engenharia)

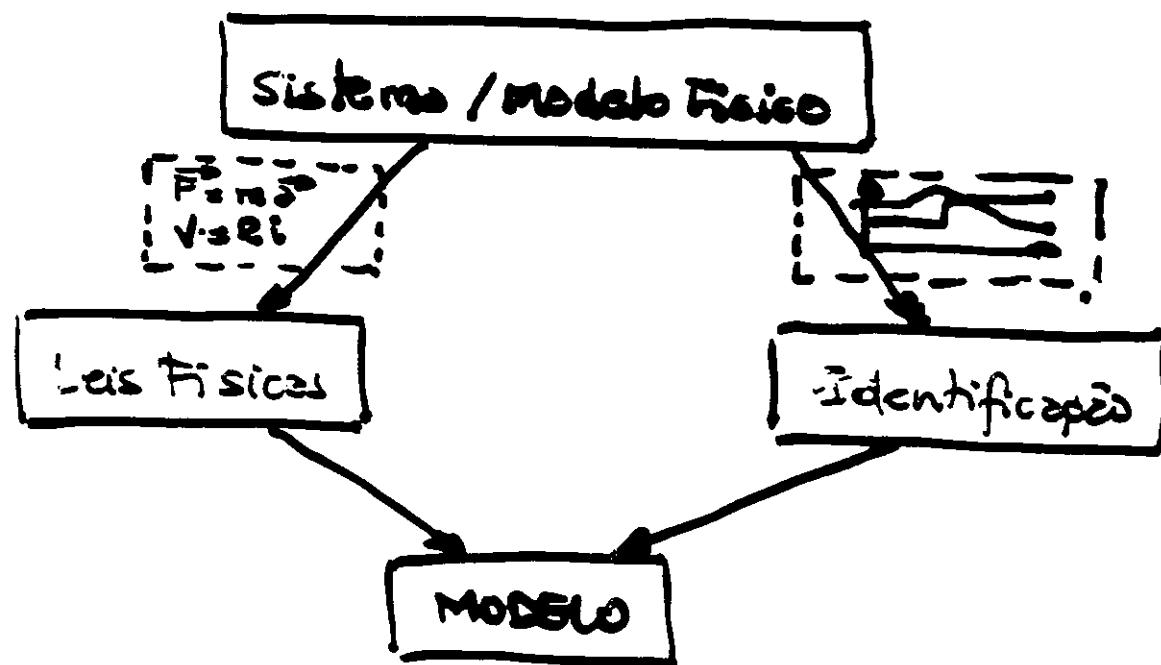


Exemplos:

- Projeto AMX (Embraer)
- Carro frenando

Degui para a frente, vamos nos atter à discussão dos modelos matemáticos.

## Construção do Modelo



## Tipos de Modelos Matemáticos

- Estáticos x Dinâmicos
  - **Contínuos** • **Discretos**
    - Tempo
    - Vari. Estado
    - Eqs. Diferenciais
    - Frequências
    - Cosseno
    - Pos. Transf.
    - Tímido
    - Vari. Estado
    - Eqs. Diferenciais
    - Frequência
    - Amostradas
    - Transf. Z
  - Lineares • Não Lineares
  - Parâmetros Concentrados (EDO) x Parâmetros Distribuídos (EDDP)
  - Determinísticos x Estocásticos
  - Eventos Discretos x Mudanças Orientadas

## 2.3. PARTES DE UM SISTEMA DE CONTROLE

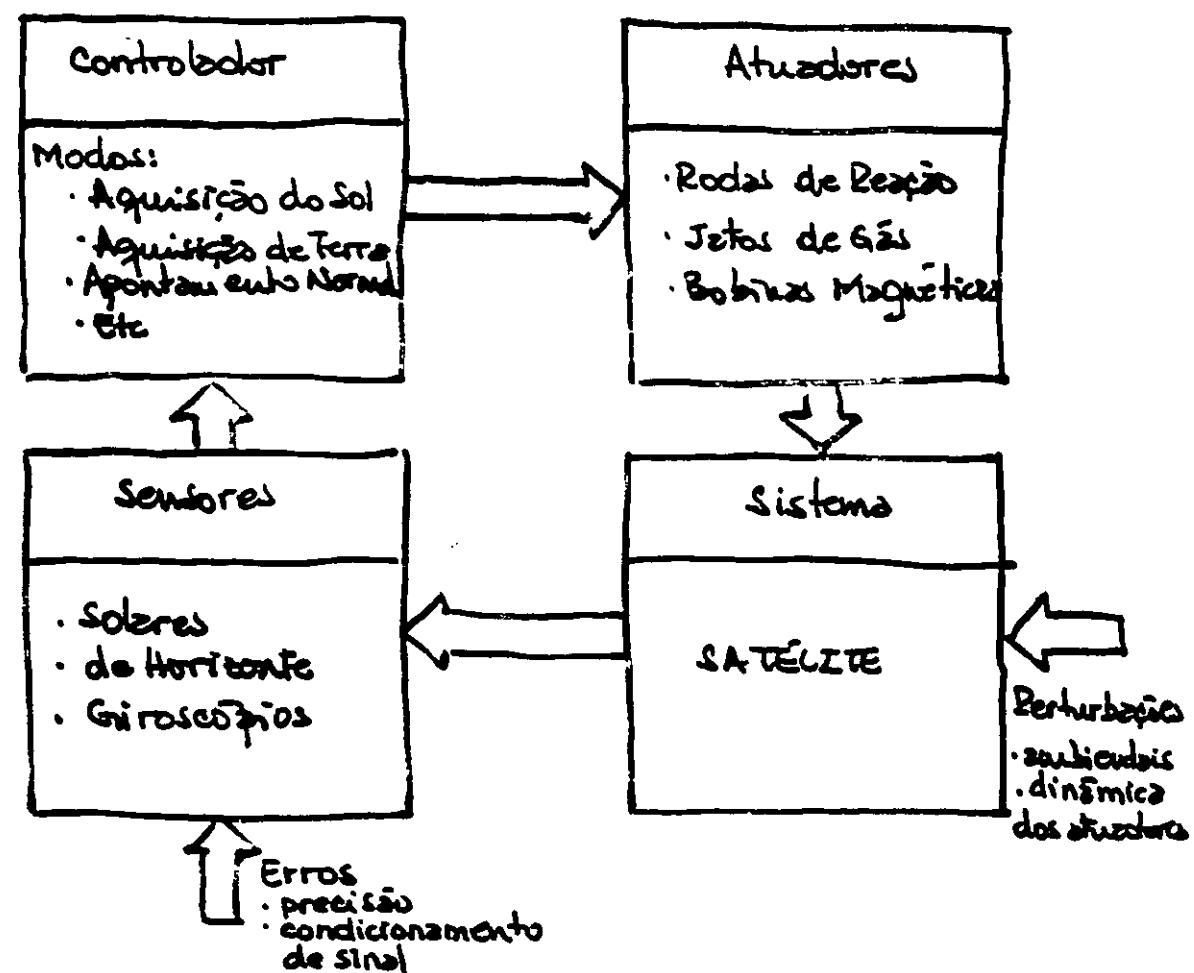
Sistema ou Planta

Sensores

Lógica de Controle

Atuadores

Exemplo: Sistema de Controle de Altitude de um Satélite  
em 3 eixos



## 2.4. CONTROLE CLÁSSICO E CONTROLE MODERNO

### MODERNO

#### Representação

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = x_0 \\ y(t) &= Cx(t)\end{aligned}$$

- Descrição da dinâmica interna
- Determina uma única Fcêo.
- Transferência
- Recomendado:
- \* Sistemas de múltiplas entradas e múltiplas saídas (MIMO)
- \* Entradas são gerais
- \* Otimizações de desempenho
- \* Incluir cond. iniciais no estudo

### CLÁSSICO

#### Representação

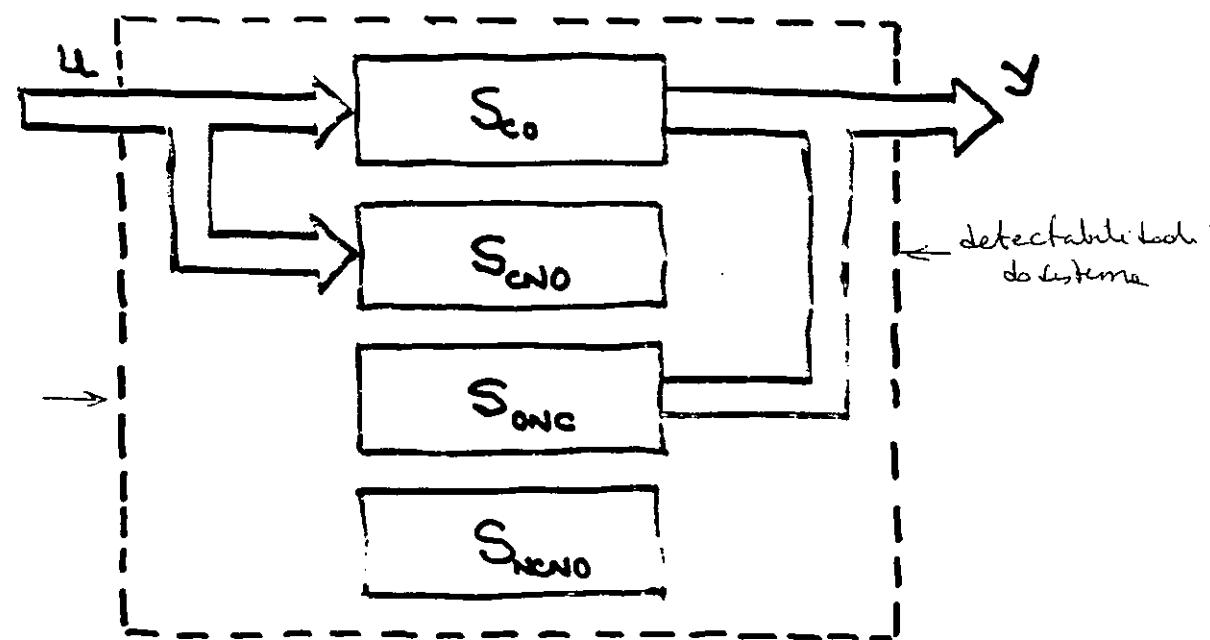
$$Y(s) = G(s) U(s)$$

- Não há descrição da dinâmica
- Dada  $G(s)$ , Ex. infinitas representações de estado
- Recomendado:
- \* Sistemas com uma entrada e uma saída (SISO)
- \* Entradas tipo impulsivas, degrau e senoides

### 3. CONTROLAIBILIDADE E OBSERVABILIDADE

Controlabilidade: habilidade do controle (entrada) agir sobre cada uma das variáveis de estado

Observabilidade: habilidade de cada variável de estado influenciar a saída do sistema



Função de transferência  $y(s) = G(s)U(s)$ ;  $G(s)$  depende apenas de  $S_{00}$

$$\text{sistema} \quad \dot{x} = Ax + Bu$$

$$y = Cx$$

a) Controlável se

$$[B | AB | A^2B | \dots | A^{n-1}B] \text{ for de posto } n$$

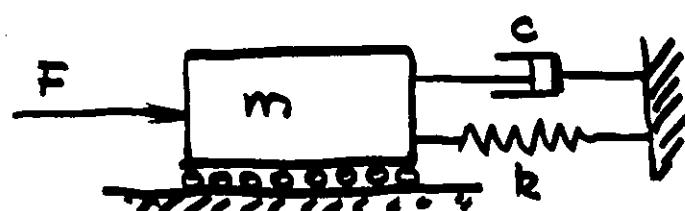
b) Observável se

$$[C^T | A^TC^T | (A^T)^2C^T | \dots | (A^T)^{n-1}C^T] \text{ for de posto } n$$

## ESPAÇO DE ESTADOS

- Interesses: sistemas contínuos, de parâmetros concentrados, determinísticos ou estocásticos.
- Sistemas contínuos, de parâmetros concentrados são descritos por EDO's
- Na abordagem por espaço de estados, todas as EDO's do modelo são reduzidas a ODE's de 1ª ordem.
- Variáveis dinâmicas que operam no sistema de 1ª ordem: VARIÁVEIS DE ESTADO
- A ordem do sistema  $\Rightarrow$  nº de eqs. de 1ª ordem
- São necessárias tantas cond. iniciais quanto a ordem do sistema.

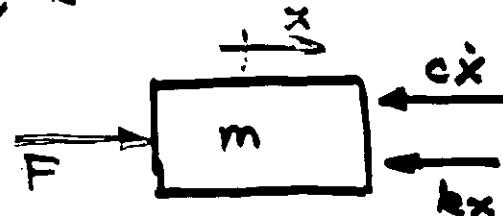
Ex.:



Lei de Newton:

$$m\ddot{x} = F - cx - kx$$

$$m\ddot{x} + cx + kx = F \quad (1)$$



Representação em espaço de estado:

• define-se  $x_1 = x$  (posição)

•  $\frac{dx}{dt} = x_2$  (velocidade)  $\Rightarrow \dot{x}_1 = x_2 \quad (2)$

$$\text{em (1): } m\ddot{x}_2 + cx_2 + kx_1 = F \quad (3)$$

(2) e (3):

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -\frac{k}{m}x_2 - \frac{b}{m}x_1 + F/m$$

ou:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{b}{m} & -\frac{k}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ F/m \end{bmatrix}$$

Se admitirmos  $F$  - força de controle

$$F = u(t)$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ u \end{bmatrix}$$

$$\dot{x}(t) = A x(t) + B u(t)$$

Para completar é preciso decidir como o sistema é observado (sensores). As medidas ( $y(t)$ ) são agrupadas em um vetor  $y(t)$  e estabelece-se a relação entre os nsa observados e os variáveis de estado (e, eventualmente, de controle)

$$y(t) = C x(t) (+ D u(t))$$

No exemplo, supondo que se mede:

• só a posição:  $y(t) = [1 \ 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

• só a velocidade:  $y(t) = [0 \ 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

• a posição e a velocidade simultaneamente:

$$y(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Em resumo, representação em variáveis de estado ou no espaço de estados:

• Sistema linear:

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t) && \text{mod. dinâm.} \\ y(t) &= Cx(t) && \text{(observ.) mod. obs.}\end{aligned}$$

$x$  - vetor ( $n \times 1$ ) de variáveis de estado

$u$  - vetor ( $m \times 1$ ) de variáveis de controle

$y$  - vetor ( $p \times 1$ ) de variáveis observadas

$A$  é  $n \times n$ ;  $B$  é  $n \times m$ ;  $C$  é  $p \times n$

• Sistema não linear

$$\begin{aligned}\dot{x}(t) &= f(x(t), u(t), t) && \text{mod. dinâmico} \\ y(t) &= g(x(t), t) && \text{mod. observ.}\end{aligned}$$

$x$  - vetor ( $n \times 1$ ) de variáveis de estado

$u$  - vetor ( $m \times 1$ ) de variáveis de controle

$y$  - vetor ( $p \times 1$ ) de variáveis observadas

$f$  é um vetor  $n \times 1$  de funções não lineares

$g$  é um vetor mal de funções não lineares.

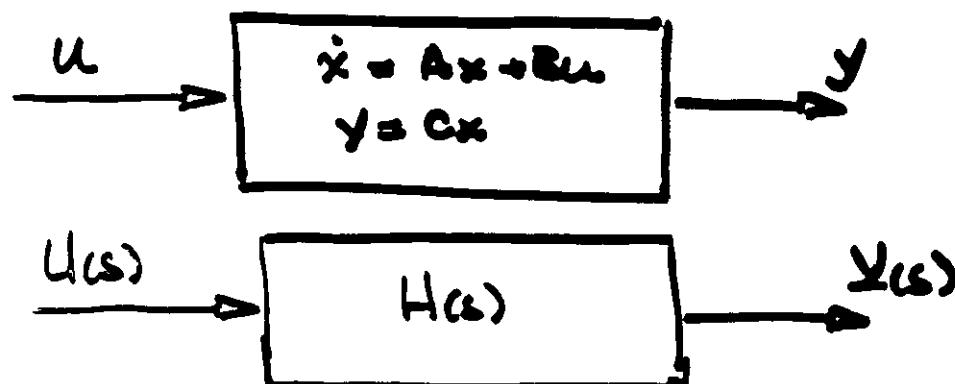
Variáveis de estado definem completamente a resposta de um sistema em qualquer instante

Variáveis de controle representam o esforço de agir sobre o sistema.

Variáveis observadas representam o esforço de medir o sistema.

## MODELOS ENTRADA - SAÍDA

### Funções de transferência



$$Y(s) = H(s) U(s)$$

$$H(s) = C(sI - A)^{-1} B \quad (+\delta)$$

Sistemas múltiplas entradas - múltiplas saídas (MEMO):

$$A(n \times n), B(n \times m), C(p \times n) \Rightarrow H(s) \in (p \times m)$$

Sistemas uma entrada - uma saída (SISO):

$$A(n \times n), B(1 \times n), C(1 \times n) \Rightarrow H(s) \text{ é escalar}$$

Fgôes transferência mescaram a dinâmica interna: dados  $(A, B, C)$  existe uma única matriz  $H(s)$ ; dados  $H(s)$  existem os termos  $(A, B, C)$  que compõe  $H(s)$ .

## MATRIZ DE TRANSIÇÃO DE ESTADOS

Consider o sistema:

$$\dot{x}^o = Ax + Bu$$

onde  $x \in \mathbb{N}^k$  e  $u \in \mathbb{M}^{k \times l}$

Como é uma EDO, procuramos uma solução geral como a soma de uma solução particular com a solução da homogênea. Começamos com a homogênea:

$$\dot{x} = Ax$$

e a solução é escrita primeiramente como uma série infinita com coeficientes não determinados:

$$x(t) = \underline{\phi}(t)x(0) = (I + c_1 t + c_2 t^2 + \dots + c_n t^n + \dots) x(0) \quad (1)$$

onde  $\Phi(t)$  é a solução e  $x(0)$  a condição inicial

Para determinar os coeficientes, substituiras (1) na equação homogênea:

$$\frac{d}{dt} \Phi(t) x(0) = (C_1 + 2C_2 t + 3C_3 t^2 + \dots + n C_n t^{n-1} + \dots) x(0) \quad \leftarrow ①$$

$$= A \Phi(t) X(0) = (A + A C_1 t + A C_2 t^2 + \dots + A C_n t^n + \dots) X(0)$$

Igualando os coeficientes:

$$C_1 = \Delta$$

$$C_2 = \frac{1}{2} AC_1 = \frac{1}{2} A^2$$

$$C_3 = \frac{1}{3} A C_2 = \frac{1}{3 \cdot 2} A^3, \dots, C_n = \frac{1}{n!} A^n$$

D21

$$\hat{\phi}(t) = e^{At} = I + At + \frac{1}{2!} A^2 t^2 + \frac{1}{3!} A^3 t^3 + \dots + \frac{1}{n!} A^n t^n + \dots \quad (2)$$

$\hat{Q}(t) = e^{At}$  — Matriz de Transição de Estados

Propriedades:

$$\text{I) } \frac{d}{dt} \underline{\Phi}(t) = \frac{d}{dt} e^{At} = A e^{At} = e^{At} A \quad (3)$$

$$\dot{\underline{\Phi}} = A \underline{\Phi}$$

$$\text{II) } x(t+\tau) = \underline{\Phi}(t) x(0) \quad (4)$$

$$x(\tau) = \underline{\Phi}(\tau) x(0) \quad (5)$$

$$x(t+\tau) = \underline{\Phi}(t) \underline{\Phi}(\tau) x(0) \quad (6)$$

$$\Rightarrow \underline{\Phi}(t+\tau) = \underline{\Phi}(t) \underline{\Phi}(\tau) \text{ ou } e^{A(t+\tau)} = e^{At} \cdot e^{A\tau} \quad (7)$$

$$\text{III) } \tau = -t \Rightarrow \underline{\Phi}(0) = \underline{\Phi}(t) \underline{\Phi}(-t) = e^{A \cdot 0} = I$$

$$\underline{\Phi}(-t) = \underline{\Phi}^{-1}(t) \quad (8)$$

$$\text{IV) } \dot{\underline{\Phi}} = A \underline{\Phi} \text{ é a matriz da solução com } \underline{\Phi}(0) = I$$

$$\text{V) } \underline{\Phi}(t) \dot{\underline{\Phi}}(\tau) = \underline{\Phi}(\tau) \dot{\underline{\Phi}}(t) \quad (9)$$

Consideremos agora a solução particular de

$$\dot{x} = Ax + Bu$$

Também:  $x_p(t) = \underline{\Phi}(t) p(t) = e^{At} p(t)$ . devido  
com  $p(t)$  descontinuado e tal que  $p(t)=0$  para  $t=0$

$$\dot{x}_p = \left( \frac{d}{dt} \underline{\Phi}(t) \right) p(t) + \underline{\Phi}(t) \left( \frac{d}{dt} p(t) \right)$$

$$\dot{x}_p = \widehat{A \underline{\Phi}(t)} p(t) + \underline{\Phi}(t) \frac{d}{dt} p(t) \quad (10)$$

$$\text{Mas } \dot{x}_p = A x_p + Bu$$

$$\text{De (10) e (11) } \Rightarrow \underline{\Phi}(t) \frac{d}{dt} p(t) = Bu$$

$$\frac{d}{dt} p(t) = \underline{\Phi}^{-1}(t) Bu$$

$$\Rightarrow p(t) = \int_0^t \underline{\Phi}^{-1}(s) Bu(s) ds \quad (12)$$

$$\text{Com isso: } x_p(t) = \underline{\Phi}(t) \int_0^t \underline{\Phi}^{-1}(\tau) B u(\tau) d\tau$$

$$- \int_0^t \underline{\Phi}(t) \underline{\Phi}^{-1}(\tau) B u(\tau) d\tau \quad \text{(8)}$$

Como já vimos em (8)

$$\underline{\Phi}(t) \underline{\Phi}^{-1}(\tau) = \underline{\Phi}(t-\tau)$$

e a solução particular fica:

$$x_p(t) = \int_0^t \underline{\Phi}(t-\tau) B u(\tau) d\tau$$

e a solução geral da E.D.O.:

$$x(t) = \underline{\Phi}(t) x(0) + \int_0^t \underline{\Phi}(t-\tau) B u(\tau) d\tau$$

## I. ÁLGEBRA LINEAR : DEVISÃO

### II. Definição de uma Matriz

Matriz é um conjunto de  $m \times n$  quantidades dispostas num arranjo retangular de  $m$  linhas e  $n$  colunas.

$$A = [A] = [z_{ij}] = \begin{bmatrix} z_{11} & z_{12} & \cdots & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & \cdots & z_{2n} \\ \vdots & \vdots & & & \vdots \\ z_{m1} & z_{m2} & \cdots & \cdots & z_{mn} \end{bmatrix}$$

$A$  - matriz  $m \times n$

II.a Se  $m=n$ , então  $A$  é uma matriz quadrada

II.b. Se  $m=1$ , então  $A$  é um vetor, ou vetor linha ou matriz linha

$$\{A\} = a = \{z_{11}; \dots, z_{1n}\}$$

II.c. Se  $n=1$ , então  $A$  é um vetor, ou vetor coluna ou matriz coluna

$$\{A\} = a = \begin{Bmatrix} z_{11} \\ \vdots \\ z_{m1} \end{Bmatrix}$$

### I.2. Igualdade de Matrizes

$$[A] = [B] \iff z_{ij} = b_{ij}, \forall i=1, \dots, m \\ \forall j=1, \dots, n$$

### I.3. Adição de Matrizes

$$[C] = [A] + [B] \iff c_{ij} = z_{ij} + b_{ij}, \forall i=1, \dots, m \\ \forall j=1, \dots, n$$

Propriedades da Adição:

1.3.a. Comutativa :  $A + B = B + A$

1.3.b. Associativa :  $A + (B + C) = (A + B) + C$

1.4. Multiplicação por um escalar  
Seja  $\alpha \in \mathbb{R}$

$$\alpha [A] = [B] \iff b_{ij} = \alpha a_{ij} \quad \begin{matrix} i=1, \dots, m \\ j=1, \dots, n \end{matrix}$$

Propriedades da Multiplicação por Escalar:

1.4.a.  $\alpha [A] = [\alpha] A$

1.4.b.  $\alpha(A + B) = \alpha A + \alpha B$

1.5. Multiplicação de Matrizes

Sejam  $A$  matriz  $(m \times n)$  e  $B$  matriz  $(n \times p)$ .

$C$ , matriz  $(m \times p)$  é o produto de  $A$  por  $B$  se

$$[A][B] = [C] \iff c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \text{ com}$$

$$i = 1, \dots, m \text{ e } j = 1, \dots, p.$$

Propriedades da Multiplicação de Matrizes

1.5.a.  $\alpha([A][B]) = (\alpha[A])[B] = [A](\alpha[B])$

1.5.b.  $[A]([B] + [C]) = [A][B] + [A][C]$

1.5.c.  $[A][B][C] = [A]([B][C]) = ([A][B])[C]$

1.5.d. Em geral  $[A][B] \neq [B][A]$  sendo que  $[B][A]$  pode ser exclusivamente não definido.

1.6. Matriz Identidade

Matriz identidade de ordem  $n$  é uma matriz quadrada tal que:

$$\begin{cases} a_{ij} = 1 & \text{para } i=j \\ a_{ij} = 0 & \text{para } i \neq j \end{cases}, \quad i, j = 1, \dots, n$$

1.7. Matriz Triangular Inferior ou Superior  
Matriz triangular inferior é tal que

$$a_{ij} = 0 \quad \text{para } i > j$$

$$\begin{bmatrix} z_{11} & 0 & 0 & \cdots & 0 \\ z_{21} & z_{22} & 0 & \cdots & 0 \\ z_{31} & z_{32} & z_{33} & \cdots & 0 \\ \vdots & \vdots & & \ddots & 0 \\ z_{n1} & z_{n2} & z_{n3} & \cdots & z_{nn} \end{bmatrix}$$

Matriz triangular superior é tal que  
 $a_{ij} = 0 \quad \text{para } i < j$

$$\begin{bmatrix} z_{11} & z_{12} & z_{13} & \cdots & z_{1n} \\ 0 & z_{22} & z_{23} & \cdots & z_{2n} \\ 0 & 0 & z_{33} & \cdots & \\ 0 & 0 & 0 & \ddots & z_{nn} \end{bmatrix}$$

1.8. Matriz Transposta

$[B]$  é transposta de  $[A] \Leftrightarrow b_{ji} = a_{ij} \quad \forall i=1, \dots, n; \forall j=1, \dots, m$

1.9. Matriz Simétrica

$[A]$  é simétrica  $\Leftrightarrow [A]^T = [A]$

ou  $a_{ji} = a_{ij} \quad \forall i=1, \dots, n, \forall j=1, \dots, m$

1.10. Matriz Antissimétrica

$[A]$  é antissimétrica  $\Leftrightarrow [A]^T = -[A]$

ou  $a_{ji} = -a_{ij} \quad \forall i=1, \dots, n; \forall j=1, \dots, m$

1.11. Matriz Diagonal

Seja  $[A]$  matriz quadrada:

$[A]$  é diagonal  $\Leftrightarrow a_{ij} = 0, \forall i \neq j, i, j = 1, \dots, n$

1.12. Matriz Inversa

$[A]$  é matriz quadrada.

$[A]^{-1}$  é a inversa de  $[A] \Leftrightarrow \det[A] \neq 0$  e  
 $[A]^{-1}[A] = [I]$

Se  $\det[A] = 0$ ,  $A$  é uma matriz singular

### 1.13. Propriedades de Matrizes Transpostas e Inversas

1.13.a.  $(A + B)^T = A^T + B^T$

1.13.b.  $(A \cdot B)^T = B^T A^T$

1.13.c.  $(A \cdot B)^{-1} = B^{-1} A^{-1}$

1.13.d.  $(A^{-1})^T = (A^T)^{-1}$

1.13.e.  $[(AB)^{-1}]^T = (A^T)^{-1} (B^T)^{-1}$

1.13.f. Se  $A$  é simétrica, ento  $B^T A B$  é também simétrica.

### 1.14. Formas Quadráticas

$[A]$  é uma matriz definida positiva se e somente se  $[A]$  matriz  $(n \times n)$  e  $\{x\}$  um vetor qualquer  $(n \times 1)$  satis que:

1.14.a.  $\{x\}^T [A] \{x\} \geq 0$  qualquer  $\{x\}$

1.14.b.  $\{x\}^T [A] \{x\} = 0 \Leftrightarrow \{x\} = 0$

A definição acima é rigorosa mas não é prática. Prazo prática:  $[A]$  é matriz definida positiva sse todos os seus menores principais têm determinantes positivos.

$$[A] = \begin{bmatrix} z_{11} & z_{12} & z_{13} & \dots & z_{1n} \\ z_{21} & z_{22} & z_{23} & \dots & z_{2n} \\ z_{31} & z_{32} & z_{33} & \dots & z_{3n} \\ \vdots & & \vdots & & \vdots \\ z_{n1} & z_{n2} & z_{n3} & \dots & z_{nn} \end{bmatrix}$$

Menores principais:

$$[A_1] = [z_{11}]$$

$$[A_2] = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix}$$

$$[A_3] = \begin{bmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \\ z_{31} & z_{32} & z_{33} \end{bmatrix}$$

⋮

$$[A_n] = [A]$$

Então:  $\det[A_i] > 0$ ,  $i=1, 2, \dots, n$

$[B]$  será uma matriz definida negativa se

$$\det[B_i] \begin{cases} < 0 & i \text{ ímpar} \\ > 0 & i \text{ par} \end{cases}$$

### 1.15. Inversão de Matrizes

Um dos aspectos mais importantes na implementação do Filtro de Kalman é o condicionamento numérico das matrizes envolvidas. A menos do uso de hipóteses simplificadoras, um programa de estimação com uso do filtro de Kalman deverá conter uma subrotina de inversão de matrizes e uma subrotina de produto de matrizes.

Com relação à inversão de matrizes lembraremos aqui alguns conceitos fundamentais e tentaremos desenvolver algumas aplicações recentes como a decomposição de Choleski, de uso bastante comum nas modernas técnicas de propagação do filtro, e o uso do filtro de raiz quadrada.

#### 1.15.2. Menor Mij (não principal)

A partir de uma matriz  $(n \times n)$   $A$  se retirarmos a  $i$ -ésima linha e  $j$ -ésima coluna, teremos uma matriz  $[(n-1) \times (n-1)]$ . O determinante dessa matriz  $[(n-1) \times (n-1)]$  é chamado o

menor  $M_{ij}$  da matriz  $A$ .

1.15.b. O cofator  $\bar{A}_{ij}$  do elemento  $a_{ij}$  da matriz  $A$  é:  
 $\bar{A}_{ij} = (-1)^{i+j} M_{ij}$

Note que o cofator  $\bar{A}_{ij}$  do elemento  $a_{ij}$  é o coeficiente do termo  $a_{ij}$  na expansão do determinante  $|A|$  já que:

$$a_{11}\bar{A}_{11} + a_{12}\bar{A}_{12} + \dots + a_{1n}\bar{A}_{1n} = |A|$$

1.15.c. Matriz Adjunta

A matriz  $B$ , adjunta de  $A$ , é a transposta da matriz cujos elementos são os cofatores de  $A$ . Portanto

$$\text{adj } A = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{21} & \dots & \bar{A}_{n1} \\ \bar{A}_{12} & \bar{A}_{22} & \dots & \bar{A}_{n2} \\ \vdots & \vdots & & \vdots \\ \bar{A}_{1n} & \bar{A}_{2n} & \dots & \bar{A}_{nn} \end{bmatrix}$$

Dadas essas definições, a matriz inversa de  $A$  de  $A$  é calculada por

$$A^{-1} = \frac{\text{adj } A}{|A|} = \begin{bmatrix} \frac{\bar{A}_{11}}{|A|} & \frac{\bar{A}_{21}}{|A|} & \dots & \frac{\bar{A}_{n1}}{|A|} \\ \frac{\bar{A}_{12}}{|A|} & \frac{\bar{A}_{22}}{|A|} & \dots & \frac{\bar{A}_{n2}}{|A|} \\ \vdots & \vdots & & \vdots \\ \frac{\bar{A}_{1n}}{|A|} & \frac{\bar{A}_{2n}}{|A|} & \dots & \frac{\bar{A}_{nn}}{|A|} \end{bmatrix}$$

A decomposição de Choleski que expomos a seguir tem sido utilizada em Teoria de Estimação para aumentar a estabilidade numérica dos algoritmos. Uma constatação verifica-se na simulação dos filtros e que propagar a raiz quadrada da covariância é menos sensível a erros que trabalhar com a matriz de covariância.

A partir de uma matriz  $A$ , de elementos  $a_{ik}$ , é possível definir uma forma quadrática definida positiva que pode ser representada como uma soma de quadrados:

$$Q(x) = x^T A x = \sum_{i=1}^n \sum_{k=1}^n z_{ik}^{(0)} x_i x_k = \sum_{i=1}^n \left( \sqrt{z_{ii}} x_i + \sum_{k=i+1}^n \frac{z_{ik}^{(0)}}{\sqrt{z_{ii}}} x_k \right)^2$$

Nessa decomposição aparecem 2 formas lineares

$$y_i = \sum_{k=i}^n r_{ik} x_k \quad i=1, 2, \dots, n$$

cujos coeficientes  $r_{ik}$  são definidos por:

$$r_{ii} = \sqrt{z_{ii}} \quad \text{e} \quad r_{ik} = \frac{z_{ik}^{(0)}}{\sqrt{z_{ii}}} \quad k \geq i$$

Nessas definições os valores  $r_{ik}$  são dados apenas para  $k \geq i$ . Se colocarmos  $r_{ik} = 0$  para  $k < i$ , os coeficientes podem ser colocados numa matriz triangular superior  $R$  que define uma transformação linear  $y = Rx$ .

Isto sugere a seguinte sequência de igualdades:

$$Q(x) = x^T A x = \sum_{i=1}^n \left( \sum_{k=i}^n r_{ik} x_k \right)^2 = \sum_{i=1}^n y_i^2 = y^T y ; y^T = (Rx)^T$$

$$y^T y = x^T R^T R x = x^T A x$$

Como a representação é única

$$A = R^T R$$

$R^T$ , nesse caso, é triangular superior. A mesma dedução pode ser feita para matriz triangular inferior. A decomposição da matriz  $A$ , então, é uma decomposição de Choleski e  $R$  é ditz raiz quadrada da matriz  $A$ . A raiz quadrada de uma matriz não é única, visto que todas transformações ortogonais levam a uma outra matriz com as mesmas propriedades.

A decomposição da matriz  $A$  num produto de matrizes triangulares facilita enormemente o processo de inversão já que:

- \* o determinante de uma matriz triangular é o produto dos termos da diagonal
- \*\* a inversa de uma matriz triangular superior é feita facilmente por substituição já que isso implica resolver o sistema:

$$\Leftrightarrow \begin{cases} r_{11} s_{11} = 1 \\ r_{11} s_{12} + r_{22} s_{21} = 0 \\ r_{22} s_{22} = 0 \\ r_{31} s_{11} + r_{32} s_{21} + r_{33} s_{31} = 0 \\ r_{32} s_{22} + r_{33} s_{32} = 0 \\ r_{33} s_{33} = 0, \text{ etc} \end{cases} \quad (\text{det } R \neq 0)$$

onde  $s_{ij}$  são os elementos da matriz  $S$ , inversa de  $R$ .

Então, temos um procedimento para inverter a matriz  $A$

$$B = A^{-1} = (R^T R)^{-1} = R^{-1} (R^T)^{-1} = R^{-1} (R^{-1})^T$$

Consiste nos três seguintes passos:

- Decomposição de Choleski de  $A = R^T R$
- Inversão de  $R$ , gerando uma matriz triangular superior  $S = R^{-1} = (s_{ik})$  com  $s_{ik} = 0$  para  $i > k$
- Obtenção de  $B = S S^T$ .

### 1.16. Autovetores e Autovetores

Associados a todo matriz quadrada  $A$  existem um conjunto de vetores, chamados autovetores, e um conjunto de escalares, chamados autovalores.

$u$  é um autovetor de  $A$  se  $u$  é não nulo e  $A$ , que pode ser nulo, é um escalar t.q.:

$$A u = \lambda u$$

O escalar  $\lambda$  é um autovetor de  $A$  se existe um autovetor  $u$  não nulo tal que a igualdade acima seja satisfeita. Dessa igualdade:

$$(A - \lambda I) u = 0$$

e define-se a função característica da matriz  $A$  por

$$\phi(\lambda) = \det(A - \lambda I)$$

A função característica fornece os autovetores da matriz  $A$  através de

$$\phi(\lambda) = \det(A - \lambda I) = 0$$

Uma maneira de se determinar os autovetores da matriz  $A$  é obtida através da construção da matriz

adjunta  $C = \text{adj}(A - \lambda_i I)$ . Nesse caso, mostra-se que qualquer coluna não nula de  $C$  pode ser usado como autovetor correspondente a  $\lambda_i$ .

### 1.17. Partição de Matrizes.

Em muitas ocasiões se faz necessário subdividir uma matriz  $A$  em blocos retangulares e considerar as partes resultantes para efeitos de cálculo. Esse processo é chamado partição de matriz  $A$  e as partes resultantes são chamadas sub-matrizes.

$$A = \left[ \begin{array}{cccc|ccccc} a_{11} & a_{12} & \dots & a_{1k} & a_{1,k+1} & a_{1,k+2} & \dots & a_{1m} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{e1} & a_{e2} & \dots & a_{ek} & a_{e,k+1} & a_{e,k+2} & \dots & a_{em} \\ \hline a_{e+1,1} & a_{e+1,2} & \dots & a_{e+1,k} & a_{e+1,k+1} & a_{e+1,k+2} & \dots & a_{e+1,m} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nk} & a_{n,k+1} & a_{n,k+2} & \dots & a_{nm} \end{array} \right] = \begin{bmatrix} A_{11} & | & A_{12} \\ \hline A_{e1} & | & A_{e2} \\ \hline A_{n1} & | & A_{n2} \end{bmatrix}$$

Uma propriedade importante com relação às matrizes particionadas é que se considerarmos duas matrizes  $A$  e  $B$ , de dimensão  $(n \times n)$ , particionadas da mesma maneira, i.e., as correspondentes submatrizes, de  $A$  e de  $B$ , tem a mesma dimensão e  $A_{11}, B_{11}$  e  $A_{22}, B_{22}$  são quadradas, então as submatrizes  $A_{ij}$  e  $B_{ij}$  podem ser tratadas como elementos simples, do tipo  $(a_{ij})$ ,  $(b_{ij})$  para efeitos de operações. Para verificar essa afirmação consideremos:

$$A = \begin{bmatrix} A_{11} & | & A_{12} \\ \hline A_{21} & | & A_{22} \end{bmatrix} \quad A_{11} [p \times p]; A_{21} [(n-p) \times p]; A_{12} [p \times (n-p)]; \\ A_{22} [(n-p) \times (n-p)]$$

$$B = \begin{bmatrix} B_{11} & | & B_{12} \\ \hline B_{21} & | & B_{22} \end{bmatrix} \quad , \quad [C] = \begin{bmatrix} C_{11} & | & C_{12} \\ \hline C_{21} & | & C_{22} \end{bmatrix} \quad ; \quad D = \begin{bmatrix} D_{11} & | & D_{12} \\ \hline D_{21} & | & D_{22} \end{bmatrix}$$

com os blocos de  $B$ ,  $C$  e  $D$  da mesma dimensão das blocos de  $A$ . Então:

$$A + B = \begin{bmatrix} A_{11} & | & A_{12} \\ \hline A_{21} & | & A_{22} \end{bmatrix} + \begin{bmatrix} B_{11} & | & B_{12} \\ \hline B_{21} & | & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11} + B_{11} & | & A_{12} + B_{12} \\ \hline A_{21} + B_{21} & | & A_{22} + B_{22} \end{bmatrix}$$

$$E = C \cdot D = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} = \begin{bmatrix} C_{11}D_{11} + C_{12}D_{21} & C_{11}D_{12} + C_{12}D_{22} \\ C_{21}D_{11} + C_{22}D_{21} & C_{21}D_{12} + C_{22}D_{22} \end{bmatrix} = \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix}$$

Note que, nesse caso, os blocos de  $E$  têm as mesmas dimensões dos blocos de  $C$  e  $D$ .

### 1.18. Identidades Matriciais

As identidades matriciais que serão obtidas neste item tem papel importante no processo de formações mais conveniente para processamento do filtro. Consideremos uma matriz  $\Lambda$ , simétrica e não singular, particionada da forma:

$$\Lambda = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

onde  $A$  e  $C$  são submatrizes quadradas, não singulares e necessariamente simétricas. Dependendo da partição,  $B$  é retangular ou não. Procuraremos a inversa de  $\Lambda$ .

$$\Lambda^{-1} = \begin{bmatrix} L & M \\ N & P \end{bmatrix} \quad \text{tal que}$$

$$\Lambda \cdot \Lambda^{-1} = I ; \quad \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} L & M \\ N & P \end{bmatrix} = \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix} = I$$

Particularizando para  $A$  e  $C$  da mesma dimensão, i.e.,  $I_1 = I_2$ , e desenvolvendo.

$$AL + BN = I_1 \quad (1)$$

$$AM + BP = 0 \quad (2)$$

$$BL + CN = 0 \quad (3)$$

$$BTM + CP = I_1 \quad (4)$$

Resolvendo o sistema acima:

$$L = [A - BC^{-1}B^T]^{-1}$$

$$P = [C - BT\Delta^{-1}B]^{-1}$$

$$M = -A^{-1}B [C - B^T A^{-1}B]^{-1}$$

$$N = -C^T B^T [A - BC^{-1}B^T]^{-1}$$

Veja que, agora, como  $\Lambda$  é simétrica então sua inversa  $\Lambda^{-1}$  também é simétrica:

$$\Lambda \Lambda^{-1} = I$$

$$[\Lambda^{-1}]^T \Lambda^T \Lambda^{-1} = I \quad (\Lambda \text{ é simétrica})$$

$$[\Lambda^{-1}]^T \Lambda^T \Lambda^{-1} = \Lambda^{-T} \Lambda^T \Lambda^{-1} = \Lambda^{-T}$$

$$I \Lambda^{-T} = \Lambda^{-T} \Rightarrow \Lambda^{-1} = \Lambda^{-T} = (\Lambda^{-1})^T$$

e  $\Lambda^{-1}$  é simétrica. De sorte, então, que:

$$\begin{cases} A^{-T} = (A^{-1})^T = A^{-1} \\ C^{-T} = (C^{-1})^T = C^{-1} \\ M = N^T \end{cases}$$

Usando o fato de  $M = N^T$ , igualamos os resultados obtidos e obtemos a primeira identidade importante:

$$A^{-1}B(C - B^T A^{-1}B)^{-1} = (A - BC^{-1}B^T)^{-1}BC^{-1}$$

(I.1)

A segunda identidade é obtida de

$$\Lambda^{-1}\Lambda = \begin{bmatrix} I & M \\ N & P \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} = I_{(n)}$$

que nos fornece:

$$NB + PC = I$$

$$P = C^{-1} - NBC^{-1}$$

Substituindo os resultados obtidos temos:

$$(C - B^T A^{-1}B)^{-1} = C^{-1} - C^{-1}B^T(BC^{-1}B^T - A)^{-1}BC^{-1}$$

(I.2)

Finalmente, as duas últimas identidades de interesse se podem ser obtidas por manipulação algébrica conveniente de (I.1) e (I.2) gerando (I.3) e (I.4) respectivamente.

$$AB^T(C + BABA^T)^{-1} = (A^{-1} + B^TC^{-1}B)^{-1}B^TC^{-1}$$

I.3

$$(C^{-1} + BA^{-1}B^T)^{-1} = C - CB(B^TCB + A)^{-1}B^TC$$

I.4

## 2. EQUAÇÕES DIFERENCIAIS ORDINÁRIAS

É quase um caso geral em problemas de Engenharia o tratamento de sistemas físicos variando no tempo através de equações diferenciais, ordinárias ou a derivadas parciais. Isto decorre do fato de leis físicas terem sido sistematicamente colocadas nas formas do cálculo diferencial, como são exemplo as leis de Newton da Mecânica. No entanto, em alguns casos, é conveniente o tratamento dos sistemas físicos através de equações de diferenças, representando uma integração das diferenciais.

O presente capítulo pretende apresentar uma revisão dos conceitos fundamentais das equações diferenciais ordinárias com grande ênfase nos sistemas lineares e com algumas analogias, quando pertinentes, aos sistemas de equações de diferenças. Não abordaremos aqui, nem ao longo do curso, sistemas de equações a derivadas parciais, embora para esse último tipo de sistema já existam algumas aplicações dentro da Teoria de Estimação.

### 2.1. Modelos Dinâmicos Contínuos

O modelo mais simples para representar a evolução temporal de um sistema físico é o de um modelo linear, invariante no tempo com entrada e saída simples (SISO), que pode ser descrito por uma equação diferencial linear, de coeficientes constantes e de ordem  $n$ :

$$\frac{d^n z(t)}{dt^n} + a_{n-1} \frac{d^{n-1} z(t)}{dt^{n-1}} + \dots + a_0 z(t) = c_p \frac{d^p u(t)}{dt^p} + \dots + c_0 u(t)$$

onde  $u(t)$  é a entrada no sistema no instante  $t$  e  $z(t)$  é a correspondente saída

Se aplicarmos transformadas de Laplace sobre o modelo acima, admitindo condições iniciais nulas e reescrevendo a equação, geramos a função de transferência  $G(s)$  que descreve a saída  $z(t)$  em função da entrada  $u(t)$ , no domínio das frequências:

$$z(s) = G(s) u(s)$$

$$G(s) = \frac{c_p s^p + c_{p-1} s^{p-1} + \dots + c_1 s + c_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}$$

A correspondente representação no espaço de estados, para o mesmo sistema, seria uma equação diferencial vetorial de 1º ordenem, associada a uma relação de saída:

$$\begin{aligned}\dot{x}(t) &= Fx(t) + bu(t) \\ z(t) &= h^T x(t)\end{aligned}$$

onde  $x$  é um vetor de dimensão  $n$ ,  $F$  é uma matriz  $(n \times n)$  de constantes e  $b$  e  $h$  são vetores de constantes.

Mostra-se, nesse caso, que a relação entre a representação no espaço de estados e a função de transferência é dada por:

$$G(s) = h^T [sI - F]^{-1} b$$

A equação acima revela que os polos da função de transferência empêcam com os autovalores da matriz  $F$ , associados à parte homogênea do sistema de equações diferenciais.

A representação no espaço de estados não é única. Um certo conjunto de variáveis de estados determina, univocamente, apenas o comportamento do sistema, mas, existem infinitos

conjuntos com essa propriedade. Existem pelo menos 4 formas comuns de representação no espaço de estados: variáveis físicas (aplicação direta das leis físicas); forma controlável padrão; forma observável padrão (diferem apenas pela definição de  $b$  e  $h$  no sistema de equações diferenciais) e forma canônica (decomposição nos modos). Essas várias formas de representação estão relacionadas através de transformações matriciais.

Representações no espaço de estados podem ser facilmente estendidas para sistemas de múltiplos entradas e múltiplas saídas (MIMO). Admitindo zíndic sistemas juntamente no tempo, com uma entrada  $u(t)$  fornecida por um vetor de dimensão  $r$  e com uma saída  $z(t)$  através de um vetor  $m$ -dimensional, o modelo toma a forma:

$$\begin{cases} \dot{x}(t) = Fx(t) + Bu(t) \\ z(t) = Hx(t) \end{cases}$$

onde  $F$ ,  $B$  e  $H$  são matrizes de constantes de dimensões  $(n \times n)$ ,  $(n \times r)$  e  $(m \times n)$ , respectivamente.

Uma representação equivalente pode ser montada em termos de uma matriz de funções de transferência cujos elementos seriam as funções de transferência dos componentes individuais dos vetores de entrada e saída

$$z(s) = G(s) u(s)$$

$$G(s) = \begin{bmatrix} G_{11}(s) & \dots & G_{1r}(s) \\ \vdots & & \vdots \\ G_{m1}(s) & \dots & G_{mr}(s) \end{bmatrix} = H[sI - F]^{-1}B$$

Sistemas variantes no tempo são gerados rapidamente através de modelos de espaço de estado: as matrizes def-

do a estrutura do sistema varia com o tempo:

$$\dot{x}(t) = F(t) x(t) + B(t) u(t)$$

$$z(t) = H(t) x(t)$$

Transformações de Laplace não são facilmente aplicáveis a esses casos, dali a preferência pelo representações de estado.

Finalmente, um modelo não linear de estado pode ser descrito através de um sistema de equações diferenciais em uma forma

$$\dot{x}(t) = f[x(t), u(t), t]$$

$$z(t) = h[x(t), t]$$

Exemplo: Satélite num órbita planar

O modelo para um satélite em órbita planar pode ser estabelecido através de uma aproximação usando um potencial de massa unitária num campo de forças que variam com o inverso do quadrado da distância.

Seja  $r$  a distância do centro da Terra, tomado como centro do campo de forças, ao satélite e seja  $\theta$  o ângulo entre um eixo de referência fixo e a linha do centro da Terra ao satélite. Admita que o satélite tem como controles um empuxo radial  $u_r$  e um empuxo tangencial  $u_t$ .

O movimento do satélite é então governado por um par de equações diferenciais acopladas de segunda ordem:

$$\ddot{r}(t) = r(t) \dot{\theta}^2(t) - \frac{G}{r^2(t)} + u_r(t)$$

$$\ddot{\theta}(t) = -\frac{2}{r(t)} \dot{\theta}(t) \dot{r}(t) + \frac{1}{r(t)} u_t(t)$$

Essas relações não lineares podem ser colocadas na

forma padrão escala usando como estados

$$\begin{aligned} x_1 &= \text{f} \\ \dot{x}_2 &= \text{r} \\ x_3 &= \Theta \\ \dot{x}_4 &= \dot{\Theta} \end{aligned} ; \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} ; \quad u = \begin{bmatrix} u_r \\ u_\theta \end{bmatrix}$$

Com isso, escrevemos

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \\ \dot{x}_4(t) \end{bmatrix} = \begin{bmatrix} f_1[x(t), u(t), t] \\ f_2[x(t), u(t), t] \\ f_3[x(t), u(t), t] \\ f_4[x(t), u(t), t] \end{bmatrix} = \begin{bmatrix} x_2(t) \\ x_1(t)x_4^2(t) - \frac{G_1}{x_1^2(t)} + u_r(t) \\ x_4(t) \\ -\frac{2}{x_1(t)}x_4(t)x_2(t) + \frac{1}{x_1(t)}u_\theta(t) \end{bmatrix}$$

## 2.2. SISTEMAS LINEARES: MATRIZ DE TRANSIÇÃO

Para os sistemas de equações diferenciais lineares é possível, em geral, obter uma descrição da resposta do sistema na forma de uma matriz que relaciona os estados entre dois instantes quaisquer de tempo. Essa matriz é chamada Matriz de Transição de Estados ou Matriz fundamental. Para o caso de modelos não-lineares é possível uma extensão das definições e conceitos envolvidos mas a geração de funções de transição de estado se tornam tão complicadas que o interesse prático fica muito restrito. No entanto, como técnicas de análise de sistemas não-lineares quase sempre envolvem linearizações em torno de pontos pré-estabelecidos, a geração de matrizes de transição é também de grande importância para os modelos não-lineares.

Consideremos, então, sistemas lineares na forma:

$$\dot{x}(t) = F(t)x(t) + G(t)u(t)$$

dou a condição inicial  $x(t_0) = x_0$  e. onde:

$F(t)$  é uma matriz de funções ( $n \times n$ )

$G(t)$  é uma matriz de funções ( $n \times m$ )

$x(t)$  é o vetor de estados ( $n \times 1$ )

e  $u(t)$  é o vetor de entradas ( $m \times 1$ )

Tomemos juntamente o caso linear homogêneo

$$\dot{x}(t) = F(t)x(t)$$

Se os elementos de  $F(t)$  são funções contínuas de tempo é possível provar que o sistema acima possue solução única correspondente ao círculo vetor de estado inicial  $x_0$ .

Para o sistema homogêneo é possível definir o chamado Conjunto Fundamental de Soluções. Para se chegar a esse conjunto, diremos antes que um conjunto de funções vetoriais com valores no tempo  $x^1(t), x^2(t), \dots, x^m(t)$  é linearmente independente se não existir nenhuma combinação linear desses vetores que seja identicamente nula, isto é:

$$\text{se } d_1x^1(t) + d_2x^2(t) + \dots + d_mx^m(t) = 0, \forall t \iff d_1 = d_2 = d_3 = \dots = d_m = 0$$

Tomemos, então,  $x^1(t), x^2(t), \dots, x^m(t)$ ; um conjunto linearmente independente de soluções para o sistema  $\dot{x}(t) = F(t)x(t)$ . Ignoramos a menor prova que esse conjunto de soluções tenha sido obtido, mas admitimos que existem essas soluções linearmente independentes formam um conjunto fundamental de soluções.

Arranjemos, agora, as soluções  $x^1(t), x^2(t), \dots, x^m(t)$  como colunas de uma matriz ( $n \times n$ )  $\bar{X}(t)$ . Nesse caso, o

matriz  $\tilde{X}(t)$  é dito Matriz Fundamental de Soluções e mostra-se que satisfaz a equação matricial diferencial:

$$\dot{\tilde{X}}(t) = F(t) \tilde{X}(t)$$

Note que na matriz fundamental não está envolvida a condição inicial do sistema. A partir da matriz fundamental é possível, então, enunciar os dois seguintes teoremas, cuja prova se encontra em Luenberger:

T.1. Uma matriz fundamental de soluções  $\tilde{X}(t)$  é não singular em qualquer instante de tempo  $t$ .

T.2. Seja  $\tilde{X}(t)$  uma matriz fundamental de soluções correspondente ao sistema  $\dot{x}(t) = F(t)x(t)$ . Podemos então montar uma matriz na forma:

$$\underline{\Phi}(t, \bar{t}) = \tilde{X}(t) \tilde{X}^{-1}(\bar{t})$$

A matriz  $\underline{\Phi}(t, \bar{t})$  é chamada Matriz de Transição de Estados e se mostra:

) T.3. A solução do sistema homogêneo  $\dot{x}(t) = F(t)x(t)$ , com a condição inicial  $x(t_0) = x_0$ , é dada pela Matriz de Transição de Estado na forma:

$$x(t) = \underline{\Phi}(t, t_0) x(t_0)$$

com  $\underline{\Phi}(t, t_0)$  dada por:

$$\dot{\underline{\Phi}}(t, t_0) = F(t) \underline{\Phi}(t, t_0)$$

e com a condição inicial  $\underline{\Phi}(t_0, t_0) = I_n$

Prova:

Claramente a condição inicial é satisfeita visto que:

$$x(t_0) = \underline{\Phi}(t_0, t_0) x(t_0) = I_n x(t_0) = x(t_0) = x_0$$

Diferenciando  $\dot{x}(t) = \underline{\Phi}(t, t_0) x(t_0)$  em relação ao tempo:

$$\dot{\dot{x}}(t) = \dot{\underline{\Phi}}(t, t_0) x(t_0) + \underline{\Phi}(t, t_0) \dot{x}(t_0) = \dot{\underline{\Phi}}(t, t_0) x(t_0)$$

Por outro lado  $\dot{\underline{\Phi}} = F \underline{\Phi}$  e:

$$\dot{\dot{x}}(t) = \dot{\underline{\Phi}}(t, t_0) x(t_0) = F(t) \underline{\Phi}(t, t_0) x(t_0) = F(t) x(t)$$

e  $\underline{\Phi}(t, t_0)$  satisfaz todos os requisitos para ser solução do sistema homogêneo.

Normalmente, o que se faz não é gerar o conjunto fundamental de soluções  $X(t)$  e sim gerar diretamente a matriz de transição. Voltaremos, para verificar esse fato, ao sistema linear homogêneo.

$$\dot{x}(t) = F(t) x(t)$$

Para  $t = t_0$ , tomemos como condições iniciais

$$\bar{x}_{(t_0)}^j = [0, 0, \dots, 0, 1, 0, \dots, 0]^T = e_j \quad (\text{elemento da } j\text{-ésima linha de } R^n, j=1, \dots, n)$$

Se integrarmos o sistema homogêneo com cada uma dessas  $n$  diferentes condições iniciais, isto é, integrarmos para todos os  $e_j$ ,  $j=1, \dots, n$ , teremos num instante  $t$  qualquer:

$$\underline{\Phi}(t, t_0) = [\bar{x}_1^1(t), \bar{x}_2^1(t), \dots, \bar{x}_n^1(t), \dots, \bar{x}_1^n(t)] \quad (n \times n)$$

com a  $j$ -ésima coluna dada por  $\bar{x}_j^1(t)$

Dessa forma temos:  $\hat{\Phi}(t_0, t_0) = I_n$

$$\text{e } x(t_0) = \sum_{j=1}^n \bar{x}_j(t_0) x_j \Rightarrow x(t_0) = \hat{\Phi}(t_0, t_0) x_0 = x_0$$

Exemplo: Consideremos um sistema de duas variáveis de estado, não autônomo, na forma

$$\dot{x}(t) = \begin{bmatrix} 0 & 0 \\ t & 1/t \end{bmatrix} x(t), \quad t \geq 1$$

$$\text{ou } \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ t & 1/t \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Por verificações imediatas

$$\dot{x}_1 = 0 \Rightarrow x_1(t) = C \text{ onde } C \text{ é constante}$$

Então: Para integrar  $x_2(t)$ , usamos a transformação  $z(t) = \frac{x_2(t)}{t}$ .

$$\dot{z} = \frac{\dot{x}_2}{t} - \frac{x_2}{t^2} \Rightarrow \dot{x}_2 = \dot{z}t^2 + x_2 \Rightarrow \dot{x}_2 = \dot{z}t + z$$

De  $\dot{x}_2 = tx_1 + \frac{1}{t}x_2$ , por substituições:

$$\dot{z}t + z = tC + z \Rightarrow \dot{z}(t) = C \Rightarrow z(t) = Ct + D$$

$$\text{e } x_2(t) = Ct^2 + Dt$$

Com isso, uma possível matriz fundamental de soluções é:

$$X(t) = \begin{bmatrix} 1 & 0 \\ t^2 & t \end{bmatrix}$$

e montamos a matriz de transição fazendo primeiro:

$$\bar{X}^{-1}(0) = \begin{bmatrix} 1 & 0 \\ 8^2 & 8 \end{bmatrix}^{-1} = \frac{1}{8} \begin{bmatrix} 8 & -8^2 \\ 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ -8 & 1/8 \end{bmatrix}$$

Dai:  $\underline{\Phi}(t,0) = \bar{X}(t)\bar{X}^{-1}(0)$  ver T1 e T2 da pg 2.7  $= \begin{bmatrix} 1 & 0 \\ t^2 & t \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -8 & 1/8 \end{bmatrix}$

$$\underline{\Phi}(t,0) = \begin{bmatrix} 1 & 0 \\ t(t-8) & t/8 \end{bmatrix} ; \quad \underline{\Phi}(t,t_0) = \underline{\Phi}(t_0,t_0) = \underline{\Phi}(0,0)$$

## 2.3. SISTEMAS LINEARES INVARIANTES NO TEMPO

Considere o sistema linear invariante no tempo:

$$\dot{x}(t) = Ax(t)$$

Nesse caso, para sistemas lineares com coeficientes constantes, a matriz fundamental de soluções satisfez:

$$\dot{\bar{X}}(t) = A\bar{X}(t) ; \quad \bar{X}(0) = I$$

e pode ser expressa por uma expansão em série de potências:

$$\bar{X}(t) = I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \dots + \frac{1}{k!}A^kt^k + \dots = e^{At}$$

Para verificar essa solução basta colocar  $t=0$  na série e:

$$\bar{X}(t_0) = \bar{X}(0) = I$$

Diferenciando em relação ao tempo cada termo da série:

$$\dot{\bar{X}}(t) = 0 + A + A^2t + \frac{A^3t^2}{2!} + \dots$$

$$\dot{\bar{X}}(t) = A(I + At + \frac{A^2t^2}{2!} + \dots) = A\bar{X}(t)$$

No desenvolvimento acima encontra-se o resultado:

da a definição de exponencial de uma matriz  $A$ , dada por:

$$e^{At} = I + At + \frac{1}{2!} A^2 t^2 + \frac{1}{3!} A^3 t^3 + \dots$$

A matriz de transição ficará, então, na forma:

$$\underline{\Phi}(t, \tau) = \bar{X}(t) \bar{X}^{-1}(\tau) = e^{At} e^{-A\tau}$$

$$\underline{\Phi}(t, \tau) = e^{A(t-\tau)}$$

Exemplos:

a) Sistema de massa pura

$$\ddot{y}(t) = u(t)$$

Definindo  $x_1 = y$ ;  $x_2 = \dot{x}_1 = u$ , o sistema fica:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

Identificando:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\text{e: } A^2 = A \cdot A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \Rightarrow A^2 = [0]$$

De mesma forma  $A^3 = A^4 = \dots = [0]$  e a matriz de transição toma o seguinte encarregamento:

$$\underline{\Phi}(t, \tau) = e^{A(t-\tau)} = I + A(t-\tau) + \cancel{\frac{A^2(t-\tau)^2}{2!}} + \dots = I + A(t-\tau)$$

$$\underline{\Phi}(t, \tau) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} (t-\tau) \Rightarrow \underline{\Phi}(t, \tau) = \begin{bmatrix} 1 & t-\tau \\ 0 & 1 \end{bmatrix}$$

b) Oscilador Harmônico

$$\ddot{y} + \omega^2 y = 0$$

Definindo  $x_1 = y$ ,  $\dot{x}_1 = x_2 = \dot{y}$ , o sistema fica:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Identificando:

$$A = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} = \begin{bmatrix} -\omega^2 & 0 \\ 0 & -\omega^2 \end{bmatrix} = -\omega^2 I$$

$$A^3 = A^2 A = \begin{bmatrix} -\omega^2 & 0 \\ -\omega^2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} = -\omega^2 I A = -\omega^2 A$$

$$A^4 = A^3 A = -\omega^2 A \cdot A = -\omega^2 A^2 = -\omega^2 (-\omega^2 I) = +\omega^4 I$$

e concluimos por indução que, para k par:

$$A^k = (-1)^{k/2} \omega^k I$$

$$A^{k+1} = (-1)^{(k+1)/2} \omega^k A$$

Montamos então a matriz de transição:

$$\Phi(t, \tau) = e^{A(t-\tau)} = I + (t-\tau) \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} - \frac{\omega^2(t-\tau)^2}{2!} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Phi(t, \tau) = \begin{bmatrix} 1 - \frac{\omega^2(t-\tau)^2}{2!} + \frac{\omega^4(t-\tau)^4}{4!} - \frac{\omega^6(t-\tau)^6}{6!} + \dots & (t-\tau) - \frac{\omega^2(t-\tau)^3}{3!} + \frac{\omega^4(t-\tau)^5}{5!} + \dots \\ -\omega^2(t-\tau) + \frac{\omega^4(t-\tau)^3}{3!} - \frac{\omega^6(t-\tau)^5}{5!} + \dots & 1 - \frac{\omega^2(t-\tau)^2}{2!} + \frac{\omega^4(t-\tau)^4}{4!} + \dots \end{bmatrix}$$

$$\underline{\Phi}(t, t_0) = \begin{bmatrix} \cos \omega(t-t_0) & \frac{1}{\omega} \sin \omega(t-t_0) \\ -\omega \sin \omega(t-t_0) & \cos \omega(t-t_0) \end{bmatrix}$$

## 2.4. PROPRIEDADES DA MATRIZ DE TRANSIÇÃO

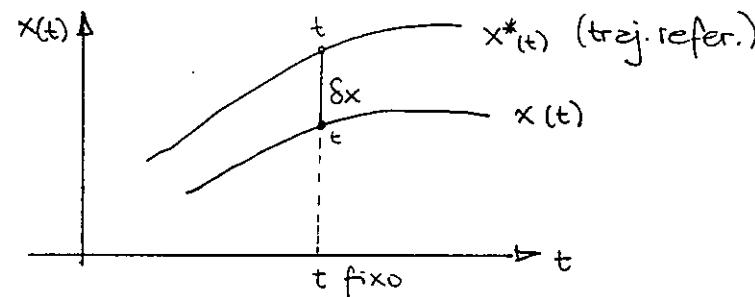
Propriedade I:

Partamos de  $\underline{x}(t) = \underline{\Phi}(t, t_0) \underline{x}(t_0)$  solução de  $\dot{x} = Fx$ .

Para 2 propagções do sistema de equações diferenciais, a partir de duas condições iniciais diferentes:

$$x^*(t) = \underline{\Phi}(t, t_0) x^*(t_0) \quad - \text{trajetória de referência}$$

$$x(t) = \underline{\Phi}(t, t_0) x(t_0)$$



Subtraindo as relações acima:

$$x^*(t) - x(t) = \underline{\Phi}(t, t_0) [x^*(t_0) - x(t_0)]$$

e, aproximando por uma perturbação:

$$\delta x(t) = \underline{\Phi}(t, t_0) \delta x(t_0)$$

Identificando o i-ésimo componente na relação acima:

$$\delta x_i = \sum_{j=1}^n \underline{\Phi}_{ij} \delta x_j(t_0)$$

$$\delta x_i = \underline{\Phi}_{i1} \delta x_1(t_0) + \underline{\Phi}_{i2} \delta x_2(t_0) + \dots + \underline{\Phi}_{in} \delta x_n(t_0)$$

Mas, pela regra de variação ou perturbação:

$$\delta x_i \triangleq \frac{\partial x_i}{\partial x_1(t_0)} \delta x_1(t_0) + \frac{\partial x_i}{\partial x_2(t_0)} \delta x_2(t_0) + \dots + \frac{\partial x_i}{\partial x_n(t_0)} \delta x_n(t_0)$$

Identificando as duas últimas equações:

$$\underline{\Phi}_{ij}(t, t_0) = \frac{\partial x_i(t)}{\partial x_j(t_0)}$$

o que nos leva a:

$$\underline{\Phi}(t, t_0) = \frac{\partial \mathbf{x}(t)}{\partial \mathbf{x}(t_0)}$$

A propriedade acima nos sugere uma maneira de gerar a matriz de transição no computador:

1º passo) Geramos a trajetória nominal  $\mathbf{x}^*$  a partir de  $\dot{\mathbf{x}} = \mathbf{F}\mathbf{x}$  e de  $\mathbf{x}(t_0)$  dada;

2º passo) Geramos uma segunda trajetória a partir de  $\dot{\mathbf{x}} = \mathbf{F}\mathbf{x}$  com uma nova condição inicial  $\mathbf{x}(t_0) + [0, \dots, 0, \Delta x_j(t_0), 0, \dots, 0]$ , com  $\Delta x_j(t_0)$  qualquer, obtendo solução  $[\bar{\mathbf{x}}(t)]_j$ ;

3º passo) Fazemos

$$\Delta x_i(t) = [\bar{\mathbf{x}}_i(t)]_j - x_i(t)$$

$$\underline{\Phi}_{ij}(t, t_0) \approx \frac{\Delta x_i(t)}{\Delta x_j(t_0)} ; i = 1, 2, \dots, n$$

e obtemos a j-ésima coluna de  $\underline{\Phi}(t, t_0)$ ;

4º passo) Retornamos ao 2º passo modificando j até que as n colunas de  $\underline{\Phi}(t, t_0)$  tenham sido geradas.

## Propriedade II (Existência da Inversa)

Consideremos um novo sistema linear na forma

$$\dot{\mathbf{A}}(t, t_0) = -\mathbf{A}(t, t_0) \mathbf{F}(t)$$

com a condição inicial  $A(t_0, t_0) = I$

Por outro lado:

$$\underline{\Phi}(t, t_0) = F(t) \underline{\Phi}(t, t_0)$$

Multiplicando por  $A(t, t_0)$

$$A(t, t_0) \underline{\Phi}(t, t_0) = A(t, t_0) F(t) \underline{\Phi}(t, t_0) = -\dot{A}(t, t_0) \underline{\Phi}(t, t_0)$$

$$\Delta(t, t_0) \underline{\Phi}(t, t_0) + \dot{A}(t, t_0) \underline{\Phi}(t, t_0) = 0$$

O que nos leva a:  $A(t, t_0) \underline{\Phi}(t, t_0) = C$ , matriz constante, e

Como  $A(t_0, t_0) = \underline{\Phi}(t_0, t_0) = I_n$  então  $C = I_n$  e:

$$A(t, t_0) \underline{\Phi}(t, t_0) = I_n$$

Logo  $\Delta(t, t_0) = \underline{\Phi}^{-1}(t, t_0)$

Por outro lado  $x(t) = \underline{\Phi}(t, t_0) x(t_0)$  nos leva a:

$$x(t_0) = \underline{\Phi}(t_0, t) x(t) = \underline{\Phi}^{-1}(t, t_0) x(t)$$

e, finalmente:

$$\underline{\Phi}^{-1}(t, t_0) = \underline{\Phi}(t_0, t)$$

Propriedade III (Transição)

$$\text{de: } x(t_1) = \underline{\Phi}(t_1, t_0) x(t_0)$$

$$\text{e } x(t_2) = \underline{\Phi}(t_2, t_1) x(t_1) = \underline{\Phi}(t_2, t_1) \underline{\Phi}(t_1, t_0) x(t_0)$$

$$x(t_2) = \underline{\Phi}(t_2, t_0) x(t_0)$$

conclui-se:

$$\underline{\Phi}(t_2, t_0) = \underline{\Phi}(t_2, t_1) \underline{\Phi}(t_1, t_0)$$

2.5. SOLUÇÃO GERAL PARA SISTEMAS NÃO HOMOGENEUS

Consideremos agora o sistema linear não homo-

gênero:

$$\dot{x}(t) = f(t)x(t) + g(t)u(t)$$

Admitimos a solução na forma:

$$x(t) = \underline{\Phi}(t, t_0) z(t)$$

e procuremos a forma de  $z(t)$

Diferencemos  $x(t) = \underline{\Phi}(t, t_0) z(t)$  em relação ao tempo:

$$\dot{x}(t) = \underline{\Phi}'(t, t_0) z(t) + \underline{\Phi}(t, t_0) \dot{z}(t)$$

$$\dot{x}(t) = F(t) \underbrace{\underline{\Phi}(t, t_0) z(t)}_{=x(t)} + \underline{\Phi}(t, t_0) \dot{z}(t)$$

$$\dot{x}(t) = F(t)x(t) + \underline{\Phi}(t, t_0) \dot{z}(t)$$

Comparando com  $\dot{x}(t) = f(t)x(t) + g(t)u(t)$ , consegue-

se:

$$G(t)u(t) = \underline{\Phi}(t, t_0) \dot{z}(t)$$

$$\dot{z}(t) = \underline{\Phi}^{-1}(t, t_0) G(t)u(t)$$

Integrando a relação acima na forma de um produto de convolução:

$$z(t) = z_0 + \int_{t_0}^t \underline{\Phi}^{-1}(s, t_0) G(s)u(s)ds$$

onde  $z_0$  é a condição inicial em  $z$ . Multiplicando por  $\underline{\Phi}(t, t_0)$ :

$$\underline{\Phi}(t, t_0)z(t) = \underline{\Phi}(t, t_0)z_0 + \int_{t_0}^t \underline{\Phi}(t, t_0) \underline{\Phi}^{-1}(s, t_0) G(s)u(s)ds$$

Como  $\underline{\Phi}(t, t_0)z(t) = x(t)$  temos que:

$$x(t_0) = \underline{\Phi}(t_0, t_0)z(t_0) = \text{In } z(t_0) = z_0$$

e, finalmente:

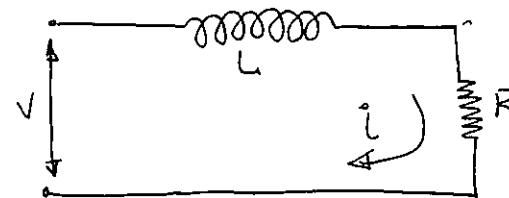
$$x(t) = \underline{\Phi}(t, t_0)x(t_0) + \int_{t_0}^t \underline{\Phi}(t, t_0) \underline{\Phi}^{-1}(s, t_0) G(s)u(s)ds$$

Lembrando que  $\underline{\Phi}^{-1}(s, t_0) = \underline{\Phi}(t_0, s)$  pela propriedade II e que  $\underline{\Phi}(t, t_0) \underline{\Phi}(t_0, s) = \underline{\Phi}(t, s)$  pela propriedade III, obtemos a expressão final para solução do sistema não homogêneo:

$$x(t) = \underline{\Phi}(t, t_0)x(t_0) + \int_{t_0}^t \underline{\Phi}(t, s)G(s)u(s)ds$$

Exemplo:

Considere o circuito composto de uma fonte de voltagem,  $v$ ; um resistor,  $R$ , e um indutor,  $L$ .



$$\text{Lei de Kirschoff : } v = iR + L \frac{di}{dt}$$

Admitamos:  $i = i_0$  em  $t = t_0 = 0$

$$\begin{cases} v = 0 & \text{para } t \leq 0 \\ v = V, \text{const.}, & \text{para } t > 0 \end{cases}$$

O sistema dinâmico fica, então, na forma

$$\frac{di}{dt} = -\frac{R}{L}i + \frac{V}{L}; \quad i(0) = i_0$$

e, identificando  $F = \left[ -\frac{R}{L} \right]$ , constante

de onde

$$\underline{\Phi}(t, t_0) = e^{-\frac{R}{L}(t-t_0)} \quad \text{para o sist. homogêneo}$$

Para a solução do sistema forçado:

$$i(t) = \underline{\Phi}(t, t_0)i(t_0) + \int_{t_0}^t \underline{\Phi}(t, s)G(s)u(s)ds$$

$$G(s) u(s) = \frac{V}{L} = \text{constante}$$

$$i(t) = e^{-\frac{R}{L}(t-t_0)} + \int_0^t e^{-\frac{R}{L}(t-s)} \frac{V}{L} ds$$

$$i(t) = e^{-\frac{R}{L}t} + \left[ -\frac{V}{R} \cdot \frac{1}{\frac{R}{L}} e^{-\frac{R}{L}(t-s)} \right]_0^t$$

$$i(t) = e^{-\frac{R}{L}t} + \frac{V}{R} [1 - e^{-\frac{R}{L}t}]$$

## 2.5. MODELOS DINÂMICOS DISCRETOS

Sistemas dinâmicos discretos ou sistemas de múltiplos estágios são aqueles descritos através de equações de diferenças ao invés de equações diferenciais.

A representação geral para um sistema discreto é dada pela forma:

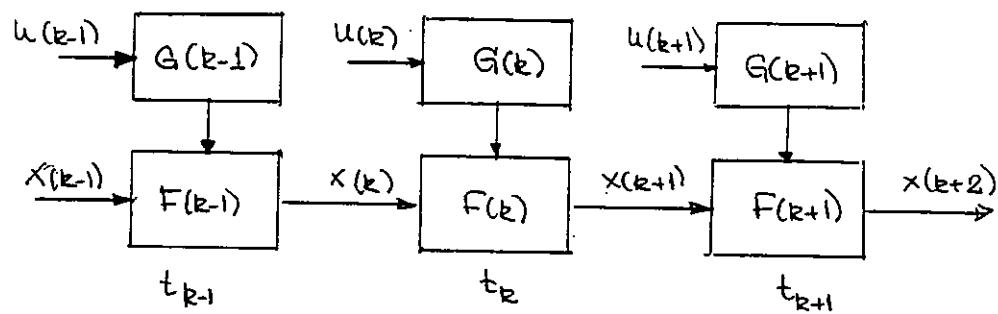
$$x(k+1) = f(x(k), u(k), t_k)$$

onde  $x$  é o vetor de estado,  $f$  um vetor de funções não-lineares,  $u$  é o vetor de entradas no sistema e  $k+1$  e  $k$  representam os instantes  $k+1$  e  $k$  do tempo em que o sistema é monitorado. A representação em sistemas discretos é uma forma muito conveniente para algumas aproximações estocásticas, como veremos no decorrer do curso.

Em particular, sistemas lineares discretos tomam a forma:

$$x(k+1) = F(k) x(k) + G(k) u(k)$$

e podem ser representados num diagrama de blocos como:



O mesmo desenvolvimento feito para as matrizes de transição dos sistemas lineares contínuos no tempo pode ser feito para os sistemas lineares discretos.

Consideremos o sistema linear discreto homogêneo

$$x(k+1) = F(k) x(k)$$

Esse sistema pode ser resolvido recursivamente desde que seja dado um valor inicial do estado. Assim, dado  $x(0)$ :

$$x(1) = F(0) x(0)$$

$$x(2) = F(1) x(1) = F_1 F(0) x(0)$$

⋮

$$x(k) = F(k-1) F(k-2) \dots F(1) F(0) x(0) = \underline{\Phi}(k, 0) x(0)$$

Definimos a matriz de transição do sistema discreto na forma:

$$\underline{\Phi}(k, l) = F(k-1) F(k-2) \dots F(l) \quad k > l$$

$$\underline{\Phi}(k, k) = I$$

Note, então, que a matriz  $F$  de definição do sistema é a própria matriz de transição entre estados consecutivos, isto é:

$$\underline{\Phi}(k+1, k) = F(k)$$

De modo análogo, podemos definir uma matriz fundamental de soluções para o sistema discreto:

$$\underline{\underline{X}}(k) = [\bar{x}^1(k), \bar{x}^2(k) \dots \bar{x}^n(k)] \quad (n \times n)$$

e mostrar que  $\underline{\underline{X}}(k)$  é não singular para todo  $k$  e também que

$$\underline{\underline{\Phi}}(k, \ell) = \underline{\underline{X}}(k) \underline{\underline{X}}^{-1}(\ell) \quad k \geq \ell$$

Para um sistema de múltiplos estágios invertente no tempo, i.e., tal que:

$$x(k+1) = A x(k)$$

a matriz de transição é calculada através de:

$$\underline{\underline{\Phi}}(k, 0) = A^k, \quad k \geq 0 \quad \text{ou} \quad \underline{\underline{\Phi}}(k+1, k) = A^{k+1-k} = A^1 = A$$

ou

$$\underline{\underline{\Phi}}(k, \ell) = A^{k-\ell}, \quad k \geq \ell$$

\* Consideramos agora o sistema discreto não homogêneo:

$$x(k+1) = \underline{\underline{F}}(k) x(k) + \underline{\underline{G}}(k) u(k)$$

A solução do sistema acima, dado o estado inicial  $x(0)$  e dadas as entradas  $u(k)$  é obtida através de:

$$x(k) = \underline{\underline{\Phi}}(k, 0) x(0) + \sum_{\ell=0}^{k-1} \underline{\underline{\Phi}}(k, \ell+1) \underline{\underline{G}}(\ell) u(\ell)$$

e as mesmas propriedades válidas para as matrizes ~~baseadas~~ de transição dos sistemas lineares contínuos permitem validadas para as matrizes de transição dos sistemas discretos.

## 2.6. EQUAÇÕES VARIACIONAIS DE 1ª ORDEM

A aproximação de resultados para sistemas não li-

nezes baseado na teoria de sistemas lineares representa a grande maioria das aplicações em Filtros de Kalman. Esses métodos estão sempre associados em equações variacionais de 1ª ordem que, em síntese, representam uma aproximação dos desvios de um sistema ~~linear~~ não linear em relação a uma trajetória gerada.

Para seu desenvolvimento, consideremos um sistema não linear dado por:

$$\dot{x}(t) = f[x(t), u(t), t]$$

com a condição inicial  $x(t_0) = x_0$  conhecida.

Imaginemos que uma solução nominal possa ser conhecida para o sistema acima. Gerando essa trajetória, podemos considerar perturbações em torno dessa solução, que chamaremos  $x_0(t)$ , provocado por mudanças nas condições iniciais ou por mudanças nos vetores de entrada.

Levando em conta, então, pequenas perturbações  $\Delta x_0$  nas condições iniciais e pequenos desvios  $\Delta u$  nos vetores de entrada, se admitirmos que esses fatores provocarão pequenas perturbações sobre a solução, podemos expandir a nova trajetória numa série de Taylor em torno da trajetória nominal  $x_0(t)$ . Tencemos, então, para a nova trajetória:

$$\dot{x}(t) = f[x(t), u(t), t]$$

$$\dot{x}(t) = f[x_0(t), u_0(t), t] + F(t)[x(t) - x_0(t)] + G(t)[u(t) - u_0(t)] + \mathcal{O}(2) + \dots$$

onde

$$F(t) = \left. \frac{\partial f}{\partial x} \right|_{x_0(t), u_0(t), t} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_n} & \dots & \frac{\partial f}{\partial x_n} \end{bmatrix}_{x_0(t), u_0(t), t} \quad (n \times n)$$

e:

$$G(t) = \frac{\partial f}{\partial u} \Big|_{x_0(t), u_0(t), t} = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial u_1} & \cdots & \frac{\partial f_n}{\partial u_m} \end{bmatrix}_{x_0(t), u_0(t), t}^{(n \times m)}$$

Desde que  $\dot{x}(t) = f[x_0(t), u_0(t), t]$  as equações acima:

$$\dot{x}(t) = \dot{x}_0(t) + F(t)[x(t) - x_0(t)] + G(t)[u(t) - u_0(t)] + O(2)$$

onde, desprezando-se os termos de ordem igual ou superior a 2, chega-se a:

$$\frac{d}{dt} [x(t) - x_0(t)] = F(t) [x(t) - x_0(t)] + G(t) [u(t) - u_0(t)]$$

Se admitirmos que a perturbação é pequena, o termo  $[x(t) - x_0(t)]$  confunde-se com a variação da função  $x(t)$  de 1ª ordem, isto é:

$$\delta x(t) \approx x(t) - x_0(t)$$

E, com isso, obtemos as equações variacionais de 1ª ordem do sistema não linear. Essas equações variacionais são então, um conjunto de n equações diferenciais lineares na forma:

$$\delta \dot{x}(t) = F(t) \delta x(t) + G(t) \delta u(t)$$

Exemplo:

Retornemos ao exemplo do satélite num órbita planar da seção 2.1. Tínhamos obtido o sistema não linear:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1(t)x_4^2(t) - \frac{Q}{x_1^2(t)} + u_r(t) \\ x_4(t) \\ -\frac{2}{x_1(t)}x_4(t)x_2(t) + \frac{1}{x_1(t)}u_t(t) \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}$$

As equações variacionais de 1ª ordem para esse sistema são obtidas por

$$F(t) = \frac{\partial f}{\partial x} \Big|_{x_0, u_0, t} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_4} \\ \vdots & & \vdots \\ \frac{\partial f_4}{\partial x_1} & \dots & \frac{\partial f_4}{\partial x_4} \end{bmatrix} \Big|_{x_0, u_0, t}$$

$$F(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \left[ x_{04}^2 + \frac{\partial G}{\partial x_1} \right] & 0 & 0 & 2x_{01}, x_{04} \\ 0 & 0 & 0 & 1 \\ \left[ \frac{2x_{02}x_{04}}{x_{01}^2} - \frac{u_{0t}}{x_{01}^2} \right] & -\frac{2x_{04}}{x_{01}} & 0 & -\frac{2x_{02}}{x_{01}} \end{bmatrix}$$

$$G(t) = \frac{\partial f}{\partial u} \Big|_{x_0, u_0, t} = \begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial u_t} \\ \vdots & \vdots \\ \frac{\partial f_4}{\partial u} & \frac{\partial f_4}{\partial u_t} \end{bmatrix}$$

$$G(t) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & \frac{1}{x_{01}} \end{bmatrix}$$

## 2.7. MÉTODOS DE INTEGRAÇÃO NUMÉRICA

Desde que muitas equações diferenciais podem ser resolvidas analiticamente, é usual aproximar as soluções desejadas numericamente.

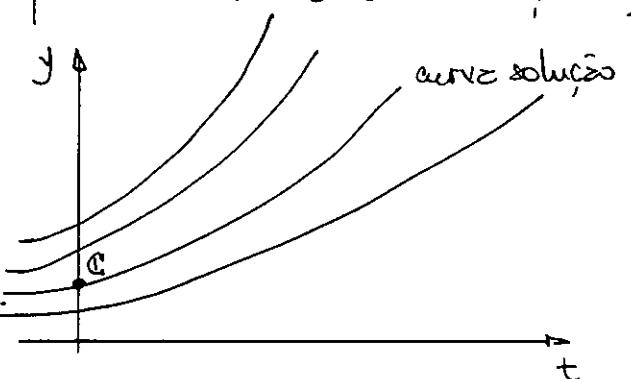
Assim, dada uma equação diferencial ordinária de 1ª ordem, escalar:

$$\dot{y} = f(y, t)$$

a solução é uma família de curvas  $y(t)$ , onde cada uma

condição inicial (valor inicial)  $y(0)$  fica automaticamente selecionada uma das curvas da família gerada.

Por exemplo, se  $\dot{y} = f(y, t) = y$  então  $y(t) = Ce^t$  é uma solução para uma dada condição  $y(0) = C$ .



O valor inicial da variável dependente pode ser especificado para qualquer valor da variável independente ( $t_0$ ). No entanto, vale-se admitir que foi feita sempre uma transformação de tal modo que  $t_0 = 0$ .

As mesmas considerações podem ser feitas para um sistema de equações diferenciais de 1ª ordem, como, por exemplo:

$$\begin{aligned}\dot{y} &= f(y, z, t) \\ \dot{z} &= g(y, z, t)\end{aligned}$$

na hipótese de  $\partial f/\partial y$ ,  $\partial f/\partial z$ ,  $\partial g/\partial y$  e  $\partial g/\partial z$  existirem em todo intervalo de integração e de serem conhecidas duas constantes de integração  $y(t_0)$ ,  $z(t_0)$ .

O problema de determinar os valores de  $y$  e  $z$  para  $t > t_0$  é chamado PROBLEMA DE CONDIÇÕES INICIAIS e sua solução numérica são necessárias nuzis duas informações:

o uma expressão do erro que se pode tolerar na solução;

o uma expressão do grau de se deseja pegar por tal solução.

Pretende-se abordar aqui os métodos passo a

passo, também chamados métodos de diferença ou métodos de variável discreta. Nesse caso, uma sequência de pontos discretos  $t_0, t_1, t_2, \dots$  é gerada, possivelmente com esforço constante variável,  $h_n = t_{n+1} - t_n$ . A cada ponto  $t_n$ , a solução  $y(t_n)$  é aproximada por um número,  $y_n$ , que é calculado a partir dos valores anteriores.

Um método de diferenças que fornece uma regra para calcular  $y_{n+1}$  usando  $k$  valores anteriores  $y_n, y_{n-1}, \dots, y_{n-k+1}$  é chamado método de  $k$ -passos. Se  $k=1$  então o método é dito de passo único. Se  $k>1$  então teremos um método de multipassos.

Exemplo de método de passo único é o método de Euler. Nesse caso, o valor  $y_{n+1}$  é calculado por uma extensão em linha reta a partir do ponto anterior,  $y_n$ . Então, a equação é do tipo

$$\dot{y} = f(y, t) \quad ; \quad y(t_0) = y_0$$

O método consiste em se fazer:

$$y(t_1) \approx y_1 = y_0 + h_0 f(y_0, t_0)$$

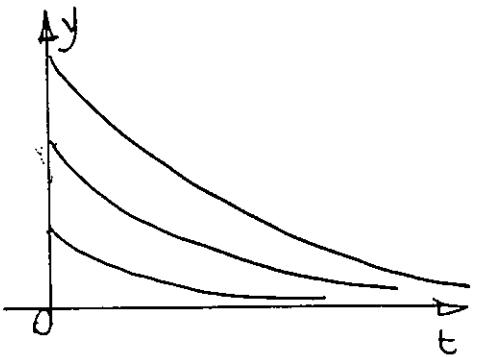
ié, retomar os 2 primeiros termos da expansão em série de Taylor e, sucessivamente:

$$y_{n+1} = y_n + h f(y_n, t_n)$$

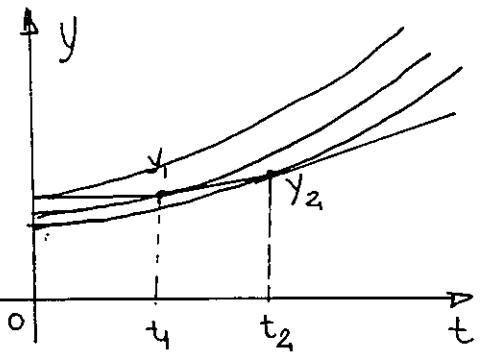
Examinemos a equação  $\dot{y} = \lambda y$ . Nesse caso, os erros iniciais na solução serão aumentados de  $e^{\lambda t}$  à medida que  $t$  cresce, como se verifica nas figuras. Então, se:

- \*  $\lambda \leq 0$  - os erros iniciais no uso do método não são aumentados  $\Rightarrow$  estabilidade da equação diferencial

\*\*  $\lambda > 0$  - os erros iniciais aumentam e a equação diferencial é ditz justável.



$$y' = -y ; \lambda = -1$$



$$y' = y , \lambda = 1$$

Os erros envolvidos na solução numérica de um problema de valores iniciais provem de duas fontes:

- a) erros de discretização ou erros de truncamento são propriedades diretamente ligadas ao método utilizado e, se, todos os cálculos no computador tiverem precisão infinita só teríamos erros de discretização;
- b) erros de arredondamento que são propriedade do computador e do programas utilizados.

Como a solução exata da equação diferencial geralmente não é conhecida e não pode ser calculada, o erro de discretização é usualmente estimado ou limitado.

O erro de discretização local é o erro cometido em um passo se todos os valores anteriores forem exatos e não houver erros de arredondamento. O erro local de discretização é então definido a partir de:

$$i_n = f(u_n, t) \quad ; \quad u_n(t_n) = y_n$$

$$d_n = y_{n+1} - u_n(t_{n+1})$$

onde  $y_{n+1}$  é a solução calculada pelo programa e  $u_n(t_{n+1})$  é a solução teórica a partir de  $u_n(t_n) = y_n$ .

O erro de discretização global é a diferença entre a solução computada, ignorando-se o arredondamento, e a solução verdadeira determinada pelas informações exatas, i.e:

$$e_n = y_n - u_n(t_n)$$

Essa distinção entre erros de discretização local e global torna-se clara quando  $f(y, t)$  não depende de  $y$  explicitamente. Nesse caso, a solução é simplesmente uma integral

$$y(t) = \int_{t_0}^t f(\tau) d\tau$$

$$\text{e } \int_{t_0}^{t_N} f(\tau) d\tau \approx \sum_{n=0}^{N-1} h_n f(t_n)$$

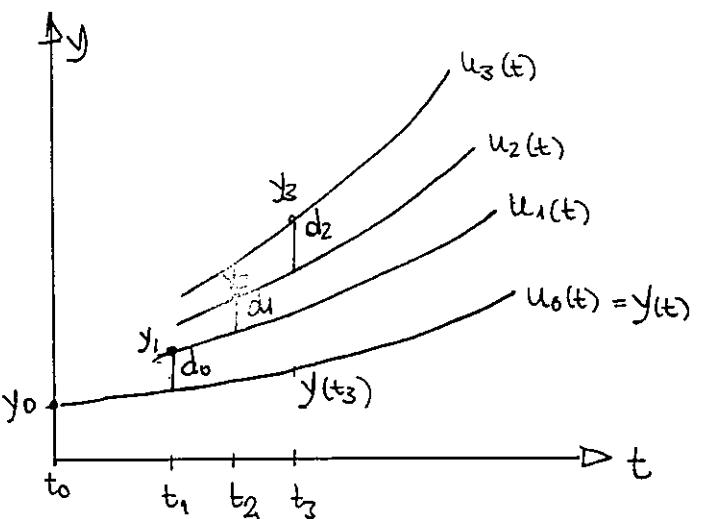
$$\text{Dai: } d_n = h_n f(t_n) - \int_{t_n}^{t_{n+1}} f(\tau) d\tau$$

$$e_n = \sum_{n=0}^{N-1} h_n f(t_n) - \int_{t_0}^{t_N} f(\tau) d\tau$$

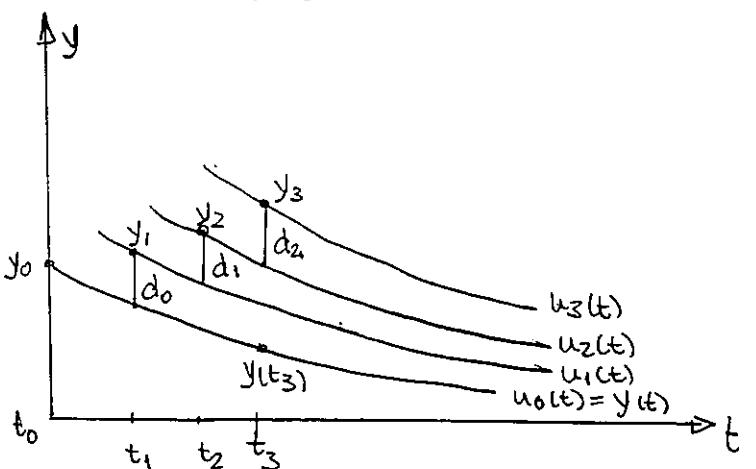
e, como cada uma das subintegrais é independente das demais, o erro global é a soma dos erros locais:

$$e_n = \sum_{n=0}^{N-1} d_n$$

No caso de uma  $f(y, t)$  dependente explicitamente de  $y$ , o erro em cada intervalo depende das soluções computadas nos intervalos anteriores. Consequentemente o erro global será geralmente maior do que a soma dos erros locais se a equação diferencial for instável e será menor no caso da equação diferencial ser estável, como mostrado nas figuras a seguir.



Equação diferencial instável  
 $e_3 = y_3 - y(t_3) > d_0 + d_1 + d_2$



Equação diferencial estável  
 $e_3 = y_3 - y(t_3) < d_0 + d_1 + d_2$

Verifica-se, portanto, que o efeito do erro local sobre o global é aproximadamente governado pelo sinal de  $\alpha f/y$ . Para os sistemas não lineares  $\alpha f/y$  pode mudar seu sinal e a equação é instável em algumas regiões e estável em outras, fato que torna ainda mais complexa a análise dos erros.

O erro de discretização local define a ordem do método numérico. Diz-se que um método é de ordem p se existir um número C tal que

$$|d_n| \leq C h_n^{p+1}$$

C pode depender das derivadas envolvidas na equação diferencial e do comprimento do intervalo, mas tem de ser independente do número de passos, n, e da tamanho do passo,  $h_n$ , isto é:

$$d_n = O(h_n^{p+1})$$

Como exemplo, tomemos o método de Euler. Temos:

$$y_{n+1} = y_n + h_n f(y_n, t_n)$$

$$u_n(t) = u_n(t_n) + (t - t_n) u'(t_n) + O[(t - t_n)^2]$$

$$u_n(t_{n+1}) = y_n + h_n f(y_n, t_n) + O(h_n^2)$$

$$u_n(t_{n+1}) = y_{n+1} + O(h_n^2)$$

$$d_n = y_{n+1} - u_n(t_{n+1}) = O(h_n^2)$$

Conclui-se, então, que o método de Euler é de primeira ordem, isto é,  $p=1$ .

Quanto ao erro global,  $\epsilon_N$ , pode ser expresso como soma de  $N$  erros locais somados a fatores descrevendo a estabilidade da equação, sendo  $N$  o número total de passos. Normalmente, então, se o erro local é  $O(h^{p+1})$  então o erro global será  $N \cdot O(h^{p+1}) = O(h^p)$

Outro tipo de erro que deve ser levado em consideração é o erro de arredondamento. Se admitirmos um erro extremo num erro de arredondamento é a cada passo pelo método de Euler, por exemplo, temos

$$y_{n+1} = y_n + h f(y_n, t_n) + \epsilon$$

e, consequentemente, um erro total de arredondamento  $N\epsilon$ , onde  $N$  é o número de passos.

No prática,  $\epsilon$  é um erro eletrônico comum, sendo seu efeito acumulado, erro total de arredondamento, estimado como  $\epsilon\sqrt{N}$  se a máquina arredonda ou uma função linear de  $\epsilon$  se a máquina truncar

Somando-se, então, os erros de arredondamento e de discretização global, temos uma expressão aproximada para o erro total de integração, na forma:

$$\text{erro total} \approx (t_f - t_0) \left( C h + \frac{\epsilon}{h} \right)$$

O erro total será minimizado por um valor ótimo de  $h$ , tamanho do passo, dado aproximadamente por

$$h_{\text{ótimo}} = \sqrt{\frac{\epsilon}{C}}$$

Finalmente, abordaremos neste tópico, os principais tipos de métodos de solução numérica das equações diferenciais. Existem pelo menos 3 grandes categorias em que estão divididos esses métodos: série de Taylor, Runge-Kutta e outros tipos.

Métodos de série de Taylor se baseiam na expansão da solução  $y(t)$  numa série de Taylor

$$y(t+h) = y(t) + h y'(t) + \frac{h^2}{2!} y''(t) + \dots$$

Se as derivadas de ordem maior puderem ser calculadas, um método de ordem  $p$  é obtido usando-se:

$$y_{n+1} = y_n + h y'_n + \frac{h^2}{2!} y''_n + \frac{h^3}{3!} y'''_n + \dots + \frac{h^p}{p!} y^{(p)}_n$$

O primeiro termo desprezado dê uma estimativa do erro de discretização local e pode ser usado para calcular o tamanho do passo,  $h$ .

As derivadas  $y'$ ,  $y''$ , etc podem ser expressas em

termos das derivadas parciais de  $f$ . Ao se fazer isso é importante distinguir as funções

$f(y, t)$  de duas variáveis judependentes  
e  $f(y(t), t)$  de uma única variável judependente obtida pelas substituições de  $y(t)$  em  $f$ .

No segundo caso, se tivermos

$$y'(t) = f(y(t), t)$$

a diferenciação total nos levará a

$$y'' = fy' y' + ft = fy f + ft = \left(\frac{\partial f}{\partial y}\right) f + \left(\frac{\partial f}{\partial t}\right)$$

$$y''' = fyy f^2 + f^2 f + fy ft + 2f_{ty} f + f_{tt}$$

etc.

e todas as derivadas aparecerão explicitamente em  $f$ .

Porz uma função de duas variáveis judependentes, como, por exemplo:

$$f(y, t) = y^2 + t^2$$

terímos

$$y' = y^2 + t^2$$

$$y'' = 2y^3 + 2yt^2 + 2t = 2yy' + 2t$$

$$y''' = 6y^4 + 8y^2t^2 + 4yt + 2t^2 + 2,$$

etc

e verifica-se que embora  $f$  seja um polinomio em duas variáveis, as derivadas totais se tornam mais complicadas à medida que a ordem cresce.

Por equações diferenciais baseadas em funções em que se possa desenvolver facilmente as derivadas totais, os métodos de série de Taylor podem ser bastante eficientes. Já para uma aplicação mais geral o seu uso é bem restrito.

Os métodos Runge-Kutta são deduzidos de

modo a aproximar o método de série de Taylor sem precisar de definições explícitas ou cálculos de derivadas além da primeira. Essa aproximação é conseguida à custa de várias avaliação de função no intervalo de interesse. Por exemplo, o Runge-Kutta clássico de 4<sup>º</sup> ordem é dado por:

$$y_{n+1} = y_n + \frac{1}{6} (k_0 + 2k_1 + 2k_2 + k_3)$$

onde:

$$k_0 = h f(y_n, t_n)$$

$$k_1 = h f\left(y_n + \frac{1}{2} k_0, t_n + h/2\right)$$

$$k_2 = h f\left(y_n + \frac{1}{2} k_1, t_n + h/2\right)$$

$$k_3 = h f(y_n + k_2, t_n + h)$$

O método clássico de Runge-Kutta pode ser visto como uma extensão, para equações diferenciais, da regra de quadratura de Simpson:

$$\int_{t_n}^{t_{n+1}} f(\alpha) d\alpha \approx \frac{h_n}{6} \left[ f(t_n) + 4 f\left(\frac{t_n + t_{n+1}}{2}\right) + f(t_{n+1}) \right]$$

A medida que  $h \rightarrow 0$  o método acima converge assintoticamente, até termos de ordem de  $h^4$ , para a série de Taylor. Nesse tipo de método não se encontra disponível uma estimativa do erro de discretização local para ajudar na escolha do tamanho do passo, mas tem várias vantagens:

- facilmente codificáveis;
- numericamente estáveis para uma grande classe de problemas;
- o método se autojuiciiza, uma vez que se precisa apenas  $y_n$  para calcular  $y_{n+1}$ ;
- o tamanho do passo,  $h_n$ , pode ser alterado em qualquer posição durante a integração.

Métodos multipassos são aqueles que utilizam as informações nos pontos anteriores,  $y_{n-1}, y_{n-2}, \dots$ , etc e  $f_{n-1}, f_{n-2}, \dots$ , etc onde  $f_i = f(y_i, t_i)$  faz calcular o valor de  $y_{n+1}$ , sendo muito eficientes. Para obter altas precisões requerem, geralmente, menos esforço que os métodos de passo único e é quase sempre possível obter uma estimativa do erro de discretização.

A ordem de um método multipasso pode ser selecionada automaticamente e mudada dinamicamente, tendo-se assim métodos que trabalham numa grande variedade de problemas. De mesma forma, uma programação apropriada de um método multipasso pode fornecer saídas em pontos arbitrários sem alterar o valor do passo,  $h$ . Além disso, alguns tipos de erros podem ser detectados durante a solução do problema.

Todas essas vantagens são obtidas às custas de um aumento na complexidade do programa e, em alguns casos, na possibilidade de instabilidade numérica.

Métodos multipasso lineares podem ser considerados casos especiais da fórmula geral:

$$y_{n+1} = \sum_{i=1}^k \alpha_i y_{n+1-i} + h \sum_{i=0}^{k-1} \beta_i f_{n+1-i}$$

onde  $k$  é um número intiro especificado e  $\beta_k$  e  $\alpha_k$  não são simultaneamente nulos. A fórmula acima define um método de  $k$ -passos linear geral, sendo chamado linear porque cada  $f_i$  entra linearmente, podendo ser ou não uma função linear nos seus argumentos. Depois que o método foi iniciado, cada passo requer o cálculo de  $y_{n+1}$  a partir dos valores conhecidos  $y_n, y_{n-1}, \dots, y_{n-k+1}; f_n, f_{n-1}, \dots, f_{n-k+1}$ . Separar-se, então em 2 tipos de métodos:

- Métodos explícitos: se  $\beta_0 = 0$  então o cálculo é

direto, pois não há necessidade de se calcular  $f_{n+1}$ .

- Métodos implícitos: se  $f_0 \neq 0$  é necessário o cálculo de  $f_{n+1}$  para gerar  $y_{n+1}$ . Claramente, esse fato acrescenta uma dificuldade ao método que é justificado por suas melhores propriedades.

Geramente dois métodos multipassos são usados conjuntamente no cálculo de cada passo da solução. Um método explícito chamado previsor é seguido por uma ou mais aplicações de um método implícito chamado corretor, daí originando o nome métodos de previsão-correção. Um bom exemplo de método desse tipo é o método de quarta ordem de Adams onde:

$$\text{previsor: } y_{n+1} = y_n + \frac{h}{24} (55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3})$$

$$\text{corretor: } y_{n+1} = y_n + \frac{h}{24} (9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2})$$

Ambas as fórmulas, ~~são~~ quarto ordenem, são casos particulares de uma família de previsores-corretores de ordem variável. O esquema de previsão-correção para o cálculo de  $y_{n+1}$  é dado por:

Passo 1) Use o previsor para calcular  $y_{n+1}^{(0)}$ , aproximação inicial para  $y_{n+1}$ . Faça  $i=0$ ;

Passo 2) Avlie a função derivada e faça  
 $f_{n+1}^{(i)} = f(y_{n+1}^{(i)}, t_{n+1})$

Passo 3) Calcule uma melhor aproximação para  $y_{n+1}^{(i+1)}$  usando a fórmula do corretor com

$$f_{n+1} = f_{n+1}^{(i)}$$

Passo 4) Se  $|y_{n+1}^{(i+1)} - y_{n+1}^{(i)}| >$  Tolerância de convergência, então incrementa  $i$  de 1 e volta para o passo 2, senão faz  $y_{n+1} = y_{n+1}^{(i+1)}$

Existem três processos ocorrendo separadamente:

- \* o passo de previsão, 1, ( $P$ )
- \*\* o passo de avaliação, 2, ( $E$ )
- \*\*\* o passo de correção, 3, ( $C$ )

O passo 4 pode ser substituído por um passo que requer exatamente  $m$  iterações do corrector. Nesse caso, o método resultante é conhecido como  $P(EC)^m$ . Algumas considerações empíricas, relativas aos métodos de previsão-correção:

- a) Vale a pena se fazer um passo final de avaliação para obter um valor mais preciso de  $f$ . Método resultante é conhecido como  $P(EC)^m E$ ;
- b) Se a ordem do previsor é tão maior que a do corrector, então corrigir mais que duas vezes tem pouquíssimo efeito;
- c) Se, nos métodos de Adams, os erros de discretização e de estabilização são considerados, então, um esquema  $P(EC)^m$  não é tão bom quanto  $P(EC)^{m-1} E$  e  $P(EC)^1 E$  é mais estável que  $P(EC)^2 E$ . Desse modo, para o método de Adams, o melhor esquema é  $PEC E$ .

Finalmente, uma breve descrição sobre as subrotinas RKF45 e ODE.

A subrotina RKF45 é uma subrotina para resolver problemas de valores iniciais baseada em fórmulas de Runge-Kutta desenvolvidas por Fehlberg e implementadas por Shampine e Watts.

A subrotina exige seis avaliaçãoes da função por passo. Quatro desses valores são combinados com um conjunto de coeficientes para produzir um método de 4<sup>o</sup> ordem, enquanto todos os seis valores gerados são combinados com outro conjunto de coeficientes para produzir um método de quinta ordem. Uma comparação entre os dois valores obtidos gera uma estimativa do erro que é usada para controle do passo.

As seis avaliações da função por passo são feitas por:

$$k_i = h_n f \left( y_n + \sum_{j=1}^{i-1} \beta_{ij} z_j ; t_n + d_i h_n \right), \quad i=1, 2, \dots, 6$$

e o novo valor  $y_{n+1}$  é obtido usando-se uma combinação linear dos seis  $k_i$ 's:

$$y_{n+1} = y_n + \sum_{i=1}^6 \gamma_i k_i$$

Os coeficientes  $\beta_{ij}, d_i, \gamma_i$  se encontram tabelados em Forsythe: "Computer Methods for Mathematical Computations", Prentice Hall, 1977, pg. 131.

A subrotina ODE, que utiliza um método de Adams de previsão-correção, se encontra desenvolvida no livro de Shampine - Gordon: "Computer Solution of Ordinary Differential Equations", W.H. Freeman and Co., 1975.

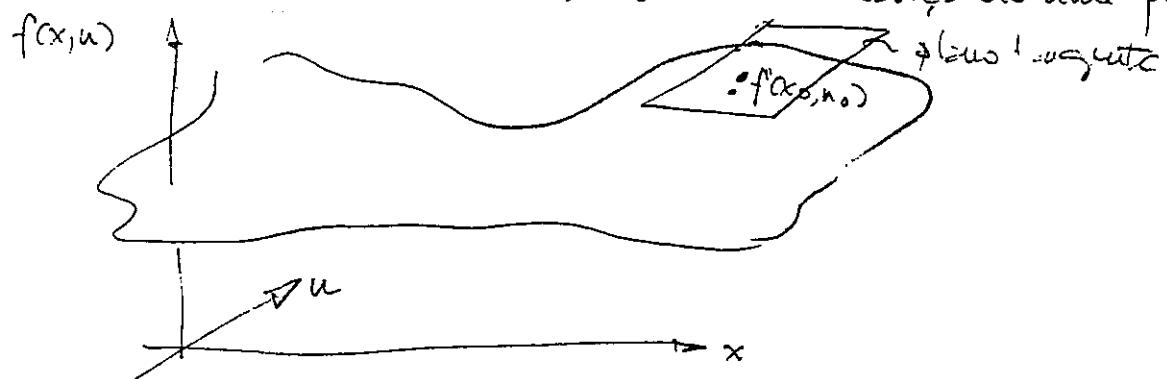
A ODE é uma subrotina extremamente sofisticada onde para atingir a precisão desejada de erro local é feita automaticamente uma decisão de ou aumentar a ordem do previsor e do corretor ou alterar o tamanho do passo. Esta seleção é feita automaticamente através de uma função de ajuste. A alteração da ordem do previsor e do corretor é também feita

ta de meiairo automática por uma geração recursiva dos coeficientes envolvidos nas fórmulas do previsor e do corretor, podendo ser atingida até ordem 14 na geração do previsor e do corretor.

## TÉCNICAS DE LINEARIZAÇÃO

### 1. Linearizações por Expansão em Série de Taylor

- Modelo linear válido na vizinhança de um ponto de operação



$$\dot{x} = f(x, u)$$

$x$  - estado  
 $u$  - controle

$$f(x, u) = \tilde{f}(x_0, u_0) + \frac{\partial f}{\partial x}(x_0, u_0)(x - x_0) + \frac{\partial f}{\partial u}(x_0, u_0)(u - u_0) + \text{l.o.sup.}$$

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

$$B = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial u_1} & \cdots & \frac{\partial f_n}{\partial u_m} \end{bmatrix}$$

Sistema linear approximado

$$\dot{z} = Az + Bu$$

$$z = x - x_0 \quad ; \quad v = u - u_0$$

Exemplo onde a técnica falha

Sist. rígido não se refere, corrigido por ajetos

esperar

$$\dot{x}_1 = a_1 x_2 x_3 + b_1 u$$

$$\dot{x}_2 = a_2 x_1 x_3 + b_2 u$$

$$\dot{x}_3 = a_3 x_1 x_2$$

$$\text{com } |a_{ij}| \leq 1, \quad a_3 \neq 0, \quad a_1 b_2^2 + a_2 b_3^2$$

Se o sistema é linearizado em torno de  $(0, 0, 0, u)$

$$\dot{x} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} u \Rightarrow \text{Não Controlável}$$

Problemas

- aqui o tamanho da vizinhança em que o modelo vale
- não dá para modelar não linearidades descontínuas
- forçar a barra pode levar a problemas sérios: bifurcações ou comportamentos caóticos

## 2. Linearização Global

Ferramenta: Geometria Diferencial (Brockett, 1978)  
sistemas SISO

$$\dot{x} = f(x) + g(x) u \quad (1)$$

Ideia: transformar (1) num sistema linear em  $z$ , projetar o controlador em  $z$ , voltar para  $x$ , obtendo o controlador não-linear.

seja  $\dot{z} = a + bv$  (2) a ser obtido a partir de:  $z = T(x)$  (3)

$$\text{de (3)} \quad \frac{dz}{dt} = \frac{dT}{dx} \frac{dx}{dt} \quad (4)$$

$$\dot{z} = \frac{dT}{dx} (f(x) + g(x) u) \quad (5)$$

Comparando (5) e (2)

$$\frac{dT}{dx} f(x) = a \quad (6)$$

$$\frac{dT}{dx} g(x) = bv \quad (7)$$

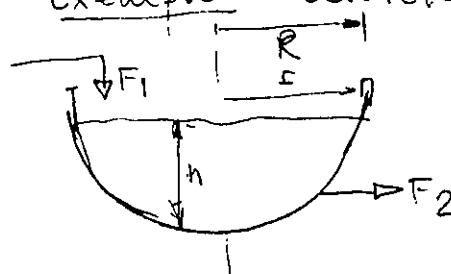
$T(x)$  é resolvida por integração a partir de (6).

A lei de controle não linear é obtida de (6) e (7)

$$u = \bar{T}(x, v) = \frac{bv f(x)}{a g(x)}, \quad a g(x) \neq 0$$

$v$  é uma lei de controle linear ( $PID$ ,  $PI$ , ... etc) para que o sistema (5) satisfaça os objetivos de controle desejados.

Exemplo: Controle de nível num tanque hemisférico



$$x = h$$

$$b = F_1$$

$$f(x) = -\frac{kx^{1/2}}{2Rx - x^2}, \quad k \text{ const}$$

$$g(x) = \frac{1}{\pi(2Rx - x^2)}$$

$$\frac{dT}{dx} f(x) = \frac{dT}{dx} \left[ \frac{-kx^{1/2}}{2Rx - x^2} \right] = a \Rightarrow T(x) = k_1 x^{3/2} \left[ \frac{4R}{3} - \frac{2x}{5} \right]$$

$$u = \frac{bv f(x)}{a g(x)} = k_2 x^{1/2} v, \quad ; \quad k_1 = -\frac{a}{k} ; \quad k_2 = \frac{\pi}{k_1}$$

$\check{v}$  é projetado por um PI, por exemplo:

$$v = k_C \left[ (z_r - z) + \frac{1}{\theta} \int_0^t (z_r - z) dt \right]$$

$$\Rightarrow \text{Polinômio característico: } s^2 + b k_C s + \frac{b k_C}{\theta} = 0$$

### SISTEMAS SISO

$$\dot{x} = f(x) + g(x)u \quad x \text{ é } nx1, u \text{ é escalar}$$

Problema: achar cond. suf. para  $f, g$  e  $u$  para que exista

$T = T(T_1, \dots, T_{n+1})$  t.q novas variáveis de estado  $z_i = T_i(x)$  e de controle  $v = T_{n+1}(x, u)$  formam sistema linear:

$$\dot{z}_1 = z_2$$

$$\dot{z}_2 = z_3$$

:

$$\dot{z}_n = v$$

Nesse caso: 1)  $T_i(0) = 0$

2) Jecobi de  $T_1, T_n$  é não singular

3)  $T_{n+1}$  pode ser invertida como função de  $u$

4)  $T_1, \dots, T_{n+1}$  são fracionárias

$$\dot{T}_1 = T_2, \dot{T}_2 = T_3, \dots, \dot{T}_n = T_{n+1}$$

cond. Necessária e suficiente (Hunt, Su, Meyer, 1983)

1) A matriz de controllabilidade

$$\begin{pmatrix} q_1 & \text{adj}_1 q_1 & \dots & \text{adj}_{n-1} q_1 & q_1 \end{pmatrix} \text{ é não singular}$$

2) O conjunto de caudas velozes

$$\{q_1, (\text{adj}_1 q_1), \dots, (\text{adj}_{n-1} q_1), q_1\} \text{ é involutivo}$$

Involutivo significa que o conjunto é integrável

### Proposição

Se  $T_i(x) = \bar{z}_i$

$$\dot{\bar{z}}_i = \frac{\partial T_i}{\partial x} \dot{x} = \frac{\partial T_i}{\partial x} [f(x) + g(x)u] = \bar{T}_{i+1}$$

Então  $\bar{T}_1, \dots, \bar{T}_n$  não dependem de  $u$ :

$$\frac{\partial T_i}{\partial x} q^{nr} = 0$$

$$\frac{\partial T_i}{\partial x} f(x) = T_{i+1} \quad i = 1, 2, \dots, n-1$$

$$\frac{\partial T_n}{\partial x} [f(x) + g(x)u] = T_{n+1}$$

Resulta o seguinte sistema de E.D.P's:

$$\left\{ \begin{array}{l} \frac{\partial T_i}{\partial x_1} g_1 + \dots + \frac{\partial T_i}{\partial x_n} g_n = 0 \\ \frac{\partial T_i}{\partial x_1} f_1 + \dots + \frac{\partial T_i}{\partial x_n} f_n = T_{i+1} \quad i=1,2,\dots,n-1 \\ \frac{\partial T_n}{\partial x_1} (f_1 + u g_1) + \dots + \frac{\partial T_n}{\partial x_n} (f_n + u g_n) = T_{n+1} \end{array} \right.$$

Parêntesis: Conceitos de Geometria Diferencial

$$\dot{x} = f(x) + g(x)u$$

$y = h(x)$  escalar

$$\text{Gradiente de } h : dh = \left[ \frac{\partial h}{\partial x_1}; \frac{\partial h}{\partial x_2}; \dots; \frac{\partial h}{\partial x_n} \right] = \frac{\partial h}{\partial x}$$

Derivada de Lie de uma função escalar  $h$  em relação a um campo vetorial  $h$  é dada por:

$$L_f h = \langle dh, f \rangle = \sum_{i=1}^n \frac{\partial h}{\partial x_i} f_i(x)$$

Derivadas de Lie de ordem superior

$$L_f^k h = L_f(L_f^{k-1}, h) = \langle dL_f^{k-1} h, f \rangle, L_f^0 h = h$$

Colchete de Lie de  $f$  e  $g$ :

$$[f, g] = \frac{\partial g}{\partial x} f - \frac{\partial f}{\partial x} g$$

$\frac{\partial f}{\partial x}, \frac{\partial g}{\partial x}$  são matrizes  $n \times n$

Interpretarlos como uma nova direção ao longo da qual a trajetória de estados pode-se mover com sua lei de contrapropriedade.

Colchete sucessivo de Lie (direções adicionais)

$$(ad_f^0, g) = 0$$

$$(ad_f^1, g) = [f, g]$$

$$(ad_f^2, g) = [f, [f, g]]$$

$$(ad_f^k, g) = [f, (ad_f^{k-1}, g)]$$

Voltando ao sistema de EDDP's:

$$\begin{cases} \langle dT_i, g \rangle = 0 \\ \langle dT_i, f \rangle = L_f T_i = T_{i+1} \quad i = 1, 2, \dots, n-1 \\ \langle dT_n, f + gu \rangle = L_f + u g T_n = T_{n+1} = v \end{cases}$$

Mostrar-se que a transformação  $T$  fica determinada se se encontre  $T_1$ , tal que:

$$\begin{cases} T_1(0) = 0 \\ \langle dT_1, (\text{adj}_f^k, g) \rangle = 0 \quad k = 0, 1, \dots, n-2 \\ \langle dT_1, (\text{adj}_f^{n-1}, g) \rangle \neq 0 \end{cases}$$

As  $T_i$  restantes se obtêm por:

$$T_{n+1} = \langle dT_1, f \rangle = L_f T_1$$

Exemplo

$$f(x) = \begin{bmatrix} \sin x_2 \\ 0 \end{bmatrix} \quad g(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} \dot{x}_1 = \sin x_2 \\ \dot{x}_2 = 1 \cdot u \end{bmatrix}$$

Matriz de Controllabilidade

$$C = \begin{bmatrix} 0, (\text{adj}_f, g) \end{bmatrix} \Rightarrow C = \begin{bmatrix} 0 & -\cos x_2 \\ 1 & 0 \end{bmatrix}$$

seja  $C \neq 0$  e  $-\pi/2 \leq x_2 \leq \pi/2 \Rightarrow$  sistema controlável

$T_1 = x_1$  (necessário)

$$\langle dT_1, g \rangle = 0, \quad \langle dT_1, (\text{adj}_f, g) \rangle \neq 0$$

$$\begin{cases} \frac{\partial T_1}{\partial x_2} = 0 \\ \frac{\partial T_1}{\partial x_1} \cos x_2 \neq 0 \end{cases}$$

$T_1 = x_1$  é solução em  $-\pi/2 \leq x_2 \leq \pi/2$  é sólida

$$T_1 = x_1$$

$$T_2 = L_f T_1 = \sin x_2$$

$$T_3 = L_f + u g, T_2 = \sin \cos x_2 = v$$

Sistema Linear Equivalente

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_2 = v \end{cases}$$

$$\begin{cases} z_1 = x_1 \\ z_2 = \sin x_2 \\ v = u \cos x_2 \end{cases}$$

Sistemas transformados

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} v$$

Com  $v = -2z_1 - 2z_2 = -2(z_1 + z_2) \Rightarrow$  polos de mtlz fechados são  
 $\lambda = -1 \pm i$

Então  $\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = A_0 z$

Equação de Lyapunov

$$A_0^T P + P A_0 = Q = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$P = \begin{bmatrix} 5/4 & 1/4 \\ 1/4 & 3/8 \end{bmatrix}$$

$V(z) = z^T P z = \frac{5}{4} z_1^2 - \frac{1}{2} z_1 z_2 + \frac{3}{8} z_2^2$  é uma função de Lyapunov

para  $z = A_0 z$  na origem

Lei de controle não linear

$$u \cos x_2 = -2z_1 - z_1 z_2$$

$$u \cos x_2 = -2x_1 - 2 \sin x_2$$

$$u = -\frac{2x_1}{\cos x_2} - 2 \operatorname{tg} x_2$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \sin x_2 \\ -\frac{2x_1}{\cos x_2} - 2 \operatorname{tg} x_2 \end{bmatrix}$$

$V(x) = \frac{5}{4} x_1^2 + \frac{1}{2} x_1 \sin x_2 + \frac{3}{8} \sin^2 x_2$  é uma função de Lyapunov

para a origem e pode-se determinar o domínio de estabilidade  
 dos pontos de equilíbrio considerados

## SISTEMAS MULTIVARIÁVEIS

$$\dot{x} = f(x) + \sum_{i=1}^m g_i(x) u_i, \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m$$

Nesse caso  $T(x, u)$  é um vetor de  $(n+m)$  transformações contínuas e inversíveis com  $T(x(0), 0) = 0$  e  $x, T_1, T_2, \dots, T_n$  funções de  $x$  e  $T_{n+1}, \dots, T_{n+m}$  funções de  $x$  e  $u$ .

Problema: gerar as transformações do sistema original para uma forma canônica

$$O_1 = k_1, \quad \dot{z}_1 = z_2$$

$$\dot{z}_2 = z_3$$

$$\vdots$$

$$\dot{z}_{\sigma_1} = u_1$$

$$O_2 = k_2 + O_1, \quad \dot{z}_{\sigma_1+1} = z_{\sigma_1+2}$$

$\vdots$

$$\dot{z}_{\sigma_2} = u_2$$

$$O_m = k_m + O_{m-1}, \quad \dot{z}_{\sigma_{m-1}+1} = z_{\sigma_{m-1}+2}$$

$$\dot{z}_{\sigma_m} = u_2$$

$k_i$  - índice de controlabilidade dos subsistemas

$$O_1 = k_1; O_2 = k_1 + k_2, \dots, O_m = k_1 + k_2 + \dots + k_m = n$$

Existem condições necessárias e suficientes para este teste:  
Luenberger, 1969, Meyer, 1985.

### DESCRIÇÃO

- Condições necessárias e suficientes são mais restritivas e difíceis de obter
- A solução do sistema de EDP's, se existir, pode ser unica
- Se as EDP's têm soluções deve-se verificar o domínio das transformações; se perde de um ponto de operação, pode não servir para nada
- Lei de controle faz ação de comando e seu vizibilidades de implementação
- Etc., etc ...

## 2. Linearizações Entrada/Saída

### 2.1. SISO

$$\dot{x} = f(x) + g(x)$$

$$y = h(x)$$

$x \in \mathbb{R}^n$ ,  $u$  e  $y$  são escalares

Derivando  $y$  no tempo:

$$\dot{y} = \frac{\partial h}{\partial x} \dot{x} = \frac{\partial h}{\partial x} (f(x) + g(x)u)$$

$$\ddot{y} = L_f h + L_g h u$$

$L_f h$ ,  $L_g h$  são os derivados de Lie com relação a  $f$  e  $g$

Se  $L_g h \neq 0 \Rightarrow u = \frac{1}{L_g h} (v - L_f h)$  produz o sistema linear  $\ddot{y} = v$

Se  $L_g h = 0$  então  $\ddot{y} = L_f^2 h + L_g (L_f h) u$  e  $\ddot{y} = L_g (L_f h) u \neq 0$

$$u = \frac{1}{L_g (L_f h)} (v - L_f^2 h) \text{ produz } \ddot{y} = v$$

Se  $r$  é o menor inteiro positivo tq  $L_g (L_f^k h) = 0$ ,  $k=0, 1, \dots, r-2$  e  $L_g (L_f^{r-1} h) \neq 0$  então

$$u = \frac{1}{L_g (L_f^{r-1} h)} (v - L_f^r h) \text{ produz } y^{(r)} = v$$

Índice de linearizabilidade é o menor  $r$  pq  $L_g (L_f^r h) \neq 0$

Deixa definição:  $\frac{d^k y}{dt^k} = L_f^k h \quad k=1, \dots, r-1$

$$\frac{d^r y}{dt^r} = L_f^r h + L_g (L_f^{r-1} h) u$$

A dinâmica linear desejada é dada por (não escolhida)

$$\frac{d^r y}{dt^r} + \beta_{r-1} \frac{d^{r-1} y}{dt^{r-1}} + \dots + \beta_0 y = v$$

$$\Rightarrow \frac{d^r y}{dt^r} = v - \sum_{k=0}^{r-1} \beta_k \frac{d^k y}{dt^k} \quad e:$$

$$L_f^r h + L_g (L_f^{r-1} h) u = v - \sum_{k=0}^{r-1} \beta_k L_f^k h$$

Portanto

$$u = \frac{v - L_f^r h - \sum_{k=0}^{r-1} \beta_k L_f^k h}{\lg(L_f^{r-1} h)}$$

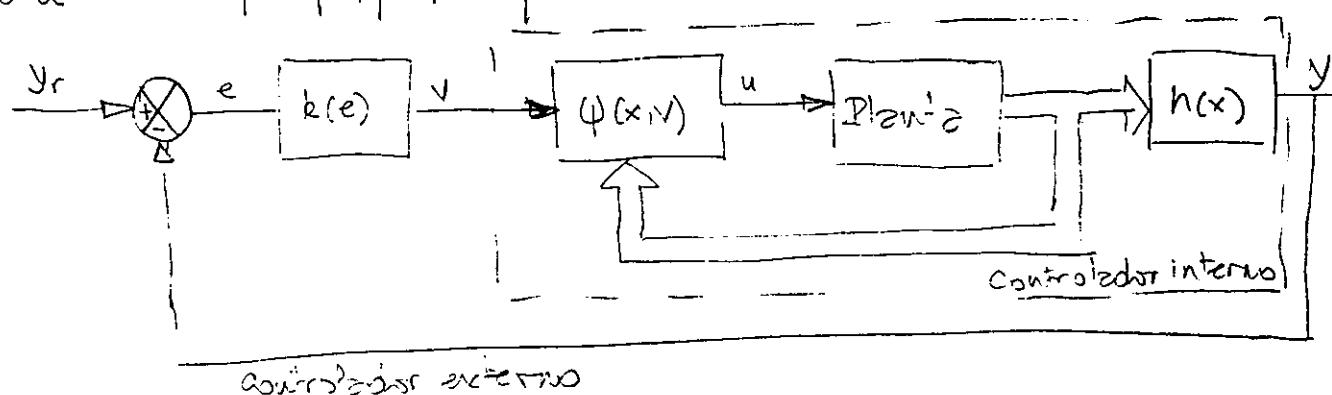
$$u = \Psi(x, v) = \frac{v - k(x)}{\lambda(x)} \quad k(x) = \sum_{k=0}^r \beta_k L_f^k h$$

$$\lambda(x) = \lg(L_f^{r-1} h)$$

transforma o sistema original no sistema linear

$$\sum_{k=0}^{r-1} \beta_k \frac{d^k y}{dt^k} = v = \mathcal{L}(x, u)$$

com os coefs  $\beta_0, \beta_1, \dots, \beta_{r-1}$  selecionados arbitrariamente



2.2.  $M \times N$

$$\dot{x} = f(x) + g_1(x)u_1 + \dots + g_m(x)u_m$$

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$g_i: \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\begin{aligned} x &\in \mathbb{R}^n \\ u &\in \mathbb{R}^m \\ y &\in \mathbb{R}^m \end{aligned}$$

interessando os indices  $i_j$  no termo:

$$\dot{y}_j = \frac{\partial h}{\partial x} [f(x) + \sum g_i(x)u_i]$$

$$y_j = l_{ij}u_i + \sum_{i=1}^m (l_{ij})u_i \quad (2.2.1)$$

Se  $\text{cod}(l_{ij}, h_j) = 0$  então a entrada não aparece em (2.2.1)

Índice de linearização  $r_j$  é o menor intiro q. s.t. umas das entradas aparece em  $l_{ij}^{(r_j)}$

$$y_j^{(r_j)} = L_f^{r_j} h_j + \sum_{i=1}^m l_{ij} (L_f^{r_j-1} h_j) u_i \quad (2.2.2)$$

com ao menos uma das  $l_{ij}(L_f^{r_j-1} h_j) \neq 0$

Se o sistema é linearizado para cada  $y_j$  existe um  $r_j$

$A(x)$ , mxm é dada por:

$$A(x) = \begin{bmatrix} L_{g_1}(L_f^{r_1} h_1) & \dots & L_{g_m}(L_f^{r_1} h_m) \\ \vdots & & \vdots \\ L_{g_1}(L_f^{r_{m-1}} h_m) & \dots & L_{g_m}(L_f^{r_{m-1}} h_m) \end{bmatrix}$$

Isto permite escrever (D.D.2) como

$$\begin{bmatrix} y_1^{(r_1)} \\ \vdots \\ y_m^{(r_m)} \end{bmatrix} = \begin{bmatrix} L_f^{r_1} h_1 \\ \vdots \\ L_f^{r_m} h_m \end{bmatrix} + A(x) \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$$

Se  $A(x)$  é não singular obtém - se  $u$ :

$$u = -\Delta^{-1}(x) \begin{bmatrix} L_f^{r_1} h_1 \\ \vdots \\ L_f^{r_m} h_m \end{bmatrix} + \Delta^{-1}(x) v$$

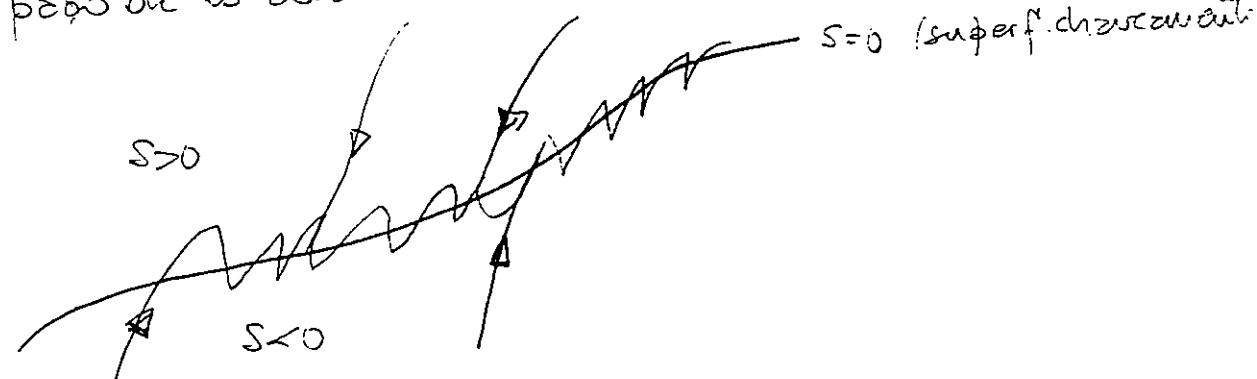
e tem - se o sistema decoplado:

$$\begin{bmatrix} y_1^{(r_1)} \\ \vdots \\ y_m^{(r_m)} \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}$$

$A(x)$  é chamada MATRIZ DE DESACOPLAMENTO

### 2.3. CONTROLE DE SISTEMAS LINEARIZADO E/S POR CONTROLE DE ESTRUTURA VARIÁVEL (SLIDING MODE)

Controladores de estrutura variável são caracterizados por lei de controlo de controlo que produz deslocamentos para que as trajetórias permaneçam sobre uma órbita superfície, no espaço de estados.



A estrutura do controlador é variado intencionalmente para se obter a dinâmica desejada. Na 1ª fase (fase de alcance), as trajetórias vêm de uma círculo arbitrário para a superfície  $s$ . Na fase seguinte (de deslizamento ou sliding mode) o movimento restringe - se à superfície.

Na fase de deslizamento, a dinâmica é determinada pelos parâmetros de projeto e pelas eqs. da superfície, sendo insensível às variações dos parâmetros e perturbações no sistema. Problema é que o sliding mode ideal não pode ser alcançado pelas interfaces do sistema de controle (descontinuidades, saltos)  $\Rightarrow$  "chattering"

Problemas a resolver:

- controle para cruzar a superfície de chavreamento
- redução do "chattering"

SISO (Exemplo)

$$\dot{x} = f(x) + g(x) u$$

$$y = h(x)$$

Objetivo: Fazer  $y$  compensar entrada  $y_d(t)$  (tracking)

Supõe-se sistema linearizado em torno da órbita  $\Sigma$

Definição

$$S(x) = \sum_{k=0}^{m-1} \beta_k e^{(k)}$$

$$e^{(k)} = \frac{d^k e}{dt^k}, \quad e \text{ é o erro } e = y_d - y$$

$S$  deve ser  $\geq 0$ , o que pode ser feito se  $S^2$  é convexa

$$1. \frac{1}{2} S^2 \geq 0$$

$$\Leftrightarrow \frac{1}{2} \frac{d}{dt} (e^T e) \geq 0 \Leftrightarrow \ddot{e} \geq 0$$

Nesse caso  $\dot{S} = -\eta \operatorname{sgn}(s)$  se torna  $\leq 0$

e mostrase que a  $s=0$  é a interface de chavreamento

$$S = \beta_0 (y_d - y)$$

$$\text{e a lei de controle é } u = \frac{1}{\beta_0 h} [y_d - \beta_0 y - \eta \operatorname{sgn}(s)]$$

### 3. Linearização Estendida

Correspondente ao "gain-chattering": obtém-se um conjunto de sistemas linearizados (por Taylor) distintos e diferentes portanto operacionais. Una vez consolidada a convergência do controlador, zera

os ganhos do sistema. Depois, mediante ajuste de curvas,  
geram-se expressões auxiliares para os ganhos, em geral, funções não-lineares de um parâmetro selecionado.

No linearização estendida procura-se, a partir de uma expressão auxiliar da sistem linearizado, "família paramétrica", uma expressão exata para os ganhos do controlador, o qual é projectado por técnicas lineares. Em relação ao gain scheduling, esta nova técnica leva a uma formulação contínua do controlo do gain-scheduling com n° finito de pontos de operação.

Caso SISO, Rezimentação de Estados

$$\dot{x} = f(x, u) \quad \begin{array}{l} x \in \mathbb{R}^n \\ u \text{ é escalar} \end{array}$$

Sistema linearizado em  $(0,0)$  é controlável, com  $\frac{\partial f}{\partial x}(0,0)$  invertível

Para  $u(t) = \varepsilon$ , a família de pontos de operação é dada por

$$f(x(\varepsilon), \varepsilon) = 0 \quad \xrightarrow{\varepsilon \rightarrow 0}$$

Rezimentação de estados:  $u = v - k(x)$   
 $k(\cdot)$  é o ganho não linear com  $k(0) = 0$

$$\dot{x} = f[x, v - k(x)]$$

#### 4. PSEUDO-LINEARIZAÇÃO

### 3. OBSERVADORES DE ESTADOS

Quando se estuda um sistema físico, o modelo matemático associado deve conter o número de variáveis de estado necessário e suficiente para determinar totalmente o estado num instante qualquer do tempo. Em vista disso, a saída do sistema deve prever medidas de todas essas variáveis de estado, seja para uso no dimensionamento do controle ou para qualquer outra finalidade. No entanto, em vários sistemas de grande importância prática, o vetor de estado não fica integralmente disponível através de medidas. Por exemplo, as medidas podem exigir o uso de sensores muito caros para se obter todas as variáveis de estado. Outra, algumas componentes do vetor de estado correspondem a variáveis internas processíveis, que não podem ser medidas.

Nessas situações, as estratégias de controle devem ser baseadas em valores de apenas um subconjunto das variáveis de estado. Existiriam, em princípio, duas possibilidades de ataque ao problema. A primeira seria procurar novos procedimentos que exigissem um número menor de medidas. A outra, mais razoável, seria descobrir alguma forma de reconstruir aproximadamente o estado baseado nas medidas disponíveis. Essa segunda forma, na hipótese de admitirmos total conhecimento, ou seja, ausência de incertezas, sobre o nosso modelo matemático e sobre o processo de medidas leva ao chamado OBSERVADOR DE ESTADOS.

Um observador é, então, um sistema dinâmico determinístico que pode, convenientemente, computar o estado, ou uma aproximação do estado, num instante qualquer. As entradas de entrada no observador são as medidas (saídas) do sistema original e o vetor de estado do observador gera as informações relevantes sobre o estado do sistema original. Para desenvolver a Teoria de Observadores, admitiremos que o sistema em estudo seja

completamente observável e completamente controlável.

### 3.1. CONTROLABILIDADE E OBSERVABILIDADE

Os conceitos de controlabilidade e observabilidade foram introduzidos por Kalman e geram condições necessárias e, em alguns casos, condições suficientes, para garantir que um problema de controle ótimo, determinístico ou estocástico, tem solução possível.

Para definirmos rigorosamente controlabilidade e observabilidade, juntaremos um sistema dinâmico com  $n$  variáveis de estado, definido num intervalo de tempo  $(T_1, T_2)$  e que obedece às equações:

$$\begin{aligned}\dot{x}(t) &= f[x(t), u(t), t] \\ y(t) &= g[x(t), u(t), t]\end{aligned} \quad (\text{sistema})$$

onde  $u(t)$ , vetor de entradas ou de controle, tem dimensão  $m$  e pertence a um deles conjunto  $\bar{U}$  de entradas admissíveis.

Tomemos agora  $t_0 \in (T_1, T_2)$  e  $x_0 \in \mathbb{R}^n$ . Então, para  $u \in \bar{U}$  existe e é única a solução do sistema dinâmico ( $\dot{x} = f[\cdot]$ ) dada por

$$x(\tau) = \Phi(\tau; u(t_0, \tau); x_0)$$

que satisfez a condição inicial

$$x(t_0) = x_0.$$

$\Phi(\tau; u(t_0, \tau); x_0)$  é dita função de transição do sistema, em analogia à matriz de transição dos sistemas lineares.

A partir daí, é possível, então, definir estados controlável e controlabilidade completa como segue:

Def.1 : Se é possível atingir o estado  $x_1 = 0$ , originariamente de  $x_0$  em  $t_0$ , então, o estado  $x_0$  é controlável em  $t_0$ . Em outras palavras,  $x_0$  é controlável

em  $t_0$  se existe uma função  $u^\circ$  tal que:

$$\varphi(T; u^\circ(t_0, \bar{t}); x_0) = x_1(T) = 0$$

para algum  $T$  finito,  $T \geq t_0$ .

Def. 2: Se cada estado  $x_0$  é controlável em cada instante  $t_0$  do intervalo  $(T_1, T_2)$  cutâo dizemos que o sistema é completamente controlável.

Estados observáveis e observabilidade completa são definidos da seguinte forma:

Def. 3: O estado  $x_0$  é observável em  $t_0$  se, dado um controle qualquer  $u$ , existe um instante  $t_1$ ,  $t_1 \geq t_0$ , tal que o conhecimento de  $u(t_0, t_1]$  e da saída  $y(t_0, t_1] = \hat{y}(x_0; u(t_0, t_1], t_1)$  é suficiente para determinar o estado  $x_0$ .

Def. 4: Se cada estado  $x_0$  é observável em todos os instantes  $t_0$  do intervalo  $(T_1, T_2)$  de interesse, então o sistema é completamente observável.

Em resumo, a controlabilidade de um sistema indica a possibilidade de transferir o sistema de um estado qualquer à origem (estado nulo), num intervalo finito de tempo, pelo aplicação de um controle conveniente. Por outro lado, a observabilidade indica a possibilidade de se determinar as componentes do estado do sistema através de observações da saída do sistema. Desse ponto em diante, admitiremos que os sistemas em estudo são completamente controláveis e completamente observáveis.

Existem regras práticas para determinar controlabilidade ou observabilidade apenas nos casos lineares. Para um sistema não linear é muito difícil a aplicação das definições e o que se faz normalmente é linearizar os sistemas e estudar controlabilidade e observabilidade em intervalos pequenos para os quais valha a linearização.

No caso de sistemas lineares a regra prática é obtida através do seguinte teorema ( prova nos livros de Maybeck ou Luenberger):

Considere um sistema linear de  $n$ -ésima ordem da forma:

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

onde  $A$  é uma matriz  $(n \times n)$ ,  $B$  é uma matriz  $(n \times m)$  e  $C$  é uma matriz  $(r \times n)$ .

Então:

1) O sistema acima é completamente controlável se e somente se a matriz  $(n \times nm)$

$$M = [B; AB; A^2B; \dots; A^{n-1}B]$$

é de ordem  $n$ , i.e., existe determinante não nulo de ordem  $n$ .

2) O sistema acima é completamente observável se e somente se a matriz  $(n \times nr)$

$$N = [C^T; A^T C^T; (A^T)^2 C^T; \dots; (A^T)^{n-1} C^T]$$

é também de ordem  $n$ .

Exemplos:

a) Consideremos o sistema:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -3 & 1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Nesse caso  $n=2$ ,  $m=1$  e  $\mathcal{T}=1$  e

$$\Delta = \begin{bmatrix} -3 & 1 \\ -2 & 0 \end{bmatrix} ; \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad e \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

Montando  $M$ ; ( $2 \times 2$ )

$$M = \left[ B; \Delta B; \Delta^2 B; \dots; \Delta^{n-1} B \right] = \left[ B; \Delta B \right]$$

$$\Delta B = \begin{bmatrix} -3 & 1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 + 1 \\ -2 + 0 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

$$M = \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \quad \text{é de classe (ordem) } 1 < n \leq 2$$

Logo o sistema é não controlável.

Montando  $N$  ( $2 \times 2$ )

$$N = \left[ C^T; A^T C^T; (A^T)^2 C^T; \dots; (A^T)^{n-1} C^T \right] = \left[ C^T; A^T C^T \right]$$

$$A^T C^T = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

$$N = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \quad \text{é de classe (ordem) } 2$$

Então o sistema é observável

Para verificarmos esses conceitos, consideremos o sistema dado:

$$\dot{x}_1 = -3x_1 + x_2 + u \quad (1)$$

$$\dot{x}_2 = -2x_1 + u \quad (2)$$

Multiplicando (2) por  $-1$  e somando com a (1) :

$$\dot{x}_1 - \dot{x}_2 = -x_1 + x_2$$

$$\frac{d}{dt}(x_1 - x_2) = -(x_1 - x_2)$$

O que mostra que a diferença (uma combinação linear) entre as variáveis de estado não é afetada pelo controle  $u$  e, portanto, o sistema é não controlável.

Por outro lado, a saída do sistema é dada por

$$y = x_1 \quad \text{e} \quad \dot{y} = \dot{x}_1 = -3x_1 + x_2 + u$$

que podemos colocar na forma

$$\begin{cases} x_1 = y \\ x_2 = 3y + \dot{y} + u \end{cases}$$

O que mostra que  $x_1$  e  $x_2$  podem ser obtidos de  $y$  e de  $u$ .

b) Consideremos agora o sistema:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

$$y = [1 \quad 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

e, da mesma forma,  $n=2$ ,  $m=1$ ,  $I=1$ , com

$$A = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad C = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

Montando  $M$ :

$$AB = \begin{bmatrix} -3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

$$M = \left[ B : AB \right] = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \quad \text{é de classe (ordem) 2.}$$



Logo, o sistema é controlável.

Montando N:

$$A^T C^T = \begin{bmatrix} -3 & 1 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 + 1 \\ -2 + 0 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

$$N = [C^T : A^T C^T] = \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \text{ é de classe (ordem) } L$$

Portanto, o sistema é não observável, o que pode ser confirmado verificando-se que:

$$y = [1 \ 1] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 + x_2$$

$$\dot{y} = \ddot{x}_1 + \ddot{x}_2 = (-3x_1 - 2x_2 + u) + (x_1)$$

ou

$$\left\{ \begin{array}{l} y = x_1 + x_2 \\ \dot{y} = -2(x_1 + x_2) + u \end{array} \right.$$

O que mostra que apenas a soma  $x_1 + x_2$  pode ser obtida.

### 3.2. O OBSERVADOR TRIVIAL

Uma solução trivial para o problema de estimação determinística do estado de um sistema dinâmico é construir um algoritmo que simule uma cópia exata do sistema original. Note que nesse caso haveria total ausência de qualquer medida ou observação sobre o sistema, isto é, mesmo que alguma medida do sistema real esteja sendo feita, ela não será utilizada na simulação.

Se, por exemplo, o sistema real tem o modelo:

$$\dot{x}(t) = Ax(t) + Bu(t)$$



o observador teria a forma:

$$\dot{\hat{z}}(t) = A\hat{z}(t) + Bu(t)$$

OBS. TRIVIAL

As entradas  $u(t)$  do sistema original são entradas ou controles aplicados externamente e, portanto, podem ser também aplicados à ótica. Agora, desde que o observador é um modelo matemático que está sendo simulado com entradas conhecidas e saída do observador fornece o estado  $\hat{z}(t)$  estimado.

Então, se  $\hat{z}(0) = x(0)$ , o modelo do observador segue o modelo do sistema de forma exata. No entanto, o que define um bom ou ruim observador é a convergência do valor estimado para o valor real, isto é, a velocidade com que o erro:

$$e(t) = \hat{z}(t) - z(t)$$

tende a zero.

No caso do observador trivial, é fácil verificar que:

$$\dot{e}(t) = \dot{\hat{z}}(t) - \dot{z}(t) = [A\hat{z}(t) + Bu(t)] - [Ax(t) + B'u(t)]$$

$$\dot{e}(t) = Ae(t)$$

e, como isso, o erro não estima-se tende a zero apenas no caso do sistema original ser estável, o que é determinado pela matriz  $A$ , e, apenas na velocidade determinada pelos autovalores associados ao modelo real, se  $x(0) \neq \hat{z}(0)$ . Esses fatos representam uma enorme limitação que torna o observador trivial de uso muito restrito.

### 3.3. O OBSERVADOR IDENTIDADE

Consideremos agora o sistema completamente ob-



servável:

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t)$$

onde  $x$  é um vetor de dimensão  $n$  e  $y$ , vetor de saída, tem dimensão  $p$ .

Dois comentários importantes devem ser feitos neste tópico: primeiro, estamos admitindo as matrizes  $A$ ,  $B$  e  $C$  como matrizes de constantes. Isso absolutamente não invalida o nosso desenvolvimento, que deve seguir as mesmas passos para  $A$ ,  $B$  e  $C$  variando no tempo. Segundo, a saída do sistema poderia ter sido escrita na forma:

$$y(t) = Cx(t) + Du(t)$$

sem alterar em modo o desenvolvimento feito.

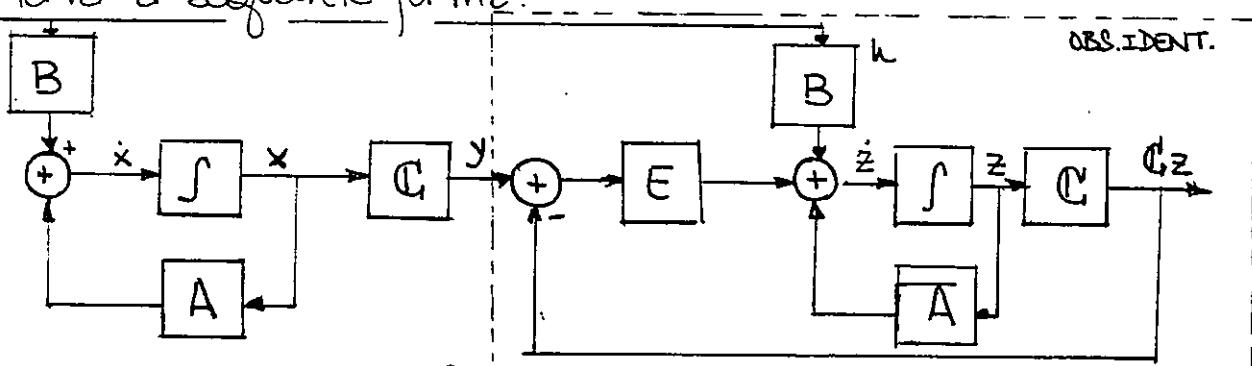
Consideremos agora para o sistema o mesmo observador que leva em conta as medidas:

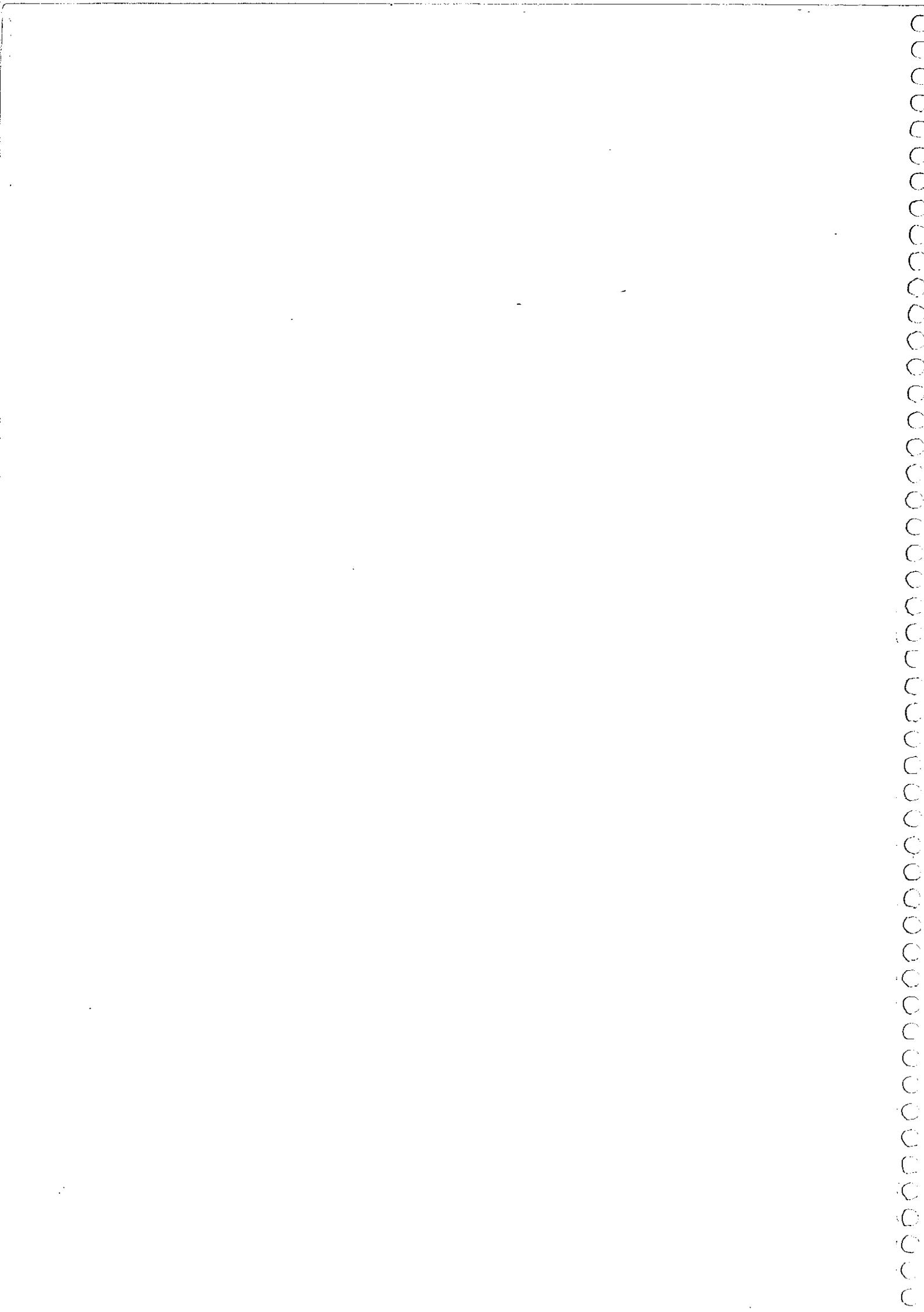
$$\dot{z}(t) = Az(t) + E[y(t) - Cz(t)] + Bu(t)$$

OBSERVADOR IDENTIDADE

onde  $z$  tem dimensão  $n$  e a matriz  $E$  ( $n \times p$ ) deve ainda ser especificada.

As entradas para o observador são de dois tipos: O primeiro é o conjunto de medidas  $y(t)$  disponíveis a partir do sistema real. O segundo é o conjunto de entradas  $u(t)$ , controlados do sistema real e, portanto, à nossa disposição. Descrevemos a matéria a seguinte forma.





Note que o observador identidade é uma generalização do observador trivial já que se  $z(t) = x(t)$  então:

$$\dot{z}(t) = A z(t) + B u(t)$$

que é o observador trivial.

Se usarmos  $y(t) = C x(t)$  na relação acima do observador identidade:

$$\dot{z}(t) = A z(t) + E C [x(t) - z(t)] + B u(t)$$

Subtraindo a equação acima da equação para  $x$ :

$$[\dot{z} - \dot{x}] = [A - EC][z - x]$$

e verificamos que se  $z(0) = x(0)$  então  $z(t) = x(t)$  para todo  $t > 0$ . O estado do observador segue o estado do sistema original e, nesse caso, o observador é dito identidade porque segue o estado inteiro do sistema real, i.e., em todas suas  $n$  componentes.

No caso geral de  $z(0) \neq x(0)$ , o vetor de erro:

$$e(t) = z(t) - x(t)$$

é dado pelo sistema homogêneo

$$\dot{e}(t) = [A - EC] e(t)$$

Se a matriz  $[A - EC]$  for assintoticamente estável, o vetor de erros tende a zero à medida que o tempo cresce. Note que os autovalores da matriz podem ser alterados conforme a escolha da matriz  $E$  e, portanto, podemos escolher como será a resposta do observador em termos de levar o erro a zero. Isso está embutido no Teorema do Observador Identidade, cuja prova se encontra em Luenberger (1970):

Dado um sistema completamente observável

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y = Cx(t)$$

um observador identidade no forma

$$\dot{z} = Az + E[y - Cz] + Bu$$

pode ser construído e os coeficientes do polinomio característico (que fornece os autovalores) do observador podem ser selecionados arbitrariamente.

Exemplo:

Considere o sistema dinâmico

$$\begin{bmatrix} \dot{x} \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

cujo vetor mensurável de saída é dado por

$$y = [1 \ 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

ou seja, só se mede  $x_1$ .

O polinomio característico do sistema é: comprimento

$$|\xi I - A| = \left| \begin{bmatrix} \xi & 0 \\ 0 & \xi \end{bmatrix} - \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} \right| = \left| \begin{bmatrix} \xi + 2 & -1 \\ 0 & \xi + 1 \end{bmatrix} \right|$$

ou seja, os autovalores são  $\xi_1 = -2$  e  $\xi_2 = -1$

Um observador identidade pode ser determinado se especificarmos

$$E = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \quad \text{matriz } (2 \times 1)$$

Montamos, então

$$[A - EC] = \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} [1 \ 0] = \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} e_1 & 0 \\ e_2 & 0 \end{bmatrix}$$

$$[A - EC] = \begin{bmatrix} -2 - e_1 & 1 \\ -e_2 & -1 \end{bmatrix}$$

Montando o polinômio característico do observador:

$$\begin{vmatrix} \lambda I - [A - EC] \end{vmatrix} = \begin{vmatrix} \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} -2-e_1 & 1 \\ -e_2 & -1 \end{bmatrix} \end{vmatrix} = \begin{vmatrix} \lambda+2+e_1 & -1 \\ +e_2 & \lambda+1 \end{vmatrix}$$

$$= (\lambda+2+e_1)(\lambda+1) + e_2 = 0$$

$$\Rightarrow \lambda^2 + (3+e_1)\lambda + 2 + e_1 + e_2 = 0 \quad (1)$$

Admitindo agora que queremos esse observador com dois autovalores iguais a  $\lambda = -3$ , scudo, portanto, mais rápido que o sistema original, teríamos para equações características:

$$(\lambda+3)(\lambda+3) = (\lambda+3)^2 = \lambda^2 + 6\lambda + 9 = 0 \quad (2)$$

Por verificações de (1) e (2):

$$3+e_1 = 6$$

$$2+e_1+e_2 = 9$$

onde:  $e_1 = 3$ ;  $e_2 = 4$

A forma final do observador seria, então:  
 $\dot{\underline{z}} = A\underline{z} + E[\underline{y} - C\underline{z}] + Bu$

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} - \begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} y + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -5 & +1 \\ -4 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} y + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

### 3.4. OBSERVADORES DE LUENBERGER OU OBSERVADORES DE ORDEM REDUZIDA

O observador identificado mostrado no item anterior tem um certo grau de redundância já que reconstrói todos os  $n$  variáveis do sistema real mesmo quando dispõe de  $p$  variáveis, medidas diretas e precisamente. Parece razoável, então, eliminar essa redundância do observador até uma ordem  $n-p$ , ao invés de  $n$ , sendo que o vetor de estados total do sistema real é obtido por  $p$  medições diretas e  $n-p$  estimativas vindas do estimador.

Consideremos, então, o sistema completamente observável

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t)$$

onde  $C$ , matriz  $(p \times n)$  é de classe  $p$ , o que significa que as  $p$  medidas são linearmente independentes (qualquer)

Tomemos uma matriz  $V_{[(n-p) \times n]}$ , tal que a matriz  $P_{(n \times n)}$ , construída através de:

$$P = \begin{bmatrix} V \\ C \end{bmatrix}$$

é não singular. Essa construção é sempre possível desde que  $C$  é de classe  $p$ .

Introduzimos agora a transformação de variáveis:

$$\tilde{x}(t) = P x(t) = \begin{bmatrix} V \\ C \end{bmatrix} x(t)$$

e consideremos o novo vetor de estados particionado da forma:

$$\tilde{x}(t) = \begin{bmatrix} w(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} Vx(t) \\ Cx(t) \end{bmatrix}$$

onde  $w(t)$  é de dimensão  $(n-p)$  e  $y(t)$  é o vetor de medições. Nessa forma, o vetor de saídas é igual às  $p$  últimas variáveis do estado transformado  $\tilde{x}(t)$

Consideremos agora a transformação aplicada ao sistema dinâmico, lembrando que  $\dot{\tilde{x}}(t) = P^{-1} \dot{x}(t)$ :

$$\dot{\tilde{x}}(t) = P \dot{x}(t) = PAx(t) + PBu(t) = PAP^{-1}\dot{\tilde{x}}(t) + PBu(t)$$

e coloquemos na forma particionada:

$$\begin{array}{l} \text{dim}^{(n-p)} \rightarrow \begin{bmatrix} \dot{w}(t) \\ \dot{y}(t) \end{bmatrix} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix} \begin{bmatrix} w(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \end{bmatrix} u(t) \\ \text{dim}(p) \rightarrow \end{array} \quad (*)$$

Podemos extrair desse sistema (\*) um subsistema de ordem  $(n-p)$  que tem como entradas as quantidades conhecidas  $w(t)$  e  $y(t)$ . Para isso, consideremos:

$$\dot{w}(t) = \bar{A}_{11} w(t) + \bar{A}_{12} y(t) + \bar{B}_1 u(t)$$

$$\dot{y}(t) = \bar{A}_{21} w(t) + \bar{A}_{22} y(t) + \bar{B}_2 u(t)$$

Multipiquemos a segunda equação por uma matriz  $E$ ,  $[(n-p) \times p]$ , arbitrária e subtraímos da primeira para obter:

$$\dot{w} - E\dot{y} = (\bar{A}_{11} - E\bar{A}_{21})w + (\bar{A}_{12} - E\bar{A}_{22})y + (\bar{B}_1 - E\bar{B}_2)u$$

A relação acima pode ser reescrita como:

$$\dot{w} - E\dot{y} = (\bar{A}_{11} - E\bar{A}_{21})[w - Ey] + (\bar{A}_{12} - E\bar{A}_{22} + \bar{A}_{11}E - E\bar{A}_{21}E)y + (\bar{B}_1 - E\bar{B}_2)u$$

Fazendo:  $v = w - Ey$ , chegamos a:

$$\dot{v}(t) = (\bar{A}_{11} - E\bar{A}_{21})v + (\bar{A}_{12} - E\bar{A}_{22} + \bar{A}_{11}E - E\bar{A}_{21}E)y + (\bar{B}_1 - E\bar{B}_2)u$$

Nesse equação  $\dot{v}(t)$ , uma transformação das variáveis de estado, é a incógnita, enquanto  $y(t)$  e  $u(t)$  servem como entradas conhecidas. Não temos nenhuma observação do vetor  $v(t)$  de dimensão  $(n-p)$ . Então, o novo modelo de observador é fornecido por meio cópia da relação acima. Portanto:

$$\hat{z}(t) = (\bar{A}_{11} - E\bar{A}_{21})z(t) + (\bar{A}_{11}E - E\bar{A}_{21}E + \bar{A}_{12} - E\bar{A}_{22})y(t) + (\bar{B}_1 - E\bar{B}_2)u(t)$$

O modelo acima é um observador para o sistema (\*). Sendo agora a dimensão de  $z(t)$  igual a  $(n-p)$ , número mínimo para completar a estimativa do estado do sistema original, a equação acima expressa um observador de ordem reduzida ou observador de Luenberger.

Agora, subtraindo a equação em  $\hat{z}$  da de  $y$ :

$$\hat{z}(t) - y(t) = (\bar{A}_{11} - E\bar{A}_{21})[z(t) - v(t)]$$

e o erro evolui na forma:

$$e(t) = (\bar{A}_{11} - E\bar{A}_{21})e(t)$$

Então, o estado  $z(t)$  do observador tende a  $v(t)$  numa velocidade determinada pelos autovalores da matriz  $\bar{A}_{11} - E\bar{A}_{21}$ .

Como nosso observador segue  $v(t)$ ,  $v(t) = w(t) - E y(t)$ , o estado do sistema real é estimado através de  $\hat{w}$  e  $\hat{y}$  dadas por:

$$\hat{w}(t) = z(t) + E y(t)$$

$$\hat{y}(t) = y(t) \quad (\text{medido})$$

Com  $w$  estimado através de  $\hat{w}$ , podemos voltar à estimativa de  $x$ . Dado que:

$$\hat{x} = \begin{bmatrix} \hat{w} \\ y \end{bmatrix} \quad \text{e} \quad \tilde{x} = P x$$

basta fazer  $\hat{x} = P^{-1} \tilde{x}$  o que pode ser feito desde que  $P$  foi montada não singular.

A eficiência do observador depende fortemente dos autovalores da matriz do observador  $[\bar{A}_{11} - E\bar{A}_{21}]$ . No entanto, pode ser mostrado que se o sistema original:

$$\dot{\tilde{x}}(t) = A\tilde{x}(t) + Bu(t)$$

$$y(t) = Cx(t)$$

for completamente observável, então o par  $\bar{A}_{21}, \bar{A}_{11}$  é também completamente observável.

Com isso, os autovalores de  $\bar{A}_{11} - E\bar{A}_{21}$  podem ser selecionados arbitrariamente por uma escolha conveniente de  $E$ , o que permite para o observador de Luenberger a mesma flexibilidade do observador identificável. Todos esses resultados podem ser resumidos no Teorema do observador de ordem reduzida enunciado a seguir e cuja prova se encontra em Luenberger (1979).

Dado um sistema completamente observável de ordem  $n$ :

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y = Cx(t)$$

com  $p$  saídas linearmente independentes, um observador de ordem  $n-p$  na forma proposta pode ser projetado e os autovalores do observador podem ser selecionados arbitrariamente.

Um resultado importante no que se refere à utilização dos observadores é o de que a introdução de um observador num sistema com regulamentações para suprir medidas não disponíveis não altera o comportamento do sistema de regulamentação. O observador simplesmente joga seus próprios autovalores aos autovalores da malha de controle. Isso significa que se temos um sistema de controle com regulamentações em que existem estados inacessíveis, podemos projetar o sistema de controle como se tivessemos todos os variáveis disponíveis e usar um observador para estimar o estado sem alterar o comportamento global do sistema de controle. Esse resultado pode ser adocido em termos do um Teorema de Separação dos Autovalores, aqui enunciado por um observador identificado que pode ser facilmente estendido para o observador

de Luenberger:

Considere o sistema completamente controlável e observável:

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t)$$

e o observador identificado

$$\dot{\hat{z}}(t) = (A - EC)\hat{z}(t) + Ey(t) + Bu(t)$$

e, ainda, a lei de controle

$$u(t) = K\hat{z}(t).$$

O polinômio característico do sistema composto (controlador e observador) é o produto dos polinômios característicos de  $(A + BK)$  e  $(A - EC)$

Exemplo:

Consideremos um sistema de segundo ordenamento. Esse modelo representa um sistema unidimensional de rastreamento, como uma antena que dore estor sempre apontando para um objeto que se move num plano. No caso da antena, as variáveis  $x_1$  e  $x_2$  representam posição e velocidade angulares do objeto rastreado, respectivamente.

$$\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & -\beta \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ l \end{bmatrix}}_{B} u$$

As medidas são feitas apenas para as posições, i.e.:

$$y = \underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_C \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Comecemos a montar o observador de Luenberger

pela correta especificação da matriz  $\underline{P}$

$$\underline{P} = \begin{bmatrix} \sqrt{-} \\ -\frac{1}{c} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ 1 & 0 \end{bmatrix}$$

Escolhemos  $\underline{P}$  definida positiva:

$$\underline{P} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix} \quad \begin{array}{l} \text{— escolhe} \\ \text{arbitraria} \end{array}$$

cujos menores principais  $\begin{bmatrix} 1 \end{bmatrix}$ ,  $\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$  tem determinantes iguais a 1. A inversa de  $\underline{P}$  será:

$$\underline{P}^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\text{Com isso: } \tilde{x} = \underline{P}x = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ x_1 \end{bmatrix}$$

$$\dot{\tilde{x}} = \underline{P}A\underline{P}^{-1}\tilde{x} + \underline{P}B u$$

$$\dot{\tilde{x}} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & -\beta \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \varrho \end{bmatrix} u$$

$$\dot{\tilde{x}} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix} \tilde{x} + \begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \end{bmatrix} u$$

A variável a ser observada será então:

$$\dot{v}(t) = (\bar{A}_{11} - E\bar{A}_{21})v(t) + (\bar{A}_{11}E - E\bar{A}_{21}E + \bar{A}_{12} - E\bar{A}_{22})y(t) + (\bar{B}_1 - E\bar{B}_2)u(t)$$

e como  $E = [e]$ , obter:

$$\dot{v} = [- (1+\beta) - e(-1)]v + [-(1+\beta)e - e(-1)e + (1+\beta) - e(-1)]y + (-\varrho + e \cdot 0)u$$

$$\dot{v} = [e - (1+\beta)] v + [e^2 - \beta e + (1+\beta)] y + (-l) u$$

O observador de Luenberger para esse caso será, então:

$$\dot{z} = [e - (1+\beta)] z + [e^2 + (1+\beta) - \beta e] y + (-l) u$$

Veja que o autoválor envolvido na equação para a variável transformada  $w(t)$  que deve ser o saída do observador é dado por  $\bar{\lambda}_{11}$ . Nesse caso

$$\bar{\lambda} = - (1+\beta)$$

Se escolhermos, por exemplo,  $e = -4(1+\beta)$ , então o autoválor associado ao observador será:

$$\lambda = e - (1+\beta)$$

$$\lambda = -5(1+\beta)$$

o que torna o observador muito mais rápido que o sistema original.

A estimativa do estado é feita através de:

$$\hat{w}(t) = z(t) + e y(t) = z(t) - 4(\beta+1)y(t)$$

$$\hat{y}(t) = y(t)$$

e, como

$$\hat{x} = P^{-1} \hat{z} = P^{-1} \begin{bmatrix} \hat{w} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} z(t) - 4(\beta+1)y(t) \\ y(t) \end{bmatrix}$$

$$\hat{x}(t) = \begin{bmatrix} y(t) \\ -z(t) + 4\beta y(t) + 5y(t) \end{bmatrix}$$

rel. entre  
 $x_2 \in \underline{z}$

E simulemos o sistema com

$$\beta = 1.0 \text{ seg}^{-1}; l = 1.0; u(t) = -1.0 \text{ ft/s}^2 (\text{const.})$$

e com condições iniciais  $x_{1(0)} = 1.0 \text{ ft}$  e  $x_{2(0)} = 1.0 \text{ ft/s}$ , po-

deveremos iniciar o observador escolhendo

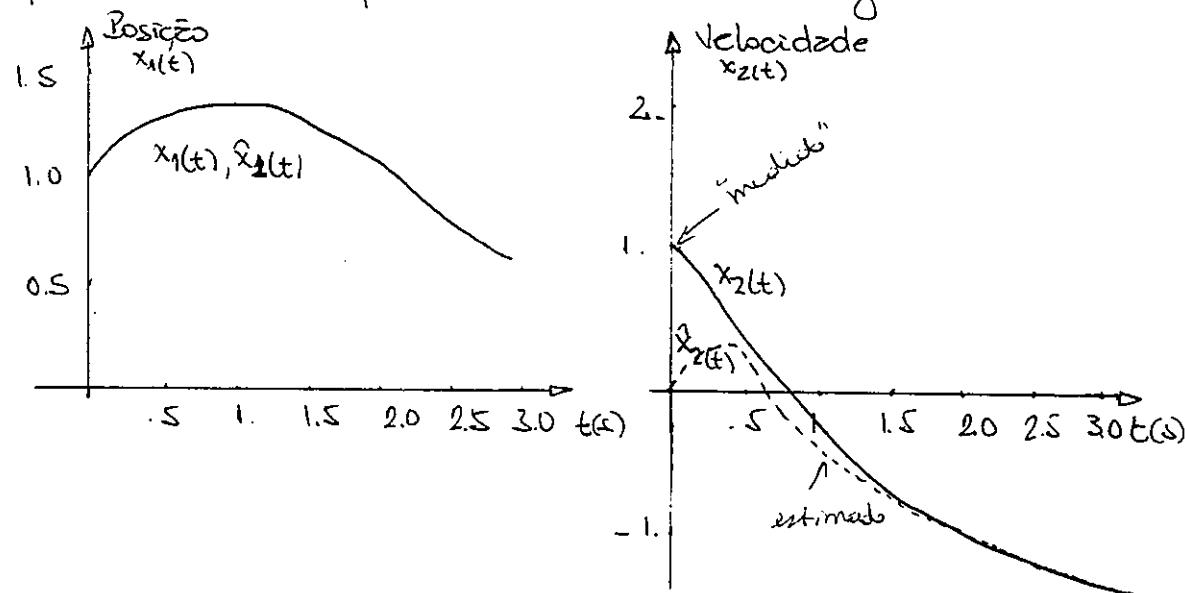
$$x_1(0) = 1.0 \text{ ft} = x_{1,0} \quad \text{e} \quad \hat{x}_2(0) = 0$$

e, como  $\dot{z}(t) = -\dot{x}_1(t) + 4\beta y(t) + 5y(t)$

$$z(0) = -1 + 4 \cdot 1 \cdot 1 + 5 \cdot 1 = 8 \text{ ft/s}$$

anterior

Resultados simulados por Gelb (1974) e Kwakernaak (1972) para esse exemplo são mostrados a seguir



Note que como  $x_1(t)$  é medido diretamente não há erro no estimativa, enquanto que o erro no estimativa de  $x_2(t)$  decréce exponentialmente a zero.

Se não tiver Matlab, pode usar:

SILAB

OCTAVE

### 3.4. Um Criterio de Otimização para Escolher Polos de Observadores

Observadores assintóticos, como o identificado e o de Luenberger, são duals dos controladores ruídos. Desde que se pode otimizar o controlador ruído, gerando o regulador ótimo quadrático, é natural que se possa pensar num dual, observador assintótico ótimo.

É claro que um critério poderia ser otimizar colocando os polos do observador o mais possível à esquerda no plano complexo. Dessa maneira, os erros do observador seriam reduzidos a zero tão rápido quanto quisermos.

No entanto, o estimador resultante teria, além de uma banda passante muito larga, uma sensibilidade muito alta aos "quase-instantâneos" ruidos de medida. Assim, não se deve tentar uma velocidade de muito alta de reconstruções do estado.

Uma solução de compromisso pode ser encontrada resolvendo um problema LQG para o sistema:

$$*\dot{\tilde{x}} = Ax + Bu + G_1 v_1(t)$$

$$\tilde{y} = Cx + G_2 v_2(t)$$

onde  $v_1$  e  $v_2$  são ruidos brancos Gaussianos, não correlacionados e de média nula

Para esse sistema define-se um observador:

$$\hat{x}(t) = Ax(t) + Bu(t) + L[y(t) - C\hat{x}(t)] \quad , \quad \hat{x}(0) = 0$$

e usz-se como critério escolher  $T_i$  (os elementos da sua diagonal, por exemplo), de modo a minimizar

$$J = E[(x(t) - \hat{x}(t))^T(x(t) - \hat{x}(t))]$$

Com isso, é possível escolher (ou obter) os polos já levando em conta o efeito dos ruídos no sistema.

## ASYMPTOTIC OBSERVERS AND COMPENSATOR DESIGN

43. R. E. KALMAN and J. BERTRAM, "General Synthesis Procedure for Computer Control of Single and Multi-Loop Linear Systems," *Trans. IEEE*, 77, pp. 602-609, 1958.
  44. R. BELLMAN, *Dynamic Programming*, Princeton University Press, Princeton, N.J., 1957.
  45. J. Bernoulli, "Solution du Problème Proposé par M. Jacques Bernoulli, etc. (1706)," in *Collected Works*, G. Cramer, ed., Vol. I, pp. 424-435, Lausanne, 1742.
  46. L. EULER, "Methodus Inveniendi Lineas Maximi Minimi Proprietate Gaudentes Sive Solutio Problemati Isoperimetri Latissimo Sensu Accepti, Lausanne, 1744," in *Collected Works*, Ser. 1, Vol. 24, Teubner, Leipzig, 1938.
  47. C. CARATHEODORY, *Calculus of Variations and Partial Differential Equations of the First Order*, Teubner, Leipzig, 1935 (English trans., Holden-Day, San Francisco, 1967).
  48. D. M. G. WISHART, "The Continuity of Ideas in the Calculus of Variations," *Control Theory Cent. Rep. No. 21*, University of Warwick, Warwick, England, 1972.
  49. R. WOODHOUSE, *A History of the Calculus of Variations in the Eighteenth Century*, Cambridge University Press, London, 1810 (reprint, Chelsea Publishing Co., New York, 1966).
  - 49a. D. J. STRUIK, ed., *A Source Book in Mathematics, 1200-1800*, Harvard University Press, Cambridge, Mass., 1969.
  50. R. E. KALMAN and R. W. KOEPKE, "Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indexes," *Trans. ASME*, 80, pp. 1800-1826, 1958.
  51. M. MORF and T. KAILATH, "Square-Root Algorithms for Least-Squares Estimation," *IEEE Trans. Autom. Control*, AC-20, pp. 487-497, Aug. 1975.
  52. L. SILVERMAN, "Discrete Riccati Equations: Alternative Algorithms, Asymptotic Properties, and System Theory Interpretations," in *Control and Dynamic Systems*, Vol. 12 (C. T. Leondes, ed.), 1976, pp. 313-386, Academic Press, New York.
  53. G. W. STEWART, *Introduction to Matrix Computations*, Academic Press, New York, 1973.
  54. G. BIERNAN, *Factorization Methods for Discrete Estimation*, Academic Press, New York, 1977.
  55. G. SEBER, *Linear Regression Analysis*, Wiley, New York, 1977.
  56. M. MORF, J. DOMBINS, B. FRIEDLANDER, and T. KAILATH, "Square-Root Algorithms for Parallel Processing in Optimal Estimation," *Automatica*, 15, pp. 299-306, May 1979.
  57. M. MORF, G. S. SIDHU, and T. KAILATH, "Some New Algorithms for Recursive Estimation in Constant, Linear, Discrete-Time Systems," *IEEE Trans. Autom. Control*, AC-19, pp. 315-323, Aug. 1974.
- In Chapter 3, we showed that if a realization  $\{\mathcal{A}, b, c\}$  is controllable, then state-variable feedback can modify the eigenvalues of  $\mathcal{A}$  at will. We shall now discuss the problem of actually obtaining the states of the realization from knowledge only of the system input  $u(\cdot)$  and system output  $y(\cdot)$ . We have already seen in Sec. 2.3 that if the realization  $\{\mathcal{A}, b, c\}$  is also observable, then a differentiation technique can be used to calculate the states. However, this technique is clearly impractical, and therefore in Sec. 4.1 we shall develop a more realistic state estimator, which is usually known as an asymptotic observer. The name arises from the fact that the states can only be obtained with an error, but one that can be made to go to zero at any specified exponential rate.
- In Sec. 4.2 we shall discuss the use of such state estimates in place of the possibly unavailable true states. We shall find that the design equations for the controller are not affected by the fact that approximate states are being used instead of the true states. More crucially, however, we shall find the overall observer-controller important and a priori nonobvious fact that the overall observer-controller configuration is *internally stable*, which was an issue not completely faced by classical design methods. However, use of the estimated instead of the true states, for feedback, may lead in general to a deterioration of the transient response (see Example 4.2-1).
- The observer studied in Sec. 4.1 was apparently first introduced in un-

published work by Bertram in 1961 (see the comments in Astrom [1, p. 158]) and by Bass in 1963 (see the comment of Kalman et al. [2, p. 55]). An independent and rather different approach was published by Luenberger in 1964 [3] (see also his Ph.D. thesis, Stanford University, Stanford, Calif., 1963). Luenberger treats a somewhat more general estimation problem which, when specialized to our situation, actually yields an observer with one less state variable than the observer of Sec. 4.1. This so-called reduced-order or Luenberger observer is discussed in Sec. 4.3 by a method due to Gopinath [4] and Cumming [5] (see also the survey papers [6] and [7]).

In Sec. 4.4, we shall briefly discuss the question of optimum observer pole locations, which will lead us to some results *that* to those developed in Sec. 3.4.1 for the optimal quadratic regulator. Analogs of the results in the other sections of Secs. 3.4 and 3.5 can be developed, though we shall not do so here. Instead, in our concluding Sec. 4.5, we shall explore at some length the fact that it is possible to obtain the combined observer-controller configuration directly by transfer function analysis, thus obviating the need for the notions of state, controllability, and observability. This was shown by Chen ([8]) although, as we shall explain in Sec. 4.5.1, without the insights provided by the state-variable design, certain judicious choices and assumptions in the arguments of Reference [8] might not have been apparent. However, once the transfer function method has been introduced, we shall find that this suggests certain extended results that are not immediately obvious in the state-space approach; moreover, this method will extend nicely to the multivariable case (Sec. 7.5).

## 4.1 ASYMPTOTIC OBSERVERS FOR STATE MEASUREMENT

We shall now begin to explore the question of methods of actually determining the states of a realization

$$\begin{aligned}\dot{x}(t) &= Ax(t) + bu(t), & x(0-) &= x_0 \\ y(t) &= cx(t) & t > 0-\end{aligned}\quad (1)$$

given knowledge only of  $y(\cdot)$  and  $u(\cdot)$ .

In Sec. 2.3, we have already described one method of determining the states at any time  $t$ . But this method involves differentiation and is therefore impractical.

When we reflect on the fact that we know  $A$ ,  $b$ ,  $c$ ,  $y(\cdot)$ , and  $u(\cdot)$ , which is really quite a lot, we wonder why  $x(\cdot)$  cannot be reconstructed by forming a dummy system  $\{A, b\}$  and driving it with  $u(\cdot)$ . The problem, of course, is that it is important to note that the dummy system need not use the same physical components as the original system—it can be a miniaturized electronic analog computer simulation or a special-purpose digital computation package. For the present purpose, quite large amounts of memory may be quite cheaply simulated in this way.

that we do not know the initial condition  $x(0-) = x_0$ . This again has to be found by the differentiation technique, but we have a slight simplification since it is reasonable to assume that  $u(t) = 0$ ,  $t \leq 0-$ . Then some calculation shows that  $x(0-)$  can be obtained from the equation

$$\dot{\theta}x(0-) = [y(0-) - \dots - y^{(n-1)}(0-)]' \quad (2)$$

and in many problems the values  $\{y(0-), \dots, y^{(n-1)}(0-)\}$  are given to us as part of the problem statement.

**An Open-Loop Observer.** Thus, by setting up a dummy system  $\{A, b, c\}$  with the proper initial conditions and driving it with the known input  $\{u(t), t > 0-\}$ , we can obtain  $\{x(t), t > 0-\}$ . Note that the values of the output  $\{y(t), t > 0-\}$  are no longer required. Unfortunately, however, this last fact is really more of a disadvantage than a help. The point is that strategies that are the same no matter what the output is (called *open-loop* strategies as opposed to *feedback* or *closed-loop* methods, which adjust parameters according to the current value of the system output or system state) are obviously susceptible to disturbances that may arise during system operation, and, furthermore, they allow no means of compensating for any errors, possibly small but still almost inevitable, in the predetermined strategy.

Thus, suppose that the initial condition that we use for our dummy

system is slightly in error—say we have not  $x_0$  but

$$\hat{x}_0 \triangleq x_0 - \epsilon, \quad \|\epsilon\| \ll \|x_0\| \quad (3)$$

where we shall use, both here and elsewhere, the caret ( $\hat{\cdot}$ ) to denote *estimate* of. What is the effect of this small error on the states calculated from our dummy system? The states will no longer be  $x(t)$  but will be a different function, say  $\hat{x}(t)$ , which satisfies the equation

$$\dot{\hat{x}}(t) = A\hat{x}(t) + bu(t), \quad \hat{x}(0-) = \hat{x}_0 - \epsilon \quad (4)$$

Clearly, there will be an error

$$\begin{aligned}\hat{x}(t) &\triangleq x(t) - bu(t), & \hat{x}(0-) &= \hat{x}_0 - \epsilon \\ \hat{x}(t) &= A\hat{x}(t) - bu(t), & \hat{x}(0-) &= \hat{x}_0 - \epsilon\end{aligned} \quad (5)$$

which will satisfy the differential equation

$$\dot{\hat{x}}(t) = A\hat{x}(t) - \epsilon \quad (6)$$

That is, the error  $\epsilon$  in the initial condition will produce an error,  $\hat{x}(\cdot)$ , at all later times. Now, as can be seen from a partial (fraction expansion of

$$\hat{x}(s) = \hat{s}sI - A)^{-1}\epsilon$$

$\tilde{x}(t)$  will be a sum of terms of the form  $\{e^{i\lambda t}, t^l e^{i\lambda t}, \dots\}$ , where the  $\{\lambda\}$  are the eigenvalues of  $A$ . Clearly, if the system is unstable (recall that we are interested in determining states to be fed back to achieve stabilization), then the error  $\tilde{x}(t)$  will become *arbitrarily large* as  $t \rightarrow \infty$ , no matter how small the initial error is. Less dramatically, even if the system is stable but some eigenvalues have real parts that are very small, the effects of errors in the initial estimates will take a long time to die out.

**A Closed-Loop Observer.** There may be situations in which open-loop estimators are fairly satisfactory, especially if periodic *resetting* is feasible so as to eliminate, or at least suitably control, the effects of initial errors and later disturbances.

However, the classical way of overcoming the potential difficulties of open-loop systems is to use *feedback*, viz., to try to *zero the error* by driving the system with a term proportional to the error in the estimate. But how can we determine this "error"? In classical feedback problems, this is provided by the *reference signal*.<sup>†</sup> In our problem, the error is  $x(\cdot) - \hat{x}(\cdot) = \tilde{x}(\cdot)$ , but of course  $x(\cdot)$  is not available! Therefore, we have to obtain a reference signal in some other way. The system output  $y(\cdot)$ , which was not used in the open-loop solution, comes in now because it is related to the quantity  $x(\cdot)$  we are interested in:  $y(\cdot) = cx(\cdot)$ , and  $y(\cdot)$  is clearly available. Therefore, an *error signal* can be generated as

$$y(t) - \hat{y}(t) = y(t) - c\hat{x}(t) = c[x(t) - \hat{x}(t)] = c\tilde{x}(t)$$

and it can be used to drive the estimator equation. These considerations lead us to consider an estimator for  $x(\cdot)$  of the form (see Fig. 4.1-1)

$$\dot{\hat{x}}(t) = Ax(t) + bu(t) + ly(t) - c\hat{x}(t), \quad \hat{x}(0) = \hat{x}_0 \quad (6)$$

where

$\hat{x}_0$  = an estimated initial state vector

and

$l$  = a *feedback gain vector*, to be suitably chosen

The system yielding  $\hat{x}(\cdot)$  is called an observer, or, actually, for reasons that will become clearer presently, an *asymptotic observer*.

We should, of course, choose  $l$  so as to properly control the error  $\tilde{x}(\cdot)$ . Now the error  $\tilde{x}(\cdot)$  obeys the differential equation

$$\begin{aligned} \dot{\tilde{x}}(t) &= \tilde{x}(t) - \hat{x}(t), \\ &= (A - lc)\tilde{x}(t), \quad \tilde{x}(0) = x_0 - \hat{x}_0 \end{aligned} \quad (7)$$

<sup>†</sup>For example, in a thermostat, the reference signal is the desired ambient temperature.

<sup>‡</sup>See also Exercise 4.1-2.

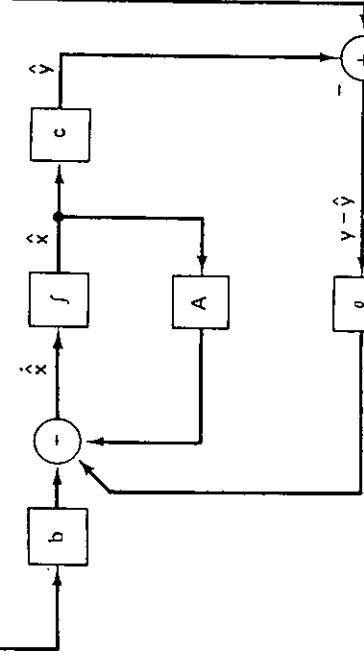
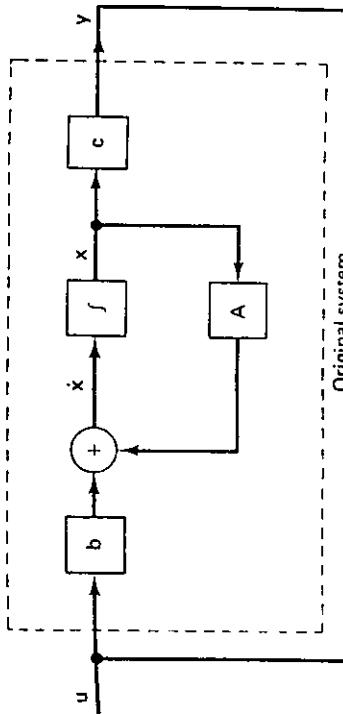


Figure 4.1-1. Block diagram of asymptotic observer. Access to the original system is assumed to be possible at the input and output terminals, while this restriction is not necessary for the observer.

Notice that when  $l = 0$  this equation reduces to the open-loop error equation (5). The effect of the *error feedback*  $l[y(\cdot) - \hat{y}(\cdot)] = l[c(x(\cdot) - \hat{x}(\cdot))]$  is to give us some control over the behavior of the error  $\tilde{x}(\cdot)$ . In fact, the natural frequencies will be the eigenvalues of  $A - lc$ , and what we can try to do is to choose  $l$  so that the error will die out as rapidly as we deem suitable. Notice that the actual value of the initial estimate  $\hat{x}_0$  is unimportant—if we have no special information, we often take  $\hat{x}_0 = 0$ . The reason for the name *asymptotic observer* should now be clear.

But can we always find a suitable  $l$ ? We shall prove that if  $\{c, A\}$  is observable, then we can choose  $l$  so that  $A - lc$  has arbitrary eigenvalues, or equivalently

$$\det(sI - A + lc) = \alpha(s) := s^n + \alpha_1 s^{n-1} + \dots + \alpha_n \quad (8)$$

where the  $\{\alpha_i\}$  are completely arbitrary. The reader will recognize that this problem is essentially the same as the one we discussed in Sec. 3.2 in determining the feedback vector  $k$  required to give a controllable realization of arbitrary dynamics. There is in fact a very close connection, and we shall actually solve our problem by *dualizing* the results of Sec. 3.2.

**Formulas for the Observer Gain.** In Sec. 3.2 we showed that given any realization  $\{A, b, c\}$  with  $\{A, b\}$  controllable we could in several ways find  $k$  so that

$$\det(sI - A + bk) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n \quad (9)$$

for any  $\{\alpha_i\}$ . For example, one formula is [cf. (3.2-13)]

$$k = (\alpha - a)\mathfrak{Q}_-^{-1}\mathfrak{C}^{-1}(A, b) \quad (10)$$

where  $a$  is the vector of coefficients of  $a(s) = \det(sI - A)$  and  $\mathfrak{Q}_-^I$  is an upper triangular Toeplitz matrix with  $[1 \ a_1 \ \cdots \ a_{n-1}]$  as the first row.

To recast our observer problem into this form we merely have to note that

$$\det(sI - A + lc) = \det(sI - A' + c'l')$$

so that if we let

$$A \longrightarrow A', \quad b \longrightarrow c', \quad k \longrightarrow l'$$

in the solution of the controller problem, we deduce that we shall be able to find  $l$  if and only if  $\mathfrak{C}(A', c')$  is nonsingular. Then

$$l' = (\alpha - a)\mathfrak{Q}_-^{-1}\mathfrak{C}^{-1}(A', c')$$

But

$$\begin{aligned} \mathfrak{C}(A', c') &= [c' \quad A'c' \quad \cdots \quad A'^{(n-1)}c'] \\ &= \mathfrak{C}'(c, A), \quad \text{the transpose of the observability matrix of } \{A, b, c\} \end{aligned} \quad (11)$$

Hence the observer gain vector  $l$  can be calculated as

$$l = \mathfrak{C}^{-1}(c, A)\mathfrak{Q}_-^{-1}(\alpha - a) \quad (12)$$

It is perhaps somewhat unexpected that asymptotic observability requires the same condition—the nonsingularity of  $\mathfrak{C}(A, c)$ —as exact observability does. But some reflection on what *unobservable states* are (cf. Sec. 2.4.2) will provide an explanation.

The duality between the asymptotic observer problem and the modal controller problem is quite striking and useful and can be used as above to

translate almost all the results of Chapter 3 to the present context. We shall leave these extensions to the reader (and partly to the exercises).

#### Example 4.1-1. The Pointer-Balancing Problem

For the pointer-balancing problem of Example 3.3-1, design an observer with modes at  $-10 \pm j10$ .

For  $\phi(0) = 0.1$ ,  $\dot{\phi}(0) = 0$ , calculate and plot the observer errors  $\tilde{\phi}(\cdot)$  and  $\hat{\phi}(\cdot)$ .

**Solution.** Recall that with  $g/L = 9$  and  $z = [\phi \ \dot{\phi}]'$  we have

$$\begin{aligned} \dot{z} &= \begin{bmatrix} 0 & 1 \\ 9 & 0 \end{bmatrix} z + \begin{bmatrix} 0 \\ -1 \end{bmatrix} u \\ y &= [1 \ 0] z \end{aligned}$$

We want  $\det(sI - A + lc) = (s + 10)^2 + 10^2 = s^2 + 20s + 200 = a_0(s)$ , and a simple calculation yields  $l' = [20 \ 209]$ . Let  $\tilde{z}(0) = 0$ ; then  $\tilde{z}(0) = z(0) = [0.1 \ 0]'$  and

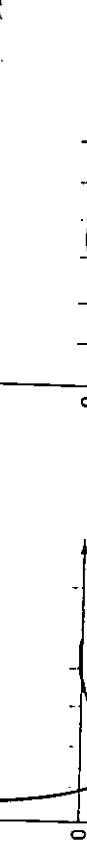
$$\tilde{z}(s) = (sI - A + lc)^{-1}\tilde{z}(0)$$

$$\begin{aligned} &= \frac{1}{a_0(s)} \begin{bmatrix} s & 1 \\ -200 & s + 20 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0 \end{bmatrix} \\ &= \frac{1}{10s} \begin{bmatrix} s & 1 \\ -200 & s + 20 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0 \end{bmatrix} \end{aligned}$$

Therefore

$$\begin{aligned} \tilde{z}_1(t) &= \tilde{\phi}(t) = \mathfrak{L}^{-1} \left[ \frac{0.1s}{a_0(s)} \right] \\ &= 0.1 \mathfrak{L}^{-1} \left[ \frac{(s+10)}{(s+10)^2 + 10^2} \right] = \frac{10}{(s+10)^2 + 10^2} \\ &= 0.1e^{-10t}(\cos 10t - \sin 10t) \end{aligned}$$

$$\tilde{\phi} = \begin{bmatrix} 0.1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -10 & 0 \end{bmatrix} \begin{bmatrix} 0.1e^{-10t}(\cos 10t - \sin 10t) \\ 0.1e^{-10t}(-\sin 10t + \cos 10t) \end{bmatrix}$$



-1

Similarly,  $\dot{\phi}(t) = -2e^{-10t} \sin 10t$ . These are sketched in the figure. Note that in about three time constants ( $3 \times \frac{1}{10}$ ) the observer error is down to a small fraction of its maximum value.

#### Example 4.1-2. The Need for Proper Modeling

Discuss the design, as  $\epsilon \rightarrow 0$ , of observers for a realization

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = x_1(t) + \epsilon x_2(t), \quad \epsilon \ll 1$$

Note that  $x_1(\cdot)$  does not depend on  $x_2(\cdot)$ , and therefore as  $\epsilon \rightarrow 0$ ,  $x_2(\cdot)$  tends to disappear completely from the observation  $y(\cdot)$ . Therefore  $x_2(\cdot)$  becomes unobservable as  $\epsilon \rightarrow 0$ , and we should expect some difficulties in trying to estimate it.

**Solution.** Since this is an observer problem, we can, without loss of generality, assume that  $u(\cdot) \equiv 0$ . We can now readily calculate the observer gain required to obtain any set of observer poles, but we shall see that the parameter  $\epsilon$  will give us indications of potential difficulties. For example,

$$\Theta = \begin{bmatrix} 1 & \epsilon \\ -1 + 2\epsilon & -\epsilon \end{bmatrix}, \quad \det \Theta = -2\epsilon^2$$

so that  $\Theta$  is almost singular for small  $\epsilon$ . Therefore we would expect trouble even if perfect differentiation were feasible. Thus note that

$$\Theta^{-1} = -\frac{1}{2\epsilon^2} \begin{bmatrix} -\epsilon & -\epsilon \\ 1 - 2\epsilon & 1 \end{bmatrix}$$

Therefore, if, for example,  $\epsilon = 0.01$ , we see that the ideal observer will give

$$\begin{aligned} \hat{x}_1(t) &= 50[y(t) + \dot{y}(t)] \\ \hat{x}_2(t) &= -4900y(t) - 5000\dot{y}(t) \end{aligned}$$

while if  $\epsilon = 0.02$ , the estimates change quite a bit to

$$\begin{aligned} \hat{x}_1(t) &= 25[y(t) + \dot{y}(t)] \\ \hat{x}_2(t) &= -1200y(t) - 1250\dot{y}(t) \end{aligned}$$

Such difficulties will persist no matter what method of state estimation is used. In this problem, when  $\epsilon$  is small we see that  $y(\cdot)$  depends almost completely on  $x_1(\cdot)$ , and it is better to represent the model by the equations

$$\begin{aligned} \dot{x}_1(t) &= -x_1(t) + u(t) \\ y(t) &= x_1(t) \end{aligned}$$

for which no observer is needed. Of course this is because the  $x_2(\cdot)$  equation is

stable so that  $x_2(\cdot)$  will tend to die out with time. If  $x_2(\cdot)$  is unstable, then we shall have trouble, and we should reexamine the original model to see if it can be modified.

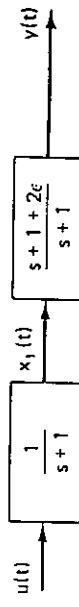
The point is that one should not just blindly plug into a mathematical formula but should also examine the solution and the model for sensitivity to parameter changes, *singular* phenomena, etc.

In this problem, some other clues to the potential difficulties might have been recognized from the transfer function,

$$H(s) = \frac{s + 1 + 2\epsilon}{(s + 1)^2}$$

which will show that, for small  $\epsilon$ , we really have a first-order system. The pole-zero cancellation for  $\epsilon = 0$  shows that we shall lose observability or controllability of the realization. A block diagram of the realization (see the figure) shows that as  $\epsilon \rightarrow 0$ , a pole of the first subsystem cancels a zero of the second subsystem, showing (according to Example 2.4-4) that we have unobservability of the overall realization.

The active reader should try to develop the *details* for controllability of the above problem and remarks.



#### Exercises

##### 4.1-1. Ackermann's Formula

Obtain a formula for  $I$  in terms of  $\{A, c\}$  and the coefficients  $\{\alpha_i\}$  of the desired characteristic polynomial,  $\alpha(s)$ .

##### 4.1-2. Why Not Feedback to the Input?

In Chapter 3, we used feedback to the input according to  $u(t) = r(t) - ky(t)$ . We could do this for the observer  $u(t) = r(t) + I[y(t) - \hat{y}(t)]$ , where, of course,  $I$  is now a scalar. What can be achieved with such feedback? Why is it that we use feedback to the states in the observer problem but not in the controller problem?

##### 4.1-3.

Is the asymptotic observer of an observable system itself observable for all possible  $I$ ? Give a proof.

##### 4.1-4. Feedback and Observability

Show that the observability of a realization  $\{A, b, c\}$  is not invariant under general state feedback  $u \mapsto r - kx$  but is invariant under linear or nonlinear output feedback  $u \mapsto r - f[x(s)], s \leq t$ .

##### 4.1-5. Another Approach to Observers [3]

If  $\dot{x}(t) = Ax(t) - bu(t)$ ,  $y(t) = cx(t)$ ,  $x(t_0) = x_0$ , let  $\hat{x}(t) = Fx(t)$ ,  $Fx(t) = gy(t)$ ,  $\hat{x}(t_0) = \hat{x}_0$ . The second equation can be said to define an observer for the first if  $x_0 = \hat{x}_0 \Rightarrow x(t) = \hat{x}(t) = f[x(s)], s \leq t$ . Show that a necessary and sufficient condition for this is that  $F = A - kc$ ,  $h = k$ ,  $g = b$ , where  $k$  is an arbitrary  $n \times 1$  vector.

#### 4.1-6. Effects of Mismatched Models

Given a realization  $\dot{x}(t) = Ax(t) + bu(t)$ ,  $y(t) = cx(t)$ , consider an observer  $\dot{\hat{x}}(t) = \hat{A}\hat{x}(t) - v(t)$ ,  $\dot{v}(t) = \hat{c}\hat{x}(t)$ ,  $v(t) = \hat{b}(\hat{y}(t) - \hat{y})$ , where the  $\{\hat{A}, \hat{b}, \hat{c}\}$  are estimates (approximations to) of  $\{A, b, c\}$ . Show that  $\dot{\hat{x}}(t) = [\hat{A} - \hat{b}\hat{c}]\hat{x}(t) - [\delta\hat{A} - \hat{b}(\delta\hat{c})]x(t) + (\delta\hat{b})u(t)$ ,  $\dot{v}(t) = \hat{x}(t) - x(t)$ , where  $\delta\hat{A} = A - \hat{A}$ ,  $\delta\hat{b} = b - \hat{b}$ , and  $\delta\hat{c} = c - \hat{c}$ . Note: Further analysis allows us to relate these results to questions of system sensitivity (as originally conceived by Bode (1949); See W. A. Porter, *IEEE Trans. Autom. Control*, AC-22, pp. 144-146, Feb. 1977).

#### 4.1-7.

Design an observer for the oscillatory system  $\dot{x}(t) = v(t)$ ,  $\dot{v}(t) = -\omega_0^2 x(t)$  using measurements of the velocity  $v(\cdot)$ . Place both observer poles at  $s = -\omega_0$ .

#### 4.1-8. Station-Keeping Satellite

For the station-keeping satellite of Example 3.3-2, design an observer using measurements  $v(\cdot)$  of azimuthal position perturbation. Place the observer poles at  $s = -2\omega$ ,  $s = -3\omega$ ,  $s = -3\omega \pm j3\omega$ , which means that the estimate errors will decay in about  $2\frac{1}{2}$  days ( $\doteq \frac{1}{\omega}$ , where  $\omega = 2\pi/29.3$  rad/day).

#### 4.1-9. Deadbeat Observers in Discrete-Time

Develop observer designs for discrete-time systems and show how to design an observer for which the error will go to zero in no more than  $n$  steps, where  $n$  is the number of states.

## 4.2 COMBINED OBSERVER-CONTROLLER COMPENSATORS

We were led to the observer problem by the need to obtain the states for use in the controller. Now we only have asymptotically correct estimates of the states rather than the states themselves, and a natural question is whether our previous result on arbitrary pole placement via state-variable feedback will continue to hold when only such estimates of the actual states are available. We have no option but to see what happens with the estimators that we have available.

In steady state there should clearly be no loss in using the asymptotic observer, since the error in the estimates will be zero. Therefore, as we shall soon confirm, the transfer function of the combined observer-controller will be just that of a pure controller (with perfect state feedback). However, this is not our major worry since, as noted in Sec. 3.1.1, there are many ways of arranging for a desired transfer function. The real question for the present scheme is whether the incorporation of the observer dynamical system into the feedback loop will affect the *stability* of the overall system--the point being that interconnections of stable subsystems may lead to unstable overall systems (cf. Sec. 3.1). However, we shall now prove the nice result that *incorporation of a stable asymptotic observer does not impair stability*.

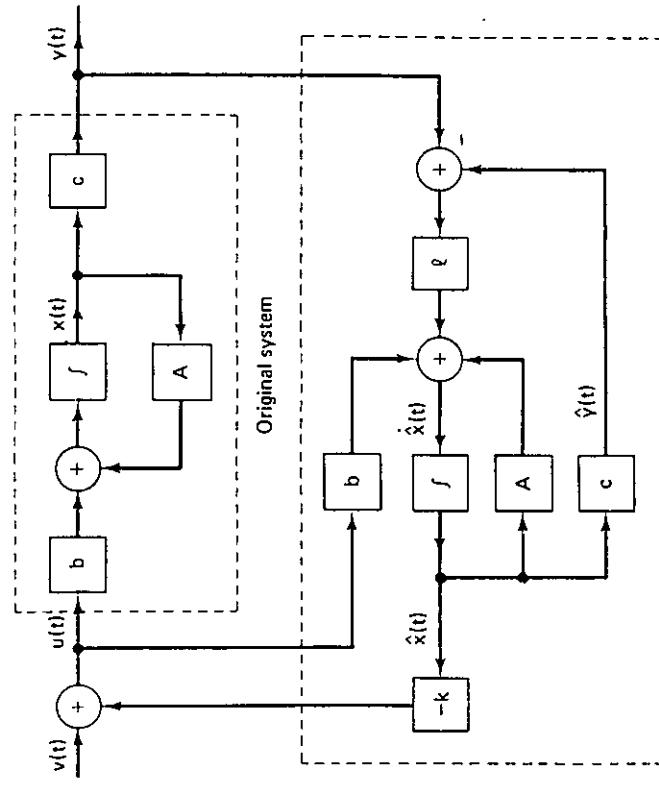


Figure 4.2-1. Combined observer-controller. Note that the observer is driven by both the output  $y(\cdot)$  and the input  $u(\cdot)$  [ $= v(\cdot) - k\hat{x}(\cdot)$ ] of the original system.

The error in the estimates obeys the equation

$$\begin{aligned}\dot{\hat{x}}(t) &= \dot{x}(t) - \dot{\hat{x}}(t) \\ &= A\dot{x}(t) - A\hat{x}(t) - lcx(t) + lc\dot{\hat{x}}(t) \\ &= (A - lc)\dot{x}(t), \quad \dot{\hat{x}}(0-) = x_0 - \hat{x}_0\end{aligned}\quad (1)$$

and the effect of the input,  $u(\cdot) = v(\cdot) - k\hat{x}(\cdot)$ , is cancelled out. To study the overall system we arrange the equations for  $x(t)$  and  $\hat{x}(t)$  in matrix form as

$$\begin{bmatrix} \dot{x}(t) \\ \dot{\hat{x}}(t) \end{bmatrix} = \begin{bmatrix} A & -bk \\ lc & A - lc - bk \end{bmatrix} \begin{bmatrix} x(t) \\ \hat{x}(t) \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} v(t), \quad \begin{bmatrix} x(0-) \\ \hat{x}(0-) \end{bmatrix} = \begin{bmatrix} x_0 \\ \hat{x}_0 \end{bmatrix} \quad (2a)$$

$$y(t) = cx(t) \quad (2b)$$

**Calculation of the Transfer Function.** When computing transfer functions we have to assume zero initial conditions. Doing this and then taking Laplace transforms, we obtain

$$\begin{aligned}sX(s) &= AX(s) - bk\hat{X}(s) + bV(s) \quad (3a) \\ s\hat{X}(s) &= lcX(s) + (A - lc - bk)\hat{X}(s) + bV(s) \quad (3b)\end{aligned}$$

Then eliminating  $\hat{X}(s)$ , we get

$$\begin{aligned}[sI - A + bk(sI - A + bk - lc)^{-1}lc]X(s) \\ = [I - bk(sI - A + bk - lc)^{-1}b]V(s)\end{aligned}\quad (4)$$

It seems difficult to proceed, but if we use the important matrix-inversion identity (Exercise A-21)

$$[I + C(sI - A)^{-1}B]^{-1} = I - C(sI - A + BC)^{-1}B$$

we can rewrite (4) as

$$\begin{aligned}bV(s) &= [I + bk(sI - A + lc)^{-1}]sI - A + lc \\ &\quad - [I - bk(sI - A + bk - lc)^{-1}]lc]X(s) \\ &= (sI - A + lc + bk - lc)X(s) \\ &= (sI - A + bk)X(s)\end{aligned}$$

Therefore

$$X(s) = (sI - A + bk)^{-1}bV(s)$$

so that the transfer function of the overall system is

$$H_{o-c}(s) = c(sI - A + bk)^{-1}b \quad (5)$$

(where the subscript  $o-c$  stands for observer-controller). These are the same equations we would have had with perfect state feedback; i.e., the overall

transfer function is just that of the controlled system and does not depend on the dynamics of the observer. In other words, the transfer function does not depend on how quickly the error in the state estimates goes to zero.

The explanation is that when the initial conditions for  $x(\cdot)$  and  $\hat{x}(\cdot)$  are both zero, and therefore the same, then of course the observer, when driven with the same inputs as the original system, will have the same outputs as the original system (see our earlier discussions in Sec. 4.1). That is,  $x(t) = \hat{x}(t)$  when  $x(0) = 0 = \hat{x}(0)$ . Therefore, when making transfer function calculations, the asymptotic observer is the same as the perfect observer!

This is a nice result, but our major concern of course is with the modes of the overall realization. These will be the roots of the characteristic equation of realization (2), namely

$$\begin{aligned}a_{o-c}(s) &= \det \begin{bmatrix} sI - A & bk \\ -lc & sI - A + lc - bk \end{bmatrix} \\ &= a_{\text{cont}}(s)a_{\text{obs}}(s), \text{ say.}\end{aligned}\quad (6)$$

We can simplify this determinant by obvious row and column transformations to obtain,

$$\begin{aligned}a_{o-c}(s) &= \det (sI - A + bk) \det (sI - A + lc) \\ &= a_{\text{cont}}(s)a_{\text{obs}}(s)\end{aligned}\quad (7)$$

That is, the characteristic polynomial of the overall system is just the product of the characteristic polynomial of the observer and the characteristic polynomial of the controlled system assuming perfect knowledge of the states. This is nice, because it means that the natural frequencies or modes of the overall system can always be arranged to be stable. In fact, they can be chosen completely arbitrarily (if the original realization is controllable and observable). For by choosing  $k$  as in Sec. 3.2, we can make  $a_{\text{cont}}(s) = \det (sI - A + bk)$  arbitrary. Therefore, if  $a_{\text{cont}}(s)$  is stable (i.e., has no roots with zero or positive real parts) and  $a_{\text{obs}}(s)$  is stable, then  $a_{o-c}(s)$  is stable. This is a fundamental result and one that is not a priori obvious because, as noted earlier, there are many situations where the interconnection of stable systems leads to an unstable system.

Another useful consequence of (7) and (8) is that the *controller and observer can be designed independently of each other*. Whether the true states are available, or only asymptotically correct estimates of the states, is immaterial to the calculation of the feedback gains  $k$ ; similarly, the dynamics of the asymptotic observer can be calculated from knowledge of  $A$  and  $c$  without caring if the observer is to be combined with a feedback controller or not. This is the so-called *separation property* of the observer-controller design procedure. However there will generally be a deterioration in the transient response of the combined system—see Example 4.2-1 below.

**Implementation of the Observer.** The fact that the observer has as many states as the original system might be disturbing—does this mean that we essentially have to replicate the original system in order to achieve the above nice results? The original system might be a rather complicated power plant, chemical processor, etc. Fortunately, we recognize that except under very special circumstances the observer need not be constructed in the same way as the original system—in particular, we can build the observer with (miniatized) electronic circuitry, with great advantages of size, ruggedness, cost, etc. In fact, with modern developments in integrated electronics, we can readily envisage the use of very high-order observers for help in the control of quite complicated physical systems. Thus, in several applications, it will be feasible to implement the observer via a microprocessor specially designed to integrate the observer state equations.

**Summary.** We have now obtained some clear and definite answers to several of the questions we raised in Sec. 3.1.1. Briefly, by using feedback of the states of a completely controllable and completely observable realization of the original transfer function, we can obtain a new internally stable realization whose natural frequencies are completely under our control. While we have obtained this result only for scalar (single-input, single-output) systems, it also has a fairly natural extension to the multivariable case (Sec. 7.4) and in this sense represents perhaps the major triumph in the state-space approach to linear systems. These points are worth further discussion, which we shall begin to pursue further in Example 4.2-3.

#### Example 4.2-1.

In this example we shall illustrate the sort of response obtained by using an observer to generate state estimates for feedback. The results were obtained by simulation on an analog computer.

The system we shall consider corresponds to the pointer-balancing problem of Example 3.3-1, except that we now take  $g/L = 1$  for convenience of simulation. We then have

$$\begin{aligned}\dot{z} &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} z + \begin{bmatrix} 0 \\ -1 \end{bmatrix} u \\ y &= \begin{bmatrix} 1 & 0 \end{bmatrix} z\end{aligned}$$

and the system modes are at  $\lambda = \pm i$ . An analog simulation of the system is shown in Fig. a.

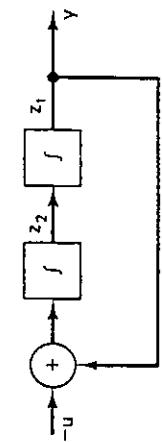


Figure a. Simulation of system.

The system is clearly unstable, as indicated in Fig. b by the response to a nonzero initial condition.

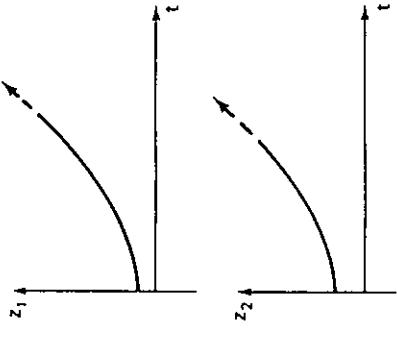


Figure b. Unstabilized response.

We use state feedback  $u = -kz$  to move the system poles to  $-0.5 \pm j0.5$ . The required  $k$  is easily obtained as  $k = [-1.5 \quad -1]$ . The decay of the system from an initial state to the origin, using this direct state feedback, is shown in Fig. c.

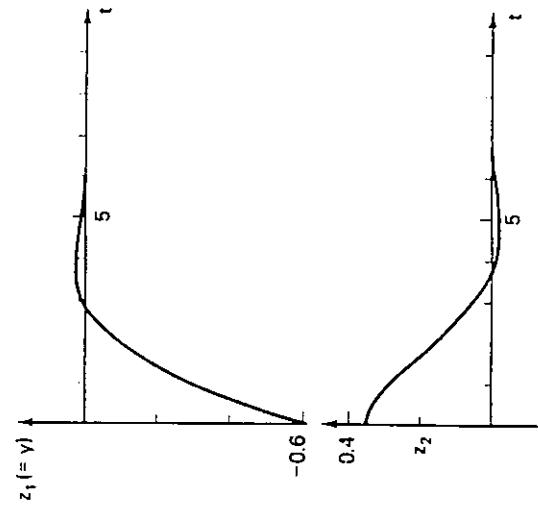


Figure c. Response with direct state feedback.

We now construct an observer for the system. The observer modes are chosen to lie at  $-1 \pm j1$ . (They are faster than the desired closed-loop poles but slow enough for us to see clearly their effect on the system response. A brief discussion of the factors involved in selecting reasonable observer modes is presented in Sec. 4.4.)

The observer gain is readily found as  $I = [2 \ -3]$ , and the observer is then easily constructed.

We now use the observer states to generate the required control,  $u = -k\hat{z}$ , and the resulting compensator is shown in Fig. d.

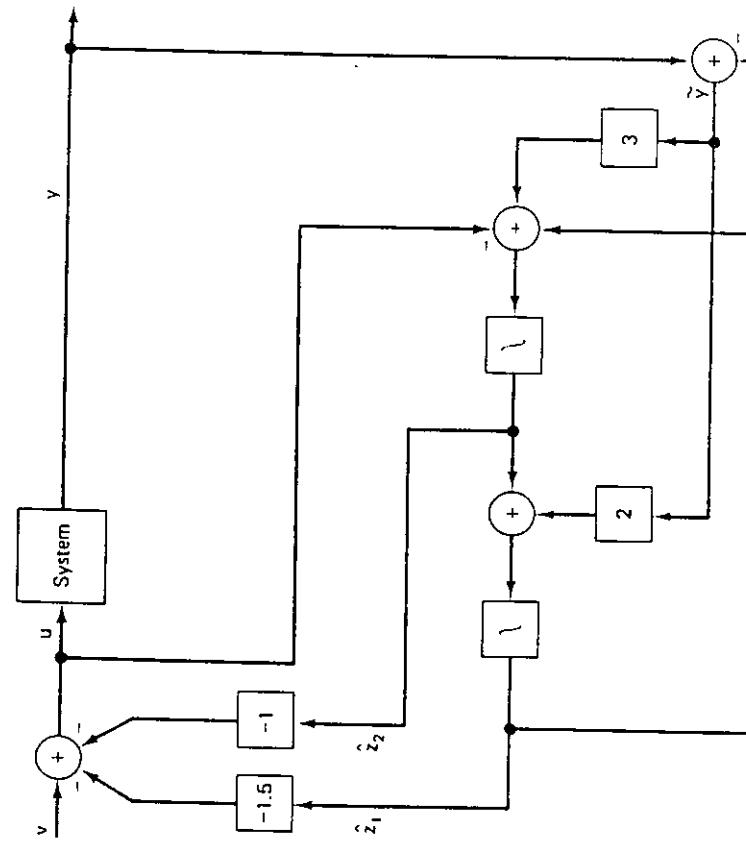


Figure d. Compensator structure.

The response of this closed-loop system is indicated in Fig. e. Note the manner in which the observer error decays; it is clear that the observer error goes to zero about twice the rate at which the system settles to zero, as is to be expected from the fact that the error poles are twice the closed-loop poles. Comparison of Figs. c and e indicates the deterioration in response due to the use of estimated rather than actual states for feedback. In Fig. f we compare the controls resulting from direct state feedback and from use of the observer-controller compensator structure; the initial large-amplitude oscillation of the control in the latter case is again a result of the large initial uncertainty as to the true state of the system.

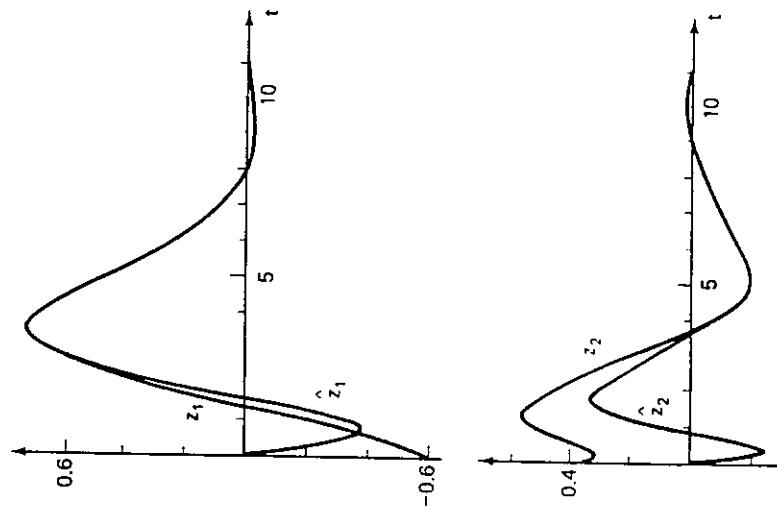


Figure c. Response using direct state feedback.

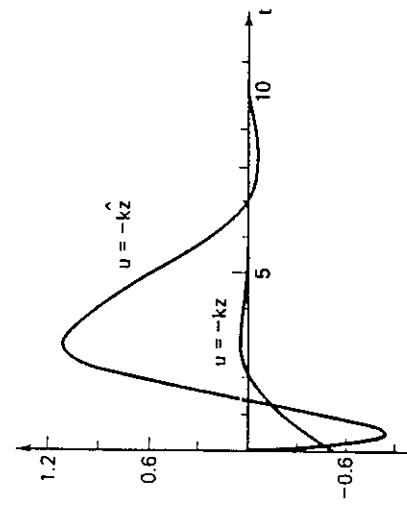


Figure e. Response using compensator of Fig. d.

Figure f. Comparison of controls generated by direct state feedback and using the compensator.

**Example 4.2-2.** Constant Disturbances and Integral Feedback

For a system driven by a *constant* unknown disturbance  $w$ , design an observer to estimate  $w$ , and use this to compensate for the disturbance.

**Solution.** We have

$$\begin{aligned}\dot{x} &= Ax + bu - bw \\ \dot{w} &= 0 \\ y &= cx\end{aligned}$$

where  $u$  is the control input and  $y$  the observed output. The constant disturbance  $w$  is modeled as the output of an undriven integrator. We then have the *augmented* system shown in Fig. a.

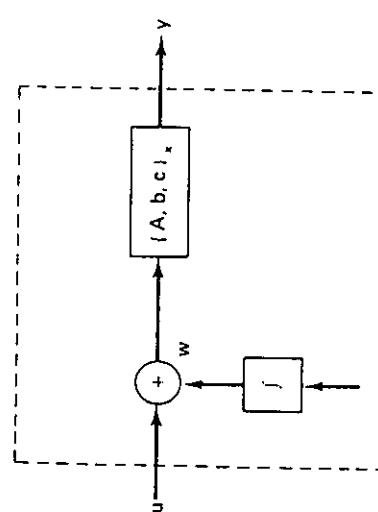


Figure a.

If now we had an estimate  $\hat{w}$  of  $w$ , we could set  $u = -\hat{w}$  to attempt to cancel out the disturbance. This motivates us to set up an observer to estimate  $w$ .

An observer for the augmented system is given by

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{w}} \end{bmatrix} = \begin{bmatrix} A & b \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \hat{w} \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} u + \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} y$$

$$\begin{aligned}x(0) &= 0, \\ \hat{w}(0) &= 0\end{aligned}$$

where  $l$  is an  $(n+1) \times 1$  vector. Partitioning  $l$  as  $[l_1 \ l_2]'$ , with  $l_2$  a scalar, we get

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{w}} \end{bmatrix} = \begin{bmatrix} A - l_1 c & b \\ -l_2 c & 0 \end{bmatrix} \begin{bmatrix} x \\ \hat{w} \end{bmatrix} + \begin{bmatrix} b \\ 0 \end{bmatrix} u + \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} y$$

$$\begin{aligned}x(0) &= 0, \\ \hat{w}(0) &= 0\end{aligned}$$

The observer structure is then as shown in Fig. b. Now if the augmented system is observable, we can choose  $l$  so as to obtain arbitrary error decay modes and thus ensure that  $\hat{w}$  approaches  $w$  asymptotically. Let us temporarily ignore the question

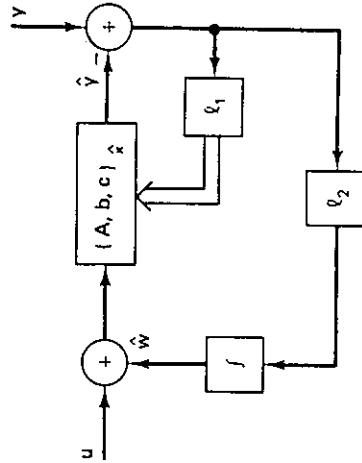


Figure b.

of observability. We also make a particular choice of  $l$  as  $[0 \ l_2]$ , which will simplify our observer considerably; for the moment we shall not worry about whether this can still ensure that the error-decay modes are stable. Now, on setting  $u = -\hat{w}$ , our observer equation reduces to

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{w}} \end{bmatrix} = \begin{bmatrix} A & 0 \\ -l_2 c & 0 \end{bmatrix} \begin{bmatrix} x \\ \hat{w} \end{bmatrix} + \begin{bmatrix} 0 \\ l_2 \end{bmatrix} y, \quad x(0) = 0, \quad \hat{w}(0) = 0$$

Since the equation for  $x$  is undriven and the initial condition is zero, we have  $x \equiv 0$ , and our observer is simply

$$\dot{\hat{w}} = l_2 y, \quad \hat{w}(0) = 0$$

The resulting overall compensation scheme is shown in Fig. c (where the dashed lines indicate parts that drop out of the compensator).

The result of the above procedure is thus precisely the technique that was presented in Sec. 3.2.3 for compensation of constant unknown disturbances, namely *integral feedback*. It arises here in a more natural and motivated manner.

The question we have so far avoided is whether proper choice of  $l_2$  can ensure that  $\hat{w}$  approaches  $w$ .

Our earlier observer equation shows that the observer error behavior is determined by the roots of (see Exercise A-11)

$$\begin{aligned}\alpha(s) &= \det \begin{bmatrix} sI - A & -b \\ l_2 c & s \end{bmatrix} = \det(sI - A) \det[s + l_2 c(sI - A)^{-1}b] \\ &= s\alpha(s) + l_2 b(s) = 0\end{aligned}$$

We assume now that the original system  $\{A, b, c\}$  was stable (or stabilized) and hence that  $\alpha(s)$  is stable, i.e., has roots with strictly negative real parts. It can then be shown that proper choice of  $l_2$  can give stable  $\alpha(s)$  if and only if  $b(s)$  has no root

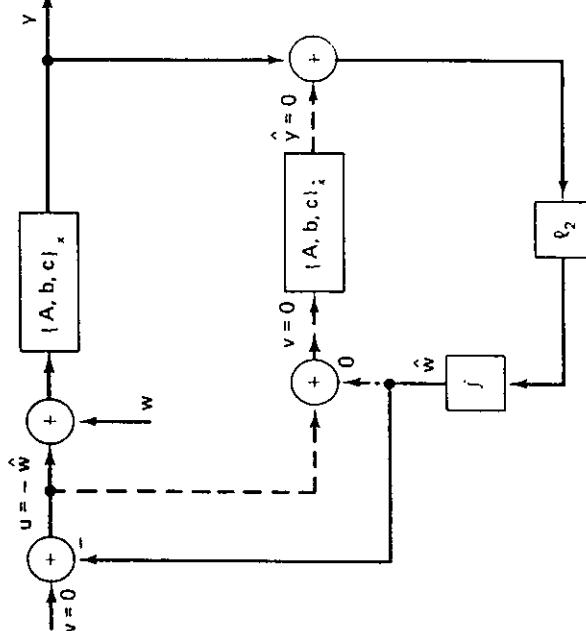


Figure c.

at the origin. [Necessity is obvious, since if  $b(s)$  contained  $s$  as a factor, then  $\alpha(s)$  would also contain this unstable factor; initial observer error would then not decay to zero. Sufficiency is easily proved using root-locus-type arguments, but we shall omit these here.]

The case where  $b(s)$  has a root at the origin (i.e., where the original system has a zero at the origin) is, however, rather trivial, since the presence of a zero at the origin makes the output of the system insensitive to a constant input anyway (see Sec. 3.2.3). Another way of understanding the above fact is to note (see Example 2.4-4) that if  $b(s)$  contained  $s$  as a factor then the augmented system (from the disturbance initial conditions to  $y(\cdot)$ ) would be unobservable due to the cancellation of the term  $s$ .

This is a simple example of the so-called "internal model" principle for the rejection of disturbances.

#### Example 4.2-3. A Review of the Overall Observer-Controller Design Problem

We are given a system with transfer function  $H(s)$ . The system may be a process, power plant, machine, market, bacterial colony, etc. We have control over one variable, the *input*, and are interested in controlling the behavior (evolution, response) of some other system variable, the *output*. The system given to us is characterized by a certain input-output behavior, say,

$$H(s) = \frac{(s-2)(s-3)}{s(s-1)(s-4)} = \frac{s^2 - 5s + 6}{s^3 - 5s^2 + 4s}$$

$$k_e = (\alpha - d) = [12 \quad 8 \quad 10]$$

Now how realistic is the assumption that the states are available? What is involved

The system may be unstable (as above), or its response to a given input may not be what we wish it to be. Suppose we wish to relocate the poles at  $-1 \pm j$  and  $-5$  (guided perhaps by the optimum root locus—draw this!).

Cascade compensation, which cancels the unstable poles, is unsatisfactory.

Output feedback,  $u \rightarrow v - ky$ , will not suffice; we have one free parameter and three quantities to be independently adjusted.

Position, velocity, and acceleration feedback,  $(k_1 s - k_2 s^2)y$ , gives us enough parameters, but the system order increases to 4.

$$H(s) = \frac{(s-2)(s-3)}{s(s-1)(s-4) - (k_2 - k_1)s^2 + k_1^2 s^3 - 2k_1 s^2 - 3s + 3}$$

Other strategies can be tried, as in Sec. 3.1.1, without leading to any obvious clear answer, and therefore we shall try to approach the problem with state-space ideas. We know that to modify the system arbitrarily we must know what the system is doing. And this (by definition, or by construction, . . . ) is told us by the state variables. Suppose then that someone gives us these state variables (or states, for short). This statement by itself is meaningless. One cannot talk of states apart from the realization they correspond to. To make sense, we have to be given measurements (terminals, . . . ) that define the state variables  $x(\cdot)$  of a realization [4, b, c]. How was the realization obtained by the person who gives us the states? Again, it is meaningless to talk of the realization. What the person had was a set of equations describing the system; these were transformed (and perhaps linearized) to a set of first-order differential equations in the variables  $x(\cdot)$ , which are now made available. If the realization is controllable, we can use linear state feedback,  $k_x(\cdot)$ , with  $k$  chosen to obtain any desired free response (natural frequencies).

Assume that we have available the states of a controller-form realization of  $H(s)$ ,

$$\begin{aligned} A_c &= \begin{bmatrix} -5 & -4 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, & b_c &= \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \\ c_c &= [1 \quad -5 \quad 6] \end{aligned}$$

We would like a new characteristic polynomial

$$\begin{aligned} \alpha(s) &= (s+5)(s+1+j)(s+1-j) \\ &= s^3 + 7s^2 + 12s + 10 \end{aligned}$$

Now with state feedback,

$$A_c - b_c k_e = \begin{bmatrix} -7 & -12 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

which shows that

in making the states available? Note that by making the states available we are in effect defining  $n$  new outputs. It is then perhaps not so surprising that we can achieve so much more (namely arbitrary pole location) with state feedback than with feed-back of the single original output. The obvious cost involved in making the states available is the cost of the sensors required to measure the states and the transducers needed to convert these measurements to useful quantities for control.

(There is a whole *dual* range of possibilities for system modification that arises from the definition of additional *inputs*, or, equivalently, from postulating access to the *outputs* of the *integrators* of the realization, but we shall not go into this here.)

So what do we do if the states are not available (or the cost of making them available is prohibitive)? If the system is observable, we can determine its state at any time by differentiation of the output and input. This is unrealistic because of noise considerations. If, however, we know the state of the system at some instant, we can construct a dummy system (realization) with this given initial state and drive it with the same input as the original system from that time on. (Note that this dummy system is only mathematically equivalent to the original—its actual implementation may be quite different, perhaps an analog-computer simulation, or often a *digital-computer algorithm* for integrating the state equations.) The (simulated) states are now available.

It is desirable to use an additional input to the dummy, namely  $y - \hat{y} = \bar{y}$  (the difference between the actual and dummy outputs). This allows asymptotic correction of errors in initial-state determination and in fact obviates the need for initial-state determination altogether. The system is still required to be observable, however, in order that the error dynamics may be arbitrarily chosen.

Returning to our example, we construct a dummy system  $\{\mathcal{A}_e, \mathcal{b}_e, \mathcal{c}_e\}$  with states  $\hat{x}_e$ , driven by an additional input  $I_e(y - c_e \hat{x}_e)$ . Suppose we want the observer poles at  $-6 \pm j6$  and  $-6$  (a "reasonable" choice for our given closed-loop pole locations). Then  $I_e$  must be chosen such that the characteristic polynomial of  $\mathcal{A}_e - I_e \mathcal{c}_e$  is

$$\alpha_e(s) = (s + 6)(s^2 + 12s + 72) = s^3 \div 18s^2 + 144s + 432$$

Some algebra then shows that

$$I_e = [807.5 \quad 127.5 \quad -24.5]$$

The combined observer-controller can now be implemented as in Figure 4.2-1. In practice, however, there are still problems to be overcome in the actual realization, one of the most important being *scaling*: We may need to transform the realization to a form where the required gains are of reasonable magnitude—see Exercise 2.2-19.

#### Exercises

##### 4.2-1. Alternative State Equations

The fact that  $\hat{X}(s) = X(s) - \hat{X}(s) = 0$  in (3a) and (3b) suggests that perhaps a more convenient set of state variables for the observer-controller system is  $[x, \hat{x}]$ . Write state equations for these variables, and use them to calculate  $H_{o,c}(s)$  and  $a_{o,c}(s)$ .

**4.2-2. Controllability and Observability of the Observer-Controller Realization**  
The realization in Fig. 4.2-1 is clearly not minimal.

- a. Show that this realization is not controllable. What are the noncontrollable state variables?
- b. Show that the realization will be nonobservable if and only if at least one of the following holds:
  - (1)  $\{\mathcal{A} - I_c\}$  is not observable,
  - (2)  $\{\mathcal{C}, \mathcal{A} - I_c\}$  is not observable, or
  - (3) A pole of the observer cancels a zero of the original transfer function  $H(s)$ .

It is, of course, assumed that  $\{\mathcal{A}, b, c\}$  is minimal.

#### 4.2-3.

In the combined controller-observer design we select  $k$  and  $I$  so that  $a_e(s)$  and  $a_o(s)$  both have poles in the left half plane. Is it true that the resulting design must be stable even if the loop is broken open at  $y$ , for instance? Explain your answer briefly.

#### 4.2-4.

Consider the undamped harmonic oscillator  $\dot{x}_1(t) = x_1(t)$ ,  $\dot{x}_2(t) = -\omega_0^2 x_1(t) \doteq u(t)$ . Using an observation of velocity,  $y(\cdot) = x_2(\cdot)$ , design an observer-state-feedback compensator to control the position  $x_1(\cdot)$ . Place the state-feedback controller poles at  $s = -\omega_0 \mp j\omega_0$  and both observer poles at  $s = -\omega_0$ .

#### \*4.3 REDUCED-ORDER OBSERVERS

The observer design method described in Sec. 4.1 is due, as noted earlier, to Bertram (1961) and Bass (1963). In his Ph.D. thesis (Stanford, 1963) Luenberger took a different approach to the observer problem. We shall not really discuss this approach here but shall focus instead on its most significant contribution. The observer obtained in Sec. 4.1 had  $n$  states, where  $n$  was the order of the realization whose states were being observed. Luenberger [3] pointed out that the order of the observer can actually be less than this because the observed output provides a linear relationship  $y(t) = cx(t)$  between the state variables. Therefore it suffices to observe  $n - 1$  of the states and then to calculate the final one from this linear relationship.

The reduction by one in observer dimension is not particularly significant, especially when the observer is implemented with (integrated-circuit) electronic logic. However, the question has some interesting theoretical aspects, which we shall explore here. Moreover, for multiooutput systems, somewhat more substantial reductions can be obtained (cf. Sec. 7.3). In this section, we shall essentially follow the historical approach to the topic, which starts with the unreduced observer of Sec. 4.1. The reduced-order observer will arise in a more direct way in the frequency-domain design procedure to be presented in Sec. 4.5.1.

We shall first explain why reduction is possible and describe Luenberger's idea [3] for exploiting it. However for the detailed calculations we shall use a different technique [4] and [5].

Let us begin with a realization  $\{A, b, c\}$  and make an easily found (and non-unique) similarity transformation such that the output vector has the form

$$c = [0 \quad \cdots \quad 1] \quad (1)$$

and

$$y = [0 \quad \cdots \quad 1]x(t) = x_r(t) \quad (2)$$

Therefore we have clearly displayed the fact that one state is directly observable,

$$\dot{x}_n(t) = y(t) = x_r(t)$$

and we only need to estimate the states  $x_r = [x_1 \quad \cdots \quad x_{n-1}]$ , where the subscript  $r$  is used to denote the states that remain to be estimated. It should be possible to estimate these  $n - 1$  states by an  $(n - 1)$ -dimensional observer. This is true, but the design has to be approached with a little care. Thus, consider the partitioned state equations

$$\begin{bmatrix} \dot{x}_r(t) \\ \dot{x}_n(t) \end{bmatrix} = \begin{bmatrix} A_r & b_r \\ c_r & a_{nn} \end{bmatrix} \begin{bmatrix} x_r(t) \\ x_n(t) \end{bmatrix} + \begin{bmatrix} g_r \\ g_n \end{bmatrix} u(t), \quad y(t) = x_n(t) \quad (3)$$

where the reasons for our notation will become clear later. Now we set up a dummy system for  $\dot{x}_r(t)$ ,

$$\dot{x}_r(t) = A_r \dot{x}_r(t) + b_r y(t) + g_r u(t) \doteq (\text{feedback term}) \quad (4)$$

The feedback term should be derived from the error  $x_r(t) - \hat{x}_r(t)$ , but not only is this term inaccessible, just as  $x(t) - \hat{x}(t)$  was in the  $n$ -th-order observer of Sec. 4.1, but the accessible quantity  $[y(t) - c\hat{x}(t)]$  used in Sec. 4.1 is identically zero [with our choice (2) for  $c$ ] if we insist that  $\dot{x}_n(t) = y(t)$ . Therefore, it seems difficult to get a feedback estimator for  $\dot{x}_r(t)$ . However, it is worthwhile to persist a bit.

Note that without feedback, the *open-loop* equation (4) for the error  $\dot{x}_r(t) = \hat{x}_r - x_r(t)$  is

$$\dot{x}_r(t) = \dot{x}_r(t) - \hat{x}_r(t) = A_r \dot{x}_r(t), \quad \dot{x}_r(0) = x_r(0) - \hat{x}_r(0) \quad (5)$$

[Note again that we might as well have assumed the *known* inputs  $b_r y(t)$  and  $g_r u(t)$  to be zero.] We had a similar equation in Sec. 4.1 for the  $n$ -dimensional open-loop observer,

$$\dot{\hat{x}}_r(t) = A_r \hat{x}_r(t), \quad \dot{\hat{x}}_r(0) = x(0) - \hat{x}_r(0) \quad (6)$$

<sup>†</sup>One such realization is obtained from our usual observer form  $\{A_0, b_0, c_0\}$  by merely labeling the states in the reverse order to our conventional order. In fact, the reader will find it helpful to interpret various results stated for observable realizations by assuming the realizations to be in observer (or related observer) form. For actual computation, however, it may be generally simpler to avoid transformation all the way to observer form and to merely find some convenient transformation that gives  $c = [0 \quad \cdots \quad 0 \quad 1]$ .

and we concluded that the error dynamics were beyond our control because the eigenvalues of  $A$  could not be changed by similarity transformations taking  $A$  to  $T^{-1}AT$ . However, it was at this point that Luenberger made the key observation that even though the eigenvalues of a matrix  $A$  could not be changed by similarity transformations, this is *not* true of eigenvalues of *submatrices* of  $A$ . (Note that if  $\tilde{A} = T^{-1}AT$ , it is unlikely that [cf. (3) for the notation]  $\tilde{A}_r$  is similar to  $A_r$ .) Luenberger showed in fact that if the original equations are observable, then state transformations can be found that will yield a *submatrix* with arbitrary eigenvalues (see Exercise 4.3-1).

Luenberger's technique is ingenious and makes interesting use of different state realizations. However, it is somewhat confusing (at least initially) to keep changing realizations, since after all we are basically interested in only one realization, viz., the one that we assumed for the original system that was to be controlled by state feedback.

**Another Approach (Gopinath [4] and Cumming [5]).** For another approach we can return to the original equations (3) and note that so far we have only really examined the subset of equations describing the evolution of  $x_r(\cdot)$ . A direct approach using this set did not succeed—see (4)—because we could not obtain a suitable *feedback term*. However, we did not use the remaining equation

$$\dot{x}_n(t) = a_{nn}x_n(t) + c_r x_r(t) + g_n u(t)$$

which certainly does provide some more information about  $x_r(\cdot)$ . Thus, note that if we define

$$\dot{y}_r(t) = \dot{y}(t) - a_{nn}y(t) - g_n u(t) \quad (7)$$

then we can write

$$\dot{x}_r(t) = A_r x_r(t) + b_r y(t) + g_r u(t) \quad (8)$$

$$\dot{y}_r(t) = c_r x_r(t) \quad (9)$$

The function  $y_r(\cdot)$  is completely determined by  $y(\cdot)$  and  $u(\cdot)$  and can therefore be regarded as an observation process. The difficulty is that  $y_r(\cdot)$  contains the derivative  $\dot{y}(\cdot)$  and cannot therefore be realistically obtained from  $y(\cdot)$ . However, if we temporarily ignore this difficulty, we see that Eqs. (8) and (9) display the  $n - 1$  states  $x_r(\cdot)$  in a form to which the feedback design of Sec. 4.1 can be applied. That is, we can set up an observer of the form

$$\dot{\hat{x}}_r(t) = A_r \hat{x}_r(t) + b_r y(t) + g_r u(t) + l_r [y_r(t) - c_r \hat{x}_r(t)] \quad (10)$$

where  $l_r$  is an  $(n - 1) \times 1$  matrix. Now

$$\dot{\hat{x}}_r(t) = (A_r - l_r c_r) \hat{x}_r(t), \quad \dot{x}_r(0) = x_r(0) - \hat{x}_r(0) \quad (11)$$

and we see that if

$$\{c_r, l_r\} \text{ is observable} \quad (12)$$

then  $\hat{x}_r(\cdot)$  can be made to go to zero arbitrarily fast (but, of course, not faster than exponentially).

The observability of  $\{c_r, A_r\}$  is to be expected from that of  $\{c, A\}$  because (3) shows that the only way  $y(\cdot) = x_n(\cdot)$  gives information about the remaining states  $x_r(\cdot)$  is through  $y_r(\cdot)$ . In any case, a simple proof of this fact can be obtained by using the PBH tests.

Therefore, we now have a reduced-order observer, except for the fact that  $y(\cdot)$  contains a derivative of the observation  $y_r(\cdot)$ , which is generally unacceptable. However, our experience with analog-computer simulations in Sec. 2.1 shows how to avoid this difficulty. The method is clear from Figs. 4.3-1 and 4.3-2. [Note that removing the differentiator (cf. Fig. 4.3-2) changes the input to the integrator from  $\dot{x}_r(\cdot)$  to  $\dot{x}_r(\cdot) - I_r y_r(\cdot)$  and changes the integrator output to  $\hat{x}_r(\cdot) - I_r y_r(\cdot)$ ; therefore,  $\hat{x}_r(\cdot)$  can be recovered by adding  $I_r y_r(\cdot)$  to the output of the integrator.] State equations can be written down from Fig. 4.3-2 to yield

$$\begin{aligned} \theta(t) &= (A_r - I_r c_r) \theta(t) + (b_r - I_r a_m \div A_r I_r - I_r c_r I_r) y(t) \\ &\quad + (g_r - I_r g_n) u(t) \end{aligned} \quad (13)$$

$$\dot{x}_r(t) = \theta(t) + I_r y(t) \quad (14a)$$

$$\dot{x}_n(t) = x_n(t) = y(t) \quad (14b)$$

as the equations for the observer.

Of course, there are other methods of implementing the equations we first obtained for the observer. Thus we could take Laplace transforms of (7)-(10) to

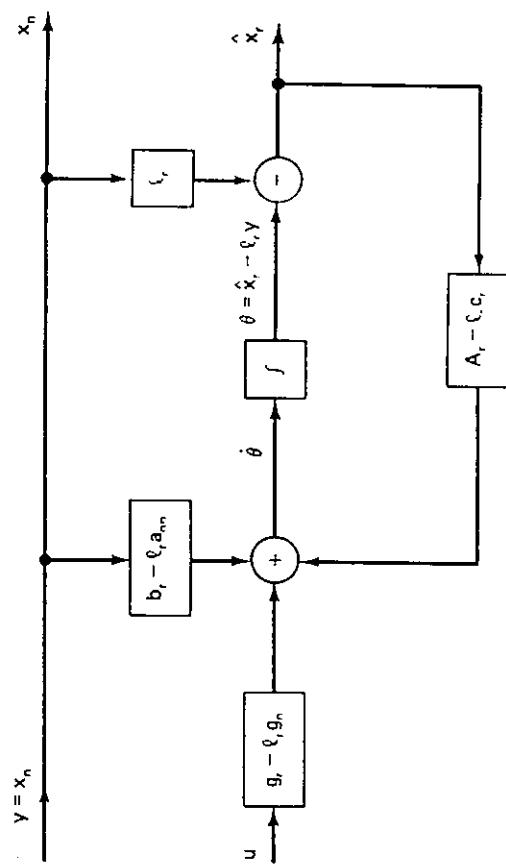


Figure 4.3-2. Equivalent implementation without differentiators.

obtain the transfer function

$$\hat{X}(s) = (sI - A_r + I_r c_r)^{-1} [s(l_r \div b_r - I_r a_m) Y(s) - (g_r - I_r g_n) U(s)] \quad (15)$$

which can be simplified and then realized in several different ways. The advantage of the first technique is that the realization (Fig. 4.3-2) it yields contains the parameters  $\{A_r, b_r, c_r, a_m, I_r\}$  in a direct way, while if (15) is used, these various parameters may be recomputed in a way that destroys their identity.

From Fig. 4.3-2 and (13)-(15) we see that the reduced-order observer can be straightforwardly incorporated into our compensator scheme.

#### Example 4.3-1. Pointer-Balancing Problem

We shall present here the results obtained by analog-computer simulation of the system of Example 4.2-1, now compensated with a reduced-order observer. Recall that the system was described by

$$\dot{z} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} z - \begin{bmatrix} 0 \\ -1 \end{bmatrix} u, \quad y = [1 \ 0] z$$

and we had determined a state-feedback gain vector to move the poles from  $\pm 1$  to  $-0.5 \pm j0.5$ , namely  $k = [-1.5 \ -1]$ .

We design from first principles a reduced-order observer with pole at  $-1.5$ . Since  $z_1$  is directly observed, we construct an observer for  $z_2$ . We thus have

$$\dot{\hat{z}}_2 = z_1 - u + I_r y \quad (\text{error signal})$$

Figure 4.3-1. Implementation using differentiators.

Noting that  $\dot{z}_1 = \dot{y} = \dot{z}_2$ , we see that a convenient error signal can be generated,

using only available measurements, as  $\hat{y} - \hat{z}_2$ , to give

$$\dot{\hat{z}}_2 = -I\hat{z}_2 + y - u + I\hat{y} \quad \text{when } \dot{\hat{z}}_2 = -I\hat{z}_2$$

Choosing  $I = 1.5$  gives the desired observer-error decay. To implement the observer without differentiation of  $y$ , we work with a new variable  $\theta = \hat{z}_2 - \hat{y}$ . Our observer equation is then

$$\begin{aligned}\dot{\theta} &= -I\theta - I^2y + y - u \\ &= -1.5\theta - 1.25y - u\end{aligned}$$

and

$$\hat{z}_1 = y, \quad \hat{z}_2 = \theta + 1.5y$$

Figure 4.3-3(a) shows the resulting compensator structure. In Fig. 4.3-3(b) we illustrate the closed-loop response using this compensator; the results shown are for  $\theta(0) = +0.5$  (a value chosen for convenience of simulation), and, as before,  $z(0) = [-0.6 \ 0.35]$ . This response may be compared with that obtained by direct state feedback and by using a full-order observer in the compensator—see Example 4.2-1. Figure 4.3-3(c) shows the control generated by the reduced-order compensator. Figure 4.3-3(d) shows the response using the compensator of (a).

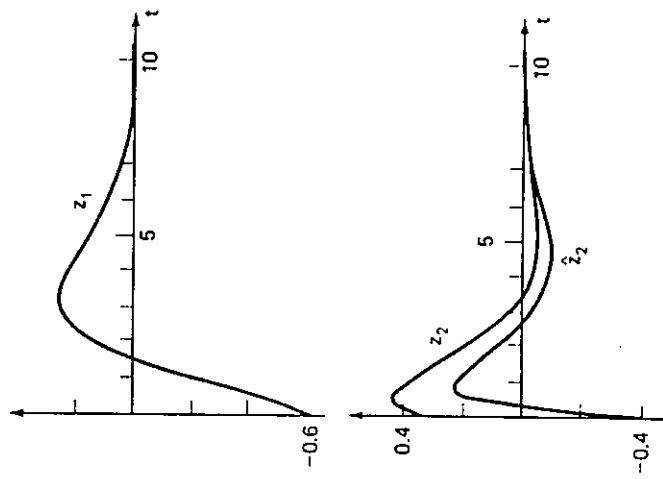


Figure 4.3-3(a). Response using compensator of (a).

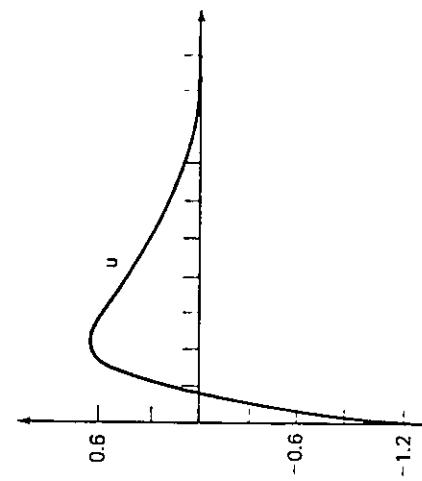


Figure 4.3-3(b). Response using compensator of (a).

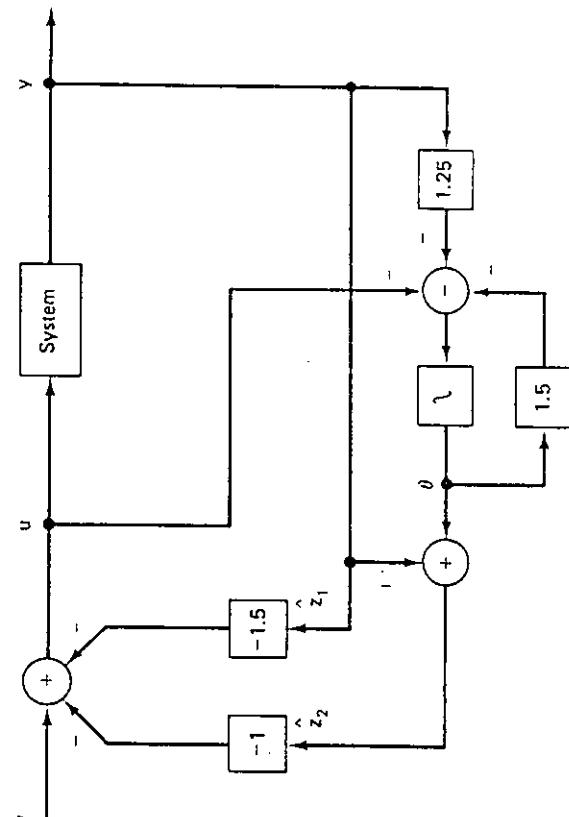


Figure 4.3-3(a). Compensator structure.

Figure 4.3-3(c). Control generated by compensator of (a).

#### Example 4.3.2. Observers and Some Classical Designs

This example is solved in a different way (using several canonical forms) in Chen [9, p. 296-298]. The problem is to design a combined observer-controller that will relocate the poles of a system with transfer function  $H(s) = 1/(s(s+1))$  at the new locations  $\{s = -1 \pm j\}$ , and to do this with a reduced-order observer that has a pole at  $s = -2$ . Because of the separation property, we can consider the controller and observer problems separately.

**Design of Controller.** We shall work with a realization of  $H(s)$  in controller canonical form. This form can be obtained by inspection,

$$A_c = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, \quad b_c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad c_c = [0 \ 1]$$

Using state feedback  $k_c x_c$ , the row vector  $k_c$  must be such that the characteristic polynomial of  $[A_c - b_c k_c]$  will be

$$a_{\text{cont}}(s) = (s + 1 + j)(s + 1 - j) = s^2 + 2s + 2$$

It follows easily that

$$k_{c1} = 1, \quad k_{c2} = 2$$

**Design of Reduced-Order Observer.** Without loss of generality we can assume that  $u(\cdot) \equiv 0$ . Then  $x_{c2}(\cdot) = y(\cdot)$ , and we only need to estimate  $x_{c1}(\cdot)$ . We have

$$\dot{x}_{c1}(t) = -x_{c1}(t), \quad \dot{y}(t) = \dot{x}_{c2}(t) = x_{c1}(t)$$

and we define

$$y_c(t) = \dot{y}(t) = x_{c1}(t)$$

Therefore the reduced-order observer is

$$\dot{x}_{c1}(t) = -x_{c1}(t) + l[y_c(t) - x_{c1}(t)]$$

where  $l$  is to be such that  $a_{\text{obs}}(s) = s + 2 = \det(sI - A_c + lC_c) = s + 1 + l$ , which gives  $l = 1$ .

**Combined Observer-Controller.** We can combine the above results as shown in Fig. 4.3-4. Note that the input  $u(\cdot) = v(\cdot) - kY(\cdot)$  has to be fed to the observer. We have also moved  $y(\cdot)$  across the integrator in the observer system so as to obviate the need for differentiation.

It will be of interest to consider various alternative configurations obtainable by transfer function manipulation. We begin by computing the transfer functions from Fig. 4.3-4 we can write the equations

$$\begin{aligned} \theta(t) &= -2\theta(t) - 2y(t) + u(t) \\ u(t) &= v(t) - 2y(t) - [r(t) - \theta(t)] \end{aligned}$$

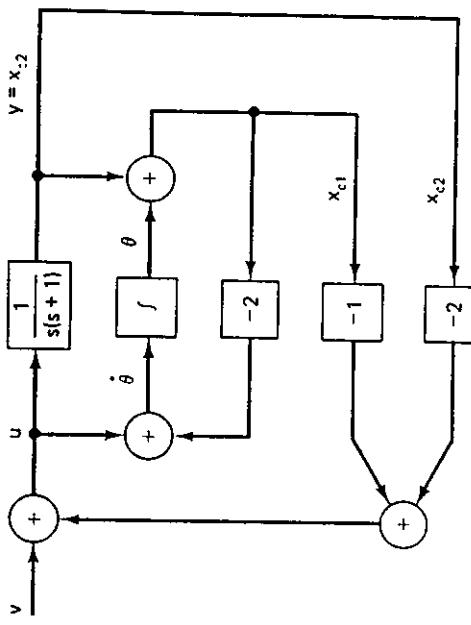


Figure 4.3-4. Combined observer-controller for Example 4.3-2.

Taking Laplace transforms and eliminating  $\theta$ , we obtain

$$U(s) = \left( -3 + \frac{2}{s-2} \right) Y(s) + V(s) - U(s)(s+2)^{-1} \quad (16)$$

Therefore

$$U(s) = \frac{-3s+4}{s-2} Y(s) + \frac{s+2}{s-2} V(s) \quad (17)$$

Relation (16) can be realized as shown in Fig. 4.3-5(a), which may be rearranged as in Fig. 4.3-5(b), a classical configuration. By the methods of Sec. 3.1.2 we can check that this realization has two hidden natural frequencies at  $s = -2$ , which are cancelled out of the transfer function from  $v(\cdot)$  to  $y(\cdot)$ . The configuration is stable but uses one integrator more than necessary.

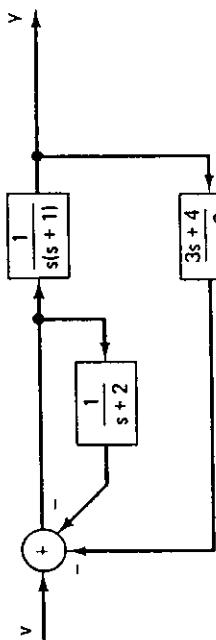
Relation (17) suggests the realization of Fig. 4.3-5(c), which can be shown to have hidden natural frequencies at  $s = -2$  and  $s = -3$ . The frequency  $s = -3$  is fixed by our original choice of transfer function poles and observer poles; in this case it happens to be a stable natural frequency, so the configuration is usable, though again it has one more integrator than necessary.

Another classical configuration, with unity feedback, is shown in Fig. 4.3-6, where  $Y(s)$  is determined by choosing it to make the overall transfer function equal to its desired value of  $(s^2 + 2s - 2)^{-1}$ . We have

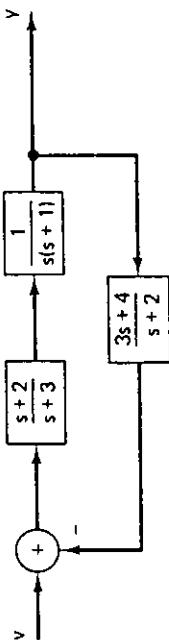
$$\frac{Y(s)H(s)}{1 + Y(s)H(s)} = \frac{1}{s^2 + 2s + 2}$$

which yields

$$Y(s) = s(s+1)^{-1}$$



(a)



(b)

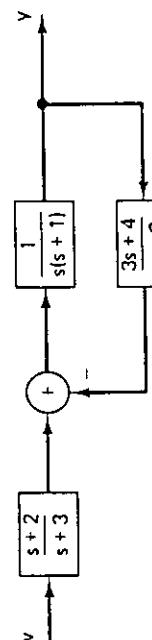


Figure 4.3-5. Various transfer function arrangements of the realization in Fig. 4.3-4. (a) Direct implementation of (16). (b) A classical configuration. (c) Direct implementation of (17).

This is a very simple solution which seems to avoid all the fuss about separate controller design, the design of reduced-order observers, and the combination of controller and observer. Why don't we just use it directly? The reason is (cf. the discussion in Sec. 3.1) that this particular procedure cannot give us full control over the hidden modes; in fact, we can see (by writing down a state-space realization or, more simply, by using the transfer function method of Exer. 2.4-7) that the configuration of Fig. 4.3-6 will have an unstable hidden mode at  $s = 0$ . However by a

slightly more complicated design procedure, we can in fact achieve fairly satisfactory results with this configuration, as we shall show in Sec. 4.5.2.

#### Exercises

##### 4.3-1. Eigenvalues of Submatrices [3]

Suppose we have a realization  $\{C, A\}$  in the (column-transposed) observer form with  $C = [0 \cdots 0 \ 1]$  and  $A$  a right-companion matrix with  $[-[a_n \ \cdots \ a_1]]^T$  in the last column. Suppose also that we wish to find a nonsingular matrix  $T$  so that  $\tilde{A}_r$ , defined as the  $(n-1) \times (n-1)$  left-hand submatrix of  $T^{-1}AT$ , has characteristic polynomial  $\det(sI - \tilde{A}_r) = \alpha(s) = s^{n-1} + \alpha_{1,n}s^{n-2} + \cdots + \alpha_{n-1}$ . Show that choosing  $T$  to have last column  $[\alpha_{n-1}, \cdots, \alpha_1]^T$  with 1s for the remaining diagonal elements and 0s elsewhere will yield  $\tilde{\alpha} = c$  and  $\tilde{A}_r$  as a right-companion matrix with last column  $-[\alpha_{n-1} \ \cdots \ \alpha_1]^T$ .

##### 4.3-2. Another Approach to Reduced-Order Observers

Show how to use the result of Exercise 4.3-1 to determine a reduced-order observer for an arbitrary but observable realization  $(A, b, c)$ . Do Examples 4.3-1 and 4.3-2 by this method.

##### 4.3-3.

Show that the subsystem  $\{C_r, A_r\}$  defined by Eqs. (4.3-3) will be observable if  $\{C, A\}$  is observable. Do this

- By using the result of Exercise 4.3-1,
- By direct algebraic manipulation of  $\mathcal{O}(C_r, A_r)$ , and
- By using the PBH tests of Sec. 2.4.3.

##### 4.3-4. Updating an Inertial Navigator with a Velocity Measurement.

An error model of the east-velocity channel of an inertial navigator is (in normalized variables)

$$\begin{bmatrix} v \\ \dot{\varphi} \\ \dot{\epsilon} \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ \varphi \\ \epsilon \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix}$$

where  $v$  = east-velocity error,  $\varphi$  = platform tilt about north axis,  $\epsilon$  = north-gyro drift, and  $w$  = gyro drift rate of change.

- Show that the open-loop eigenvalues are  $\lambda_1 = 0$ ,  $\lambda_{2,3} = \pm j$ .
- Construct an observer using  $z = v$  as the observation [made by a Doppler radar (airplane) or an EM log (ship)], placing the estimate-error poles at  $\{-10, -0.1, -0.1\}$ .
- Construct a second-order observer using  $z = v$  as the observation, placing the estimate-error poles at  $\{-0.1, -0.1\}$ .

##### 4.3-5.

We are given a system with transfer function  $H(s) = 1/s(s-2)$  and wish to use the unity-feedback configuration of Fig. 4.3-1 in order to design a compensator  $Y(s)$  that will shift both poles of  $H(s)$  to  $-1$ . No requirement is put on the zeros. You can choose any polynomial  $\beta(s)$  (of degree not greater than 2) that you wish.

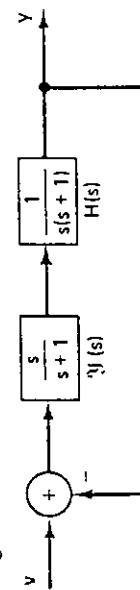


Figure 4.3-6. Classical unity-feedback configuration.

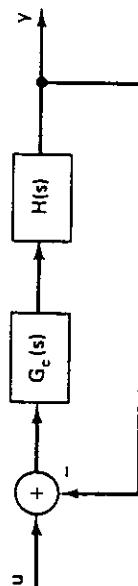


Figure a.

- a. Find such a compensator. Is  $G_c$  stable?  
 b. Do you think that the compensator procedure is satisfactory? Give reasons.  
 c. If we set up some minimal realizations for  $G_c$  and  $H$  and connect them as in Fig. a, then the overall system will be a realization of  $\beta(s)/(s+1)^2$ . Is it a minimal realization? A controllable realization? An observable realization?  
 d. Suppose that in some other design calculation we have computed a block  $G_f(s) = (13s+1)/(s+1)$  and that we now want to find  $G_c(s)$  as in Fig. b so that the overall transfer function is  $\beta(s)/(s+1)^2$  and the overall system is stable. Find such a  $G_c(s)$ .

#### 4.4 AN OPTIMALITY CRITERION FOR CHOOSING OBSERVER POLES

In Sec. 3.4.1, we discussed the relation between the modal controller and the so-called *optimal quadratic regulator*. Since the asymptotic observer is dual to the modal controller, it is natural to ask for its relationship to the dual of the quadratic regulator, and we shall briefly discuss this here.

We begin by noting that we could make the observer error decay as rapidly as we wished by putting the observer poles sufficiently far into the left half plane. In fact, one can show that in the limit the asymptotic observer reduces exactly to the ideal differentiating observer of Sec. 2.3.1. Of course, this fact brings up one reason we should not try to get too high a speed of reconstruction for the state estimates: The resulting observer will have (a high bandwidth and) a high susceptibility to the almost-inevitable measurement noise.

The conflict between speed of reconstruction and protection against measurement noise can be introduced into a quadratic optimality criterion as follows. Suppose that our actual state equations are

$$\dot{x}(t) = Ax(t) + bu(t) + gv_1(t) \quad (1a)$$

$$y(t) = cx(t) + v_2(t) \quad (1b)$$

where  $v_1(\cdot)$  and  $v_2(\cdot)$  are uncorrelated zero-mean wide-band (white) noise processes with spectral intensities unity and  $r$ , respectively; i.e., the noise power in a frequency band  $(-B/2, B/2)$  is  $B$  or  $rB$ , respectively. We assume that the unknown initial condition  $x(0)$  is also random but that it is not correlated with the noises  $v_1(\cdot)$  and  $v_2(\cdot)$ . We have no direct knowledge of  $v_1(\cdot)$  and  $v_2(\cdot)$  and therefore can only set up an observer as before,

$$\hat{x}(t) = Ax(t) + bu(t) + fv(t) - cx(t), \quad \dot{\hat{x}}(t) = 0 \quad (2)$$

But now the error will depend not only on  $x(0) - \hat{x}(0)$  but also on the noises  $v_1(\cdot)$  and  $v_2(\cdot)$ ,

$$\dot{\hat{x}}(t) = (A - fv(t))\hat{x}(t) + gv_1(t) - fv_2(t), \quad \dot{\hat{x}}(0) = x(0) \quad (3)$$

Again, by choosing  $f$  sufficiently large we can make the expected or mean value of  $\hat{x}$  tend toward zero,  $E[\hat{x}] \rightarrow 0$ , as fast as desired. But we still have to worry about the fluctuations of  $\hat{x}(\cdot)$ —choosing large values for  $f$  will accentuate the driving noise term  $fv_2(\cdot)$ . One criterion is to choose  $f$  so as to minimize the mean-square error, i.e., to minimize

$$J = E\|\hat{x}(t)\|^2 \quad (4)$$

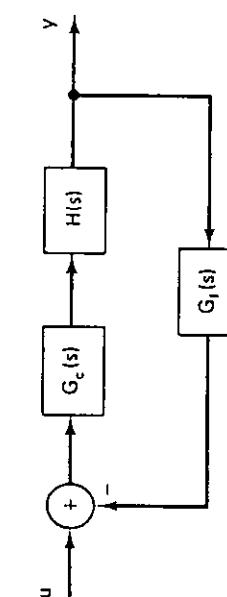


Figure b.

Compare the results of parts b and d.

#### 4.3-6.

A system is described by the transfer function  $H(s) = (s+1)/s^2 = Y(s)/U(s)$ .

- a. Find a first order dynamic feedback compensator so that the transfer function from reference input  $V$  to output  $Y$  is  $Y(s)/V(s) = [f(s^2 + 2s + 4)(s + 8)(s + 1)]$ .  
 b. Draw a block diagram of the resulting design.  
 c. Find a first order dynamic feedback compensator so that the transfer function from reference input to output is  $Y(s)/V(s) = (s^2 + 2s + 4)^{-1}(s + 1)$ .  
 d. Draw a block diagram of the design of part c.  
 e. Is the system of part a controllable from  $V$ ? Observable from  $Y$ ?

- 4.3-7. Carry out the steps required to obtain reduced-order observers for discrete-time systems.

This problem can be attacked in many different ways and in fact has an extensive literature (see, e.g., [1] and [10]–[12] among many others).

We shall not pursue the calculations here—they need some knowledge of statistics and of the solution of stochastic differential equations—but shall just quote the most relevant result. Suppose that

$$\{A, g\} \text{ is controllable, } \{c, A\} \text{ is observable} \quad (5)$$

Then the roots of the characteristic polynomial  $\det(sI - A + \bar{I}c)$  of the optimal observer can be found as the left-half-plane roots of the  $2n$ -degree polynomial

$$g(s)g(-s) + ra(s)a(-s) = 0 \quad (6)$$

where

$$a(s) = \det(sI - A), \quad g(s) = c' A \bar{d}j(sI - A)g$$

This should be compared with the rule (3.4.7)–(3.4.8) for the optimal regulator of Sec. 3.4.1—the two solutions will be identical if we make the duality interchange†

$$br^{-1}b' \longleftrightarrow gg', \quad c'c \longleftrightarrow c'r^{-1}c \quad (7)$$

The implications of (6) for the noisy observer can therefore be studied for various values of the noise parameter  $r$  by the same arguments as in Sec. 3.4.1. For example, if  $r \rightarrow \infty$  (high observation noise), the optimal observer poles are obtained as the stable roots of the original characteristic polynomial  $a(s)$  along with the unstable roots reflected across the  $j\omega$  axis. On the other hand, when  $r \rightarrow 0$ , we can have  $n - m$  poles going off to infinity in a Butterworth configuration, with the other  $m$  poles going to the zeros of  $g(s) = c' A \bar{d}j(sI - A)g$ . For intermediate values of  $r$ , the symmetric root locus plot can be useful, as we shall now illustrate.

#### Example 4.4-1. Harmonic Oscillator

For the undamped harmonic oscillator driven by a disturbance input  $w$  of unit intensity,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} w$$

determine the locus of optimum modes for an observer that uses a noisy measurement  $y = x_1 + v$ , where  $v$  is observation noise of intensity  $r$ .

It should be noted that more than one (variational) dual problem can be associated with the quadratic regulator problem and that they can be useful in obtaining approximate solutions and bounding solutions—see, for example, References [13] and [26].

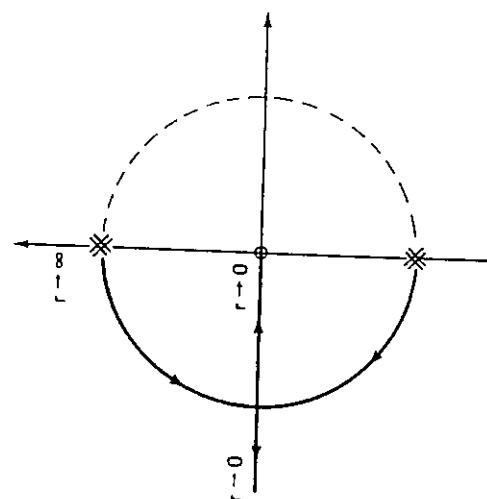
#### Solution

$$\frac{y(s)}{w(s)} = [0 \ 1] \begin{bmatrix} s & -1 \\ \omega_0^2 & s \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{s}{s^2 + \omega_0^2}$$

so the modes of the optimum observer are the stable solutions of

$$-s^2(s^2 + \omega_0^2)^{-2} + r = 0$$

The corresponding root loci are shown in the figure.



Note that as the observation noise tends toward zero one observer mode approaches the origin and hence is shielded from the driving noise  $w(\cdot)$  by the system zero at the origin; the other mode tends toward  $\infty$  and results effectively in differentiation of the system output to determine the state.

**Other Factors in Selecting Observer Poles.** It should be noted that non-stochastic factors are also important in the selection of observer dynamics. References [14]–[15] discuss how one can select them to compensate for the effect of deterministic modeling errors, component tolerances, and amplifier biases (see Exercise 4.1-6). Here we may remark that the reduced-order methods of Sec. 4.3 cannot be used when the observed process contains white (broadband) noise, because differentiation would greatly accentuate the noise. Observer designs for colored additive noise have also been studied (see, e.g., [16] and [17]).

**A Stochastic Separation Property.** As in the purely deterministic case studied in Section 4.2, the problem of optimal quadratic regulation when the states and the measurements are corrupted by noise as in (1) also reveals

a so-called *separation property*. That is, the expected cost

$$EJ = E \int_0^{\infty} \{ [c_x x(t)]^2 + r |u(t)|^2 \} dt$$

is minimized by a feedback law

$$u(t) = -\bar{k} \dot{x}(t)$$

where  $\bar{k}$  is determined as in the case of complete state availability, and  $\dot{x}(\cdot)$  is the optimal least-squares estimate of  $x(\cdot)$  as determined by (6). We shall not enter further into this result and the special conditions under which it holds, except to refer to the derivations in, for example, [1, pp. 256-292]. We may remark also that, as mentioned in Sec. 3.4.1, the use of state estimates may cause a deterioration in some of the special robustness (e.g., guaranteed gain and phase margins) properties of the optimal quadratic regulator with completely accessible states (see [18]).

#### Exercises

##### 4.4-1. Variational Approach to the Estimator

It can be shown that the solution of the following optimization problem also leads to the optimal noisy observer. Find  $x(\cdot)$  and  $v_i(\cdot)$  to minimize  $J = \int_0^{\infty} [Ly(t) - cx(t)]r^{-1}[y(t) - c'v_i(t)] + v'_i(t)v_i(t) dt$  subject to the equations  $\dot{x}(t) = Ax(t) + gv_i(t)$ ,  $y(t) = cx(t) + v_2(t)$ . It might be argued that this is a reasonable criterion for the observer problem, whether or not its minimization is equivalent to minimizing the mean-square error  $E[x(t) - \hat{x}(t)][x(t) - \hat{x}(t)]$ ; this equivalence is hard to show a priori, but it can be justified *ex post facto* by verifying that its solution yields the same result as the minimum mean-square-error problem. Therefore, solve this problem by using the (Lagrange multiplier, calculus of variations) method of Sec. 3.4.1 and obtain the solution rule described in Sec. 4.4. Note: This method yields the linear two-point equations  $[x(t_0) \text{ and } \lambda(t_f)]$  are given]

$$\begin{bmatrix} \dot{x}(t) \\ \dot{\lambda}(t) \end{bmatrix} = \begin{bmatrix} A & -gg' \\ -c'r^{-1}c & -A' \end{bmatrix} \begin{bmatrix} x(t) \\ \lambda(t) \end{bmatrix} + \begin{bmatrix} 0 \\ c'r^{-1}y(t) \end{bmatrix}$$

which can be very effectively studied by scattering-theory methods. (See Reference 27 of Chapter 3.)

##### 4.4-2. Optimum Root Loci

Consider the station-keeping satellite of Example 3.3-2, with input disturbances  $w_x$  and  $w_y$ , and  $\xi = 9x + 2y + w_x$ ,  $\dot{y} = -2\xi - 4y + w_y$ . Determine the loci of optimum observer modes when we have measurements of  $y$  corrupted by white noise of intensity  $r$  and

- a.  $w_x = \text{unit intensity white noise}$  and  $w_y = 0$  and
- b.  $w_x = 0$  and  $w_y = \text{unit intensity white noise}$ .

- c. Extend (6) to the case of systems with multiple uncorrelated white noise inputs. Now find the loci when both  $w_x$  and  $w_y$  are (uncorrelated) unit intensity white noise disturbances. Hint: If  $H_x(s)$  and  $H_y(s)$  are the transfer functions from  $w_x$  and  $w_y$ , respectively, to the output  $y$ , show that the observer modes are the left-half-plane roots of  $H_x(s)H_x(-s) + H_y(s)H_y(-s) + r = 0$ .

## 4.5 DIRECT TRANSFER FUNCTION DESIGN PROCEDURES

The combined observer-controller gives a complete solution, at least in theory, to the problem we raised in Sec. 3.1 of using feedback to modify the dynamic behavior of a given system. Let us briefly review the situation. We pointed out in Sec. 3.1 that simple output feedback or even feedback of the output and some derivatives did not give us enough information to carry out our desired objectives in all situations. Moreover, it was not very clear exactly what was to be done and exactly what could be achieved. The fact that the states gave a complete description of the system then led us to examine the possibility of feeding back not the output and its derivatives but the states of a realization of the given system. We showed in Sec. 3.2 that for a *controllable* realization such feedback could, in fact, allow us to relocate the poles (or rather, the natural frequencies) arbitrarily. But, of course, in general the states may not be directly available, and this led us in Sec. 4.1 to the use of asymptotic observers, which can provide asymptotically good estimates of the states of any *observable* realization. The difference between using output feedback (by which term we shall henceforth also imply possible feedback of some derivatives of the output) and using an observer is that in the latter both input and output are fed back (see Fig. 4.2-1, redrawn more schematically as Fig. 4.5-1). However, by some simple block diagram manipulations, Fig. 4.5-1 can be redrawn in the form of several classical output-feedback configurations, which is perhaps why it was never explicitly considered. Of course, the fact that we are starting from the special observer-controller configuration imposes certain constraints on the derived classical configurations. Moreover, we now know that the point is that such block diagram manipulations may impair stability by introducing additional *hidden modes*, which may or may not be stable. [We should stress that block diagram manipulations are not in general equivalent to similarity transformations between different realizations—in particular, block diagram manipulations can change not only the natural frequencies but also the number of states.]

However, we should not abandon the possibilities of transfer function analysis so easily. We know that by watching for cancellations in the *nominal* transfer function we can keep track of hidden modes (cf. Exercise 2.4-7) and therefore have full knowledge of the internal behavior of the system.

can see that

$$\deg n_u(s) < \deg \delta(s) \quad (32b)$$

Therefore, choosing

$$\deg \delta(s) \geq n - 1 \quad (33)$$

will suffice to make  $n_u(s)/\delta(s)$  and  $n_y(s)/\delta(s)$  realizable, i.e., proper rational functions. Now (31b) directly yields Fig. 4.5-5(b). To complete the identification with the results of Sec. 4.5.1, we just have to choose

$$m(s) = \alpha(s) - a(s) \quad (34)$$

With this substitution in (31b) we see that

$$a(s)[n_u(s) + \delta(s)] + b(s)n_y(s) = \alpha(s)\delta(s) \quad (35)$$

which is precisely Eq. (9) of Sec. 4.5.1, thus yielding the promised rederivation of the results of Sec. 4.5.1.

The insight gained in this calculation as to the significance of (23) can be applied to obtain the results of Sec. 4.5.2 and in fact some useful extensions thereof, especially to changing the relative order. However we shall not pursue these extensions here.

#### Exercises

##### 4.5-1.

Given  $H(s) = (s - 1)/s^2$ , find a compensator that will yield a new third-order realization with characteristic polynomial  $(s - 5)(s^2 + s + 1)$ . Is there a unique solution of this problem? If not, give some criteria for choosing among the different possible solutions.

##### 4.5-2. (Franklin)

Consider a system with the transfer function  $1/s(s + 1)$ , which we wish to modify using feedback so as to have a new characteristic polynomial  $s^2 + 4s + 8$ . Show that the steady-state error in the response to a unit ramp is  $1/K_v$ ,  $K_v = 2$ . It can be shown that for a system with closed-loop poles at  $\{-p_i\}$  and closed-loop zeros at  $\{-z_j\}$ ,

$$\frac{1}{K_v} = \sum \frac{1}{p_i} - \sum \frac{1}{z_j}$$

Show that by choosing  $p_1 = 0.1$  and  $-z_2 = -1/10.4$ , we can achieve  $K_v = 10$ . This illustrates one way in which flexibility in zero assignment might be used.

##### 4.5-3.

Given a system with the transfer function  $b(s)/a(s)$  and a desired stable transfer function  $b(s)\delta(s)/p(s)$ , where  $\deg p(s) - \deg \delta(s) \geq n$ , show that we

can obtain an internally stable realization as follows: Use state feedback (or an observer-controller compensator) to change  $b(s)/a(s)$  to  $b(s)/\alpha(s)$ , where  $\alpha(s)$  contains  $n$  roots of  $p(s)$ ; then cascade this with a realization of  $\delta(s)/\epsilon(s)$ , where  $\epsilon(s)\delta(s) = p(s)$ .

#### 4.5-4. Idealized Compensators

In Eqs. (4.5-2) and (4.5-3), suppose we assume that the denominator polynomial  $\delta(s)$  is absent: i.e., assume  $\delta(s) = 1$ . Show that we can achieve a desired transfer function  $H(s) = b(s)/\alpha(s)$ , where  $\alpha(s)$  is a specified  $n$ th-order polynomial, by choosing the coefficients of  $\beta(s)$  and  $\gamma(s)$  so that  $\{\beta', \gamma'\}S = [0 \ \alpha' - a']$ , where  $S$  is the Sylvester matrix of  $\{a(s), b(s)\}$ ,  $\beta'$  is the column vector of the coefficients of  $\beta(s)$ , and similarly for  $a$ ,  $\alpha$ , and  $\gamma$ . Compare with the results of Sec. 4.5.3.

#### 4.5-5. Reconciliation to State-Space Designs

Suppose the original transfer function is realized in controller form. Calculate the feedback gain  $k_c$  required to give a new characteristic polynomial  $\alpha(s)$ , and assume that the states  $x(t)$  are provided by an ideal observer. Calculate  $H_u(s)$  and  $H_y(s)$  for this compensator (cf. Fig. 4.3-1), and show that they coincide with  $\beta(s)$  and  $\gamma(s)$  as found in Exercise 4.5-4.

#### 4.5-6.

Consider the compensator scheme depicted in Fig. 4.3-6. Show that we can find a stable transfer function  $\gamma(s)$  that will make all the modes of the closed-loop system stable if and only if  $H(s)$  has the following *parity interlacing property*: for every RHP zero of  $H(s)$ , there must be an even number (counted according to multiplicity) of poles of  $H(s)$  to the right of it. Reference: D. C. Youla, J. J. Bongiorno, and C. N. Lu, *Automatica*, **10**, pp. 159-174, 1974; see also *Ibid.*, **12**, pp. 387-388, 1976.

#### REFERENCES

1. K. J. ASTROM, *Introduction to Stochastic Control Theory*, Academic Press, New York, 1970.
2. R. E. KALMAN, P. FALB, and M. A. ARIBI, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.
3. D. G. LUESBERGER, "Observing the State of a Linear System," *IEEE Trans. Mil. Electron.*, **MIL-8**, pp. 74-80, 1964. (Also see Ph.D. thesis, Stanford University, Calif., 1963.)
4. B. GORINATH, "On the Control of Linear Multiple Input-Output Systems," *Bell Syst. Tech. J.*, **50**, pp. 1063-1081, March 1971. (Also see Ph.D. thesis, Stanford University, Stanford, Calif., 1968.)
5. D. G. CUMMING, "Design of Observers of Reduced Dynamics," *Electron. Lett.*, **5**, no. 10, pp. 213-214, May 15, 1969.
6. D. G. LUESBERGER, "An Introduction to Observers," *IEEE Trans. Autom. Control*, **AC-16**, pp. 596-603, Dec. 1971.

7. A. E. BRYSON and D. G. LUENBERGER, "The Synthesis of Regulator Logic Using State-Variable Concepts," *Proc. IEEE*, **58**, pp. 1803-1811, Nov. 1970.
8. C. T. CHEN, "A New Look at Transfer Function Design," *Proc. IEEE*, **59**, pp. 1580-1585, Nov. 1971. See also Proc. Natl. Electronics Conf., **25**, pp. 46-51, 1969.
9. C. T. CHEN, *Introduction to Linear System Theory*, Holt, Rinehart and Winston, New York, 1970.
10. R. E. KALMAN and R. S. BUCY, "New Results in Linear Filtering and Prediction Theory," *Trans. ASME Ser. D. J. Basic Eng.*, **83**, pp. 95-107, Dec. 1961.
11. H. KWAKernaak and R. SIVAN, *Linear Optimal Control Systems*, Wiley, New York, 1972.
12. T. KAILATH, *Lectures on Linear Least-Squares Estimation*, CISM Courses and Lectures No. 140, Springer-Verlag, New York, 1978.
13. W. L. CHAN, "Variational Dualities in the Linear Regulator and Estimation Problems," *J. Inst. Math. Appl.*, **18**, pp. 237-248, Oct. 1976.
14. J. D. POWELL, "Mass Center Estimation in Spinning Drag-Free Satellites," *J. Spacecr. Rockets*, **9**, pp. 399-405, 1972. (Also see Ph.D. thesis, Dept. of Aeronautics and Astronautics, Stanford University, Stanford, Calif., May 1970.)
15. F. E. THAU and A. KESTENBAUM, "The Effect of Modeling Errors on Linear State Reconstructors and Regulators," *J. Dyn. Syst. Meas. Control*, pp. 454-459, Dec. 1974.
16. T. KAILATH and R. GEESEY, "An Innovations Approach to Least-Squares Estimation, Part V: Innovations Representations and Recursive Estimation in Colored Noise," *IEEE Trans. Autom. Control*, AC-18, no. 5, pp. 435-453, Oct. 1973.
17. L. M. NOVAK, "Discrete-Time Optimal Stochastic Observers," in *Control and Dynamic Systems*, Vol. 12 (C. T. Leondes, ed.), Academic Press, New York, 1976, pp. 259-311.
- 18a. J. C. DOYLE, "Guaranteed Margins for LQG Regulators," *IEEE Trans. Autom. Control*, AC-23, pp. 756-757, Aug. 1978.
- 18b. J. C. DOYLE and G. STEIN, "Robustness with Observers," *IEEE Trans. Autom. Control*, AC-24, Oct. 1979.
19. R. E. MORTENSEN, "The Determination of Compensation Functions for Linear Feedback Systems To Produce Specified Closed-Loop Poles," *Internal Technical Rep. TR-59-0000-00781*, Space Tech. Laboratories, Los Angeles, Aug. 1959. (See also *IEEE Trans. Autom. Control*, AC-8, p. 386, Oct. 1963.)
20. P. P. SHIPLEY, "A Unified Approach to Synthesis of Linear Systems," *IEEE Trans. Autom. Control*, AC-8, pp. 114-120, April 1963.
21. J. B. PEARSON, "Compensator Design for Dynamic Optimization," *Int. J. Control.*, **9**, pp. 413-482, 1969.
22. F. M. BRASCH and J. B. PEARSON, "Pole Placement Using Dynamic Compensators," *IEEE Trans. Autom. Control*, AC-15, pp. 34-43, 1970.
23. C. D. OLDS, *Continued Fractions*, Random House, New York, 1963.
24. A. YA. KHINCHIN, *Continued Fractions*, Chicago University Press, Chicago, 1964 (Russian ed., 1935).
25. V. KUČERA, *Discrete Linear Control: The Polynomial Equation Approach*, J. Wiley, London, 1979.
26. U. B. DESAI and H. L. WEINERT, "Generalized Control-Estimation Duality and Inverse Projections," *Proceedings 1979 Conference on Information Sciences and Systems*, Johns Hopkins University, Baltimore, Maryland, March 1979.

## ANS312 - Optimização em Sistemas Dinâmicos II

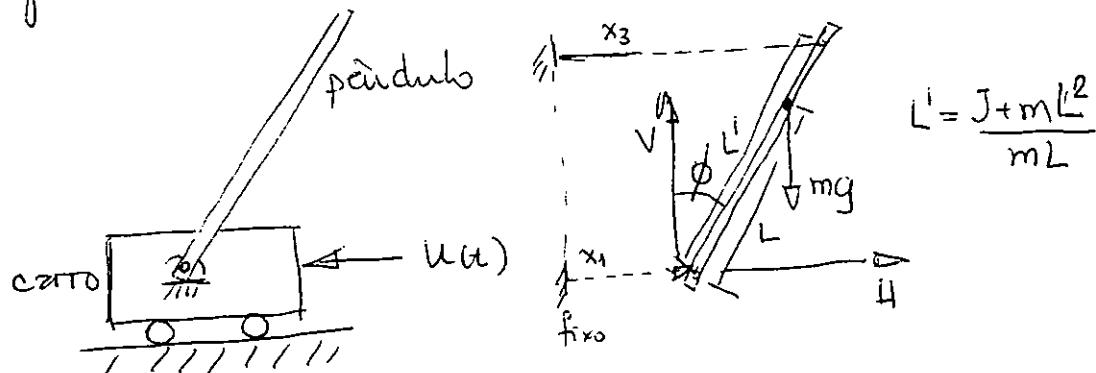
1º Prova - 21/11/88

1º questão] (Fonte: Kwakernaak e Sivan, 1972)

As equações diferenciais do movimento de um pêndulo invertido podem ser colocadas na forma vetorial:

$$\dot{x}(t) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -F/m & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -g/L & 0 & g/L & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1/m \\ 0 \\ 0 \end{pmatrix} u(t)$$

onde  $x(t) = [x_1(t), x_2(t), x_3(t), x_4(t)]^T$  com o significado da figura



$F$  é o coeficiente de atrito carro-solo,  $M$  é a massa do carro e  $u(t)$  é o controle.

Suponha que se tenha a variável observada como sendo o ângulo  $\phi$  entre o pêndulo e a vertical, i.e.,

$$y_1(t) = \left(-\frac{1}{L}, 0, \frac{1}{L}, 0\right) x(t)$$

Considere o problema de se achar um observador invariante no tempo para o sistema (identidade ou de

ordem reduzida):

- Mostre que é impossível achar um observador assintoticamente estável. Explique fisicamente.
- Mostre que se, além do ângulo  $\phi$ , o deslocamento  $s$  do carrinho é também medido, i.e:

$$y_2(t) = (1, 0, 0, 0) \times (t)$$

é adicionado ao vetor de variáveis observadas, então pode-se achar um observador assintoticamente estável.

- Estabeleça um critério e projete o observador de b).

2ª questão] (Fonte: Maybeck, vol I, 1979)

Admita que você é o responsável pela melhoria na ~~precisão~~ precisão de um sistema de rastreamento de posicionamento num laboratório de testes de voo. Atualmente você dispõe de um sistema de rastreamento por radar capaz de medir posições com  $\pm 10\text{m}$ . Você deseja duplicar a precisão para  $\pm 5\text{m}$  aproximadamente. Admitindo custos iguais você poderia:

- combinar de forma ótima os dados do radar atual com os de um novo sistema de radar onde o novo sistema, sozinho, tem capacidade de fornecer posições com  $\pm 6\text{m}$  ou;
- triplicar o sistema original e combinar os dados de maneira ótima.

Qual das propostas a) ou b) você escolheria e porquê? Estabeleça todas as suas hipóteses de modelagem explicitamente.

## CAP. 4 - TEORIA DA PROBABILIDADE

Um número muito grande de fenômenos físicos, de ocorrência bastante comum, não pode ser descrito por los modelos determinísticos a que estamos acostumados. Apenas para citar alguns exemplos: problemas de chamadas telefônicas, de controle de qualidade na fabricação de dispositivos, de falhas de um sistema, de jogos de azar, ou ainda, estudos de Mecânica Estatística, Turbulência ou Acústico, devem ser conduzidos através de modelos não determinísticos. Temos, então, fenômenos aleatórios ou experimentos, onde os resultados de qualquer forma de ocorrência do fenômeno, que chamaremos evento, não podem ser previstos deterministicamente pois são governados pelo acaso.

Por tratar de modo sistemático e formal fenômenos aleatórios como os citados, precisamos:

- a. Modelar matematicamente o problema, de maneira a representá-lo formalmente.
- b. Quantificar a possibilidade de ocorrência dos eventos.
- c. Ser capazes de fazer previsões de forma de ocorrência do fenômeno com base em (a.) e em (b).

Teoria de Probabilidade se encarrega de executar o passo a. no esquema acima, ou seja, a tradução do fenômeno físico em equações matemáticas que descrevem o fenômeno observado. Os passos b. e c. são atributos da Estatística que, fundamentalmente, se encarrega

de traduzir em números, possíveis de interpretação, o problema originalmente proposto.

#### 4.1. AS VÁRIAS DEFINIÇÕES DE PROBABILIDADE

Existem ao menos 3 maneiras sérias de se definir probabilidade: definição axiomática, definição clássica e definição de frequência relativa. Destas, a única que permite encarar probabilidade como uma teoria deductiva é a definição axiomática. No entanto, o formalismo matemático envolvido é tão grande que o seu emprego se torna problemático. Descrevemos sucintamente os conceitos envolvidos nas três definições:

##### 4.1.1. Definição Clássica

Por é montagem de uma teoria deductiva o conceito de definição clássica de probabilidade já foi abandonado há muito tempo, mas, ainda tem grande valor na determinação de dados probabilísticos e como hipótese de trabalho. Historicamente era o conceito utilizado em probabilidade até a década de 30.

De acordo com a definição clássica, a probabilidade  $P(A)$  de ocorrência de um evento A é determinada a priori, sem comprovação experimental. Assim, dado um experimento, conta-se o número total  $N$  de possíveis saídas (eventos). Se em  $N_A$  desses saídas ocorre o evento A, então  $P(A)$  é definida por:

$$P(A) = \frac{N_A}{N},$$

ié,  $\hat{P}(A)$  iguala a razão entre os números de saídas favoráveis e total de saídas.

Esse conceituação apresenta uma série de desvantagens:

(a) Só pode ser usada numa série limitada de problemas. Por exemplo, no experimento de se lançar um dado, ela é aplicável apenas se as seis faces são igualmente prováveis. Se o dado for viciado, é impossível usar a definição.

(b) Em muitos problemas o número possível de saídas é infinito. Para determinar probabilidades de alguns eventos pela definição clássica deve-se introduzir algumas medidas "definidas" como confronto e área, o que, geralmente, causa transtornos.

(c) Existe implicitamente um uso da interpretação de frequência relativa. Por exemplo, no problema do dado  $P(\text{um 6 face}) = \frac{1}{6}$  não se pensa pela simetria mas também porque foi observado ao longo da "história" que  $\frac{n_6}{n} \rightarrow \frac{1}{6}$ .

#### 4.1.2. Definição de Frequências Relativas

A conceituação de frequência relativa foi introduzida por Von Mises, em 1935. Considere a seguinte definição:

Um experimento é repetido  $n$  vezes. Se o evento A ocorre  $n_A$  vezes, então  $\hat{P}(A)$  é definida como o limite das frequências relativas, ié:

$$\hat{P}(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n}$$

Embora essa conceituação pareça bastante fácil e seja fundamental para se atribuir probabilidades a eventos no mundo físico, traz uma série de inconvenientes

do ponto de vista matemático. O maior sério delas é que qualquer número  $n$  de repetições será sempre finito e, com isso, a definição é mais uma hipótese sobre a existência do limite e não uma generalização experimental.

#### 4.1.3. Definição Axiomática

O desenvolvimento da teoria axiomática de probabilidade foi feito por Kolmogorov, sendo que seus primeiros resultados foram publicados em 1933. Partindo de 3 axiomas fundamentais, como veremos no próximo item, consegue-se construir toda uma teoria matemática que elimine os problemas de inconsistência das outras definições. A dificuldade que se encontra com esse formulário é que a teoria matemática envolvida, chamada Teoria da Medida, é extremamente complicada. Como será visto no decorrer desse capítulo, a definição axiomática não se presta a calcular probabilidades de eventos, que podem ser calculadas usando-se interpretações clássicas ou de frequências relativas, mas, uma vez atribuídas essas probabilidades, a definição axiomática checa os valores atribuídos para se poder fazer a aplicação.

#### 4.2. TEORIA AXIOMÁTICA DE PROBABILIDADE.

Dado um fenômeno aleatório (experimento) associamos um conjunto  $\Omega$ , chamado ESPAÇO AMOSTRAL, que represente todos os possíveis resultados (SAÍDAS) básicos ou elementares do fenômeno.

Exemplos:

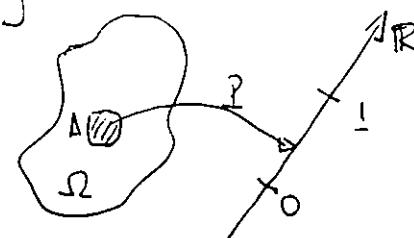
- a) Experimento: Izucar um dado e verificar a face que saiu  $\rightarrow \Omega = \{1, 2, \dots, 6\}$ , espaço amostral
- b) Experimento: leitura de uma tensão num circuito elétrico através de um voltmetro entre 10 e 15V  $\rightarrow \Omega = \{x \in \mathbb{R} \mid 10 \leq x \leq 15\}$ , espaço amostral.

Dado o fenômeno, para cada forma de ocorrência desse fenômeno associamos um subconjunto do espaço  $\Omega$  que chamaremos **EVENTO**.

Exemplos:

- a) Experimento do dado; evento: ocorrer número par equivalente a  $A = \{2, 4, 6\}$ , evento
- b) Experimento do voltmetro; evento: ocorrer leitura entre 11 e 12V equivalente ao evento  $A = \{11 \leq x \leq 12\}$

Para quantificar a possibilidade de ocorrência de um evento, definimos uma **FUNÇÃO**, dita **FUNÇÃO DE PROBABILIDADE** que fornece uma **MEDIDA** dessa possibilidade de ocorrência. A função de probabilidade,  $P$ , será definida para subconjuntos de  $\Omega$  (eventos) com valores em  $[0, 1]$ .



As propriedades convenientes da função de probabilidade  $P$ , som vistas a futuros desenvolvimentos matemáticos, são estabelecidas através dos seguintes **AXIOMAS**:

a)  $P(\Omega) = 1$

b) se  $P(\cdot)$  for definida para um evento  $A \subset \Omega$ , então:  
 $P(A) \geq 0$

c) se  $P(\cdot)$  for definida para  $A_i, A_j$  tais que  $A_i \subset \Omega$ ,  
 $A_j \subset \Omega \subset A_i \cap A_j = \emptyset$ , com  $i \neq j$ ,  $i, j = 1, 2, \dots, n, \dots$   
 (enumerável), então:

$$P(A_1 \cup A_2 \cup \dots \cup A_n \cup \dots) = P(A_1) + P(A_2) + \dots + P(A_n) + \dots$$

Notemos que a função de probabilidade  $P$  este definida sobre conjuntos (subconjuntos de  $\Omega$ , eventos). Para essa função ter um sentido matemático, é preciso especificar qual a CLASSE DE EVENTOS onde  $P$  tem seu domínio. Esse conjunto de eventos está contida num estrutura especial chamada  $\sigma$ -ÁLGEBRA ( $\sigma$ -ALGEBRA), que representa um conjunto de conjuntos ou ainda um campo de conjuntos.

Diz-se que  $\mathcal{B}$  é uma  $\sigma$ -ÁLGEBRA de subconjuntos do espaço amostral  $\Omega$  se:

a)  $\Omega \in \mathcal{B}$

b) se  $A \in \mathcal{B}$ , então seu complemento em relação a  $\Omega$ ,  $\bar{A} = \Omega - A$  é tal que  $\bar{A} \in \mathcal{B}$ .

c) se  $A_1, A_2, \dots, A_n, \dots \in \mathcal{B}$  então todo união enumerável e intersecções enumerável de conjuntos  $A_i$  também pertencem a  $\mathcal{B}$ , isto é:

$$\bigcup_i^{\infty} A_i \in \mathcal{B} \quad \text{e} \quad \bigcap_i^{\infty} A_i \in \mathcal{B}.$$

Tendo o campo de definição de  $P$ , i.e., a  $\sigma$ -álgebra  $\mathcal{B}$ , dizemos que  $P$  é uma função de probabilidade se:

a)  $P(\Omega) = 1$

b)  $P(A) \geq 0$  para todo  $A \in \mathcal{B}$

c) se  $A_i \cap A_j = \emptyset$ ;  $A_i, A_j \in \mathcal{B}$ ;  $i, j = 1, 2, \dots$  então

$$P(A_1 \cup A_2 \cup \dots) = \sum_{i=1}^{\infty} P(A_i)$$

Afirmamos, agora, sem provar, que um experimento físico aleatório fica totalmente caracterizado pelo espaço amostral  $\Omega$ , pela  $\sigma$ -álgebra  $\mathcal{B}$ , e pela função de probabilidade  $P$ , isto é, um experimento  $\mathcal{F} \triangleq (\Omega, \mathcal{B}, P)$

Como observações, deve ser acrescentado que:

a) o espaço amostral  $\Omega$  não é único para um certo experimento  $\mathcal{F}$ ;

b) a  $\sigma$ -álgebra  $\mathcal{B}$  para um experimento  $\mathcal{F}$  também não é única. Por exemplo: experimento  $\mathcal{F}$ : lançamento de um dado

b.1. Eventos de interesse: a face mostrada

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$\mathcal{B} = \{\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \Omega\}$$

b.2. Eventos de interesse: sair nº par ou ímpar

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$\mathcal{B} = \{\emptyset, \{2, 4, 6\}, \{1, 3, 5\}, \Omega\}$$

Alguns exemplos de construção de  $\{\Omega, \mathcal{B}, P\}$ .

1. Experimento de lançar um dado e verificar a face

$$\Omega = \{1, 2, \dots, 6\}$$

$\mathcal{B} = \{\text{classe de todos subconjuntos de } \Omega\}$

$$P(\{i\}) = 1/6, \quad i = 1, 2, \dots, 6$$

2. Experimento de ocorrência de uma chamada telefônica no horário no intervalo  $[0, T]$

$$2.1. \quad \Omega = \{x \in \mathbb{R}, \quad 0 \leq x \leq T\}$$

$$\mathcal{B} = \sigma\text{-álgebra gerada pelos conjuntos } C_E = \{[t_1, t_2], \quad 0 \leq t_1 \leq t_2 \leq T\}$$

$$P([t_1, t_2]) = \frac{t_2 - t_1}{T}$$

2.2.  $\Omega = \{x \in \mathbb{R}, \text{ reta real}\}$

$\mathcal{B} = \mathcal{B}' \cap \sigma\text{-álgebra de Borel na reta}$

$$P([t_1, t_2]) = \frac{\text{comprimento } ([t_1, t_2] \cap [0, T])}{T}$$

No exemplo 2.2. mencionamos  $\mathcal{B}'$ ,  $\sigma$ -álgebra de Borel na reta real. Essa  $\sigma$ -álgebra ou campo de Borel pode ser gerada por uma classe ( $\mathcal{C}_E$ ) de todos os intervalos do tipo ( $\{w \in \mathbb{R}, w \leq x, x \in \mathbb{R} \text{ qualquer}\}$ ). Nesse caso, mostra-se que a  $\sigma$ -álgebra gerada contém todos os tipos de conjuntos de interesse para aplicações da teoria de probabilidade.

#### 4.3. VARIÁVEIS ALEATÓRIAS

Como já mencionamos, com as definições anteriores podemos representar qualquer fenômeno físico se construirmos um experimento  $\mathfrak{F}$  e se representarmos esse fenômeno por  $(\Omega, \mathcal{B}_{\mathfrak{F}})$ . No entanto, essa representação não é conveniente para tratamento quantitativo por obrigar a trabalhar com conjuntos (eventos). Para facilitar a representação matemática, vamos associar os conjuntos mencionados (eventos) a números, definindo o que é chamado conceito de variável aleatória.

Então, dado um experimento  $\mathfrak{F}$ , cujos resultados são eventos  $A$ , subconjuntos do espaço amostral  $\Omega$ , a cada evento  $A$  atribuiremos, de acordo com a mesma regras, um número  $X(A)$ .

Com isso, ficará explícita uma relação entre os elementos (conjuntos)  $A$  do espaço amostral  $\Omega$  e certos números.

A relação, função, assim definida é chamada VARIÁVEL ALEATÓRIA (abreviadamente v.a.). Note que o nome variável aleatória não é dos mais felizes visto que na realidade temos uma função. Antes de definir de maneira rigorosa uma v.a. desenvolvemos dois exemplos:

1) Experimento do dado

Designamos por  $f_1, f_2, \dots, f_6$  cada uma de suas 6 faces e construimos:

$$\underline{X}(f_1) = 10 ; \underline{X}(f_2) = 20 ; \dots ; \underline{X}(f_6) = 60$$

ou

$$\underline{X}(f_i) = 10 i$$

$\underline{X}(f_i)$  como acima é uma variável aleatória.

2) Chamado Telefônico no intervalo  $[0, T]$

Nesse caso  $\Omega = \{t \mid 0 \leq t \leq T\}$

Considere o intervalo  $(t_1, t_2) \subset [0, T]$  e definimos:

$$\underline{X}(t) = \begin{cases} 1 & t_1 \leq t \leq t_2 \\ 0 & \text{fora do intervalo } (t_1, t_2) \end{cases}$$

$\underline{X}(t)$  como acima é uma variável aleatória.

Def.: Dado um experimento  $\mathcal{F} \triangleq (\Omega, \mathcal{B}, \mathbb{P})$  e definido  $X: \Omega \rightarrow \mathbb{R}^+$  função real, finita, diz-se que  $X$  é v.a. escalar se e somente se

$$\{w \in \Omega : X(w) \leq \alpha, \alpha \in \mathbb{R}'\} \in \mathcal{B}$$

De definição acima verificamos, então que dado o experimento  $\mathcal{F} \triangleq (\Omega, \mathcal{B}, \mathbb{P})$ , para cada resultado elementar  $w$  do experimento associamos um número  $\underline{X}(w)$  através de uma

v.z. O ponto importante dessa definição é que os conjuntos de  $\omega$ 's onde se define a v.z. sejam da  $\sigma$ -álgebra  $\mathcal{B}$ , sen o que a definição não é válida.

O valor que essa função toma para um particular  $\omega$  é chamado uma "realização" da v.z.  $X(\cdot)$

Voltamos ao exemplo 1, do experimento do dado. Nesse caso  $X(f_2) = 20$  é uma realização da v.z.  $X(f_1) = 10$ . Por outro lado  $\{X(\omega) \leq 35\} = \{f_1, f_2, f_3\} \in \mathcal{B}$ , com  $X(\omega) \leq 35$  só para  $\omega = f_1$  ou  $f_2$  ou  $f_3$ . Da mesma forma:  $\{X(\omega) \leq 5\} = \{\emptyset\} \in \mathcal{B}$ . e assim por diante.

Construimos mais um exemplo com o experimento:

3) lançamento de uma moeda e verificação de cara ou coroa. Podemos fazer:

$$\Omega = \{C_0, C_1\}$$

$$\mathcal{B} = \{\emptyset, \Omega, \{C_0\}, \{C_1\}\}$$

$$P(C_0) = P(C_1) = 1/2$$

Vamos definir  $X: \Omega \rightarrow \mathbb{R}'$  fazendo, por exemplo:

$$X: \Omega \rightarrow \mathbb{R}' ; \quad X(C_0) = 1 ; \quad X(C_1) = 0$$

e associemos:

$$P(C_0) = P(\{\omega : X(\omega) = 1\}) = P_X(\{X(\omega) = 1\}) = 1/2$$

$$P(C_1) = P(\{\omega : X(\omega) = 0\}) = P_X(\{X(\omega) = 0\}) = 1/2$$

Se  $\alpha \geq 1$ , então  $\{X(\omega) \leq \alpha\}$  é o evento certo já que

$$X(C_0) = 1 \leq \alpha \quad \text{e} \quad X(C_1) = 0 \leq \alpha$$

Então:

$$P(\{\omega : X(\omega) \leq \alpha, \alpha \geq 1\}) = P(\Omega) = 1$$

Verificamos, então, que existe uma relação entre as funções de probabilidade  $P(A)$  e as realizações  $X(\omega)$  da v.z.  $X(\cdot)$ .

Essa relação é dada pela função distribuição de probabilidades.

#### 4.4. FUNÇÃO DISTRIBUIÇÃO DE PROBABILIDADES

Como já vimos no desenvolvimento do item anterior, a definição de v.z.  $X$  leva a que todos os subconjuntos da forma

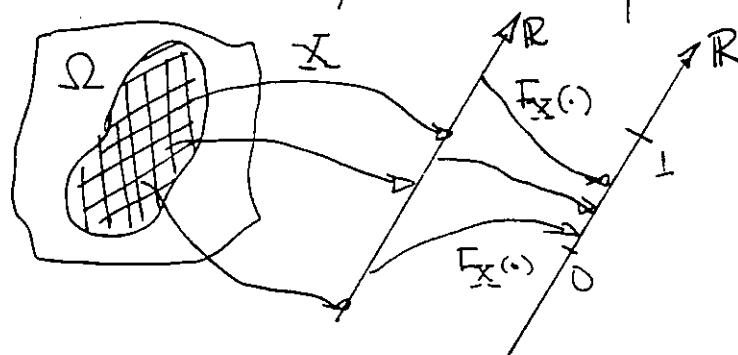
$$A = \{ \omega : X(\omega) \leq z \}$$

têm probabilidades associadas desde que  $A \subset \Omega$  e  $A \in \mathcal{B}$ .

Em consequência, definimos a função ~~DISTRIBUIÇÃO~~ de probabilidade  $F_X(\cdot)$ , função escalar com valores reais através de:

$$F_X(z) \triangleq P(\{\omega : X(\omega) \leq z, z \in \mathbb{R}\})$$

Nstemos então que uma v.z.  $X(\cdot)$  faz um mapeamento do espaço amostral  $\Omega$  no retângulo real e que a função distribuição leva de retângulo real ao intervalo de probabilidades  $[0,1]$ .



Para tentar sintetizar todo o desenvolvimento até agora recordamos ao seguinte exemplo:

4) Novamente o experimento é o lançamento de um dado. Agora, estamos interessados em obter dois eventos:

$$\Delta_1 = \{ saída de um n=1 ou de um n=2\}$$

$$\Delta_2 = \{ saída de um n=3\}$$

O espaço amostral  $\Omega$  é dado por:

$$\Omega = \{1, 2, \dots, 6\}$$

Uma  $\sigma$ -álgebra possível  $\mathcal{B}$  é compacta por  $\emptyset, \Omega, A_1, A_2$  e todas as possíveis uniões, intersecções e complementos dos subconjuntos acima. Nesse caso, a  $\sigma$ -álgebra gerada é dita minimal e contém:

$$\mathcal{B} = \left[ \emptyset; \Omega; \{1 \cup 2\}; \{3\}; \overline{\{1 \cup 2\}} = \{3 \cup 4 \cup 5 \cup 6\}; \overline{\{3\}} = \{1 \cup 2 \cup 4 \cup 5 \cup 6\}; \{1 \cup 2\} \cup \{3\} = \{1 \cup 2 \cup 3\}; \overline{\{1 \cup 2\} \cup \{3\}} = \{4 \cup 5 \cup 6\} \right]$$

Todos os elementos de  $\mathcal{B}$  tem uma probabilidade associada. Nesse caso, a função de probabilidade é estabelecida supondo o dado uniformizado:

$$P(\emptyset) = 0$$

$$P(\Omega) = 1$$

$$P(A_1) = P(\{1 \cup 2\}) = P_1 = \frac{1}{3}$$

$$P(A_2) = P(\{3\}) = P_2 = \frac{1}{6}$$

$$P(\bar{A}_1) = P(\{3 \cup 4 \cup 5 \cup 6\}) = 1 - P_1 = \frac{2}{3}$$

$$P(\bar{A}_2) = P(\{1 \cup 2 \cup 4 \cup 5 \cup 6\}) = 1 - P_2 = \frac{5}{6}$$

$$P(A_1 \cup A_2) = P(\{1 \cup 2 \cup 3\}) = P_1 + P_2 = \frac{1}{2}$$

$$P(\overline{A_1 \cup A_2}) = P(\{4 \cup 5 \cup 6\}) = 1 - P_1 - P_2 = \frac{1}{2}$$

Definimos agora uma v.z.  $X$  associada a esse experimento da forma:

$$X(\omega) = \begin{cases} 0 & \text{se } \omega \notin A_1 \text{ e } \omega \notin A_2 & 3,4,5 \\ 1 & \text{se } \omega \in A_1 & 1,2 \\ 2 & \text{se } \omega \in A_2 & 3 \end{cases}$$

e construimos sua função distribuição de probabilidades:

$$F_X(x) = P(\{\omega : X(\omega) \leq x\})$$

$$\text{Para } x < 0 \Rightarrow P(\{\omega : X(\omega) < 0\}) = P(\emptyset) = 0$$

Para  $0 \leq z < 1 \Rightarrow$

$$P(\{\omega : X(\omega) \leq z, z \in [0,1]\}) = P(\overline{A_1 \cup A_2}) = P(X(\omega) = 0) = \frac{1}{2}$$

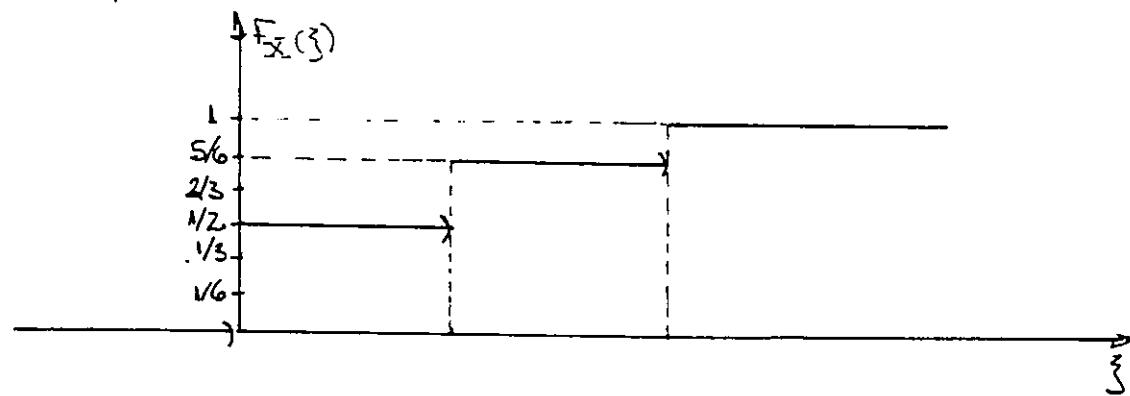
Para  $1 \leq z < 2 \Rightarrow$

$$\begin{aligned} P(\{\omega : X(\omega) \leq z, z \in [1,2]\}) &= P(\overline{A_2}) = P(X(\omega) = 0 \text{ ou } X(\omega) = 1) \\ &= P(\{1 \cup 2 \cup 4 \cup 5 \cup 6\}) = \frac{5}{6} \end{aligned}$$

Para  $2 \leq z < \infty \Rightarrow$

$$P(\{\omega : X(\omega) \leq z, z \in [2,+\infty)\}) = P(\Omega) = 1$$

O gráfico abaixo resume o comportamento da função distribuição nesse caso:



Propriedades da função distribuição:

I)  $F_X(\cdot)$  é monótona não decrescente, isto é:

$$F_X(z_2) \geq F_X(z_1) \quad \text{se } z_2 > z_1$$

II)  $\lim_{z \rightarrow -\infty} F_X(z) = 0$  ;  $\lim_{z \rightarrow +\infty} F_X(z) = 1$

III)  $P(\{X(\omega) = z\}) = F_X(z) - F_X(z^-)$

$$\text{Prov: } \{ \omega : X(\omega) \leq z \} = \{ \omega : X(\omega) = z \} \cup \{ \omega : X(\omega) < z \}$$

$$\text{Através de um dos axiomas: } P(\{X(\omega) \leq z\}) = P(\{X(\omega) = z\}) + P(\{X(\omega) < z\})$$

$$F_X(z) = P(\{X(\omega) = z\}) + F_X(z^-)$$

$$P(\{X(\omega) = z\}) = F_X(z) - F_X(z^-)$$

Def.: Uma v.z.  $X(\cdot)$  é chamada v.z. discreta se existir uma função distribuição tali que:

$$F_X(z) = \sum_{\substack{\epsilon_i < z \\ i}} P(\{X(\omega) = \epsilon_i\}) = \sum_i m_X(\epsilon_i)$$

Nesse caso, a função distribuição  $F_X(z)$  tem a forma de uma escada com descontinuidades nos pontos  $\epsilon_i$  e  $m_X(\epsilon_i)$  tem o nome de função de massa da v.z.  $X(\cdot)$ .

Propriedades das v.z. discretas:

I) Para  $\epsilon$  tal que  $X(\omega) = \epsilon$

$$F_X(\epsilon) - F_X(\epsilon^-) = P(\{X(\omega) = \epsilon\}) = m_X(\epsilon)$$

II)  $\sum_i m_X(\epsilon_i) = 1$  para todos os  $i$

$$F_X(\infty) - F_X(-\infty) = 1$$

Exemplo: Voltemos ao exemplo do lançamento de uma moeda (3, pg. 4.10). Tínhamos:

$$\Omega = \{C_0, C_1\}$$

$$\mathcal{B} = \{\emptyset, \Omega, C_0, C_1\}$$

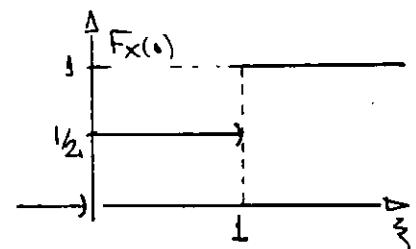
$$P(C_0) = P(C_1) = \frac{1}{2},$$

e definimos uma v.a. discreta

$$\bar{X}(\omega) = 1 \quad ; \quad \bar{X}(\omega) = 0$$

A função distribuição será:

$$F_{\bar{X}}(\xi) = \begin{cases} 0 & \xi < 0 \\ \frac{1}{2} & 0 \leq \xi < 1 \\ 1 & \xi \geq 1 \end{cases}$$



A correspondente função de massa é:

$$m_{\bar{X}}(\varepsilon) = \begin{cases} \frac{1}{2} & \varepsilon = 0 \\ \frac{1}{2} & \varepsilon = 1 \\ 0 & \varepsilon \neq 0,1 \end{cases}$$

Def. Uma variável aleatória é dita contínua se existir uma função densidade de probabilidades  $p_{\bar{X}}(\cdot)$  tal que:

$$F_{\bar{X}}(\xi) = \int_{-\infty}^{\xi} p_{\bar{X}}(s) ds$$

Propriedades das v.a. contínuas:

I) Se  $F_{\bar{X}}(\xi)$  é contínuo, então  $P(\{\bar{X}(\omega) = \xi\}) = 0$

II) Se  $F_{\bar{X}}(\cdot)$  for diferenciável  $p_{\bar{X}}(\cdot) = \frac{d}{d\xi} F_{\bar{X}}(\cdot)$

III)  $p_{\bar{X}}(\cdot) \geq 0$

IV)  $\int_{-\infty}^{\infty} p_{\bar{X}}(\xi) d\xi = 1$

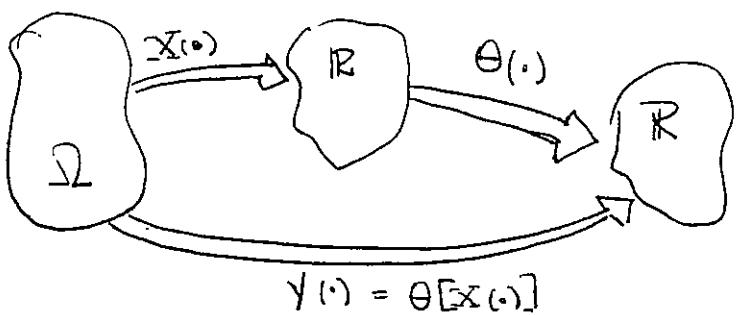
V) para um  $\Delta \xi$  suficientemente pequeno  
 $P(\{\xi \in \bar{X}(\omega) \leq \xi + \Delta \xi\}) = p_{\bar{X}}(\xi) \Delta \xi$

Dois observações importantes devem ser feitas e entenda. A primeira se refere à existência da função distribuição de probabilidades. Essa função, como está ligada diretamente à probabilidade, sempre existe. A função densidade de probabilidades pode existir ou não. A segunda observação é que a função distribuição como definida acima completa a transformação do experimento  $\mathcal{I} \triangleq (\Omega, \mathcal{B}, P)$ , um espaço abstrato, para um espaço mais físico  $(R, \mathcal{B}_R, f_X)$  pela aplicação de uma conveniente v.a.  $X(\cdot)$ . Esta transformação tornar-se-á mais clara adiante.

Conceituamos agora o que seja uma função de variável aleatória:

$$Y(\cdot) = \Theta[X(\cdot)]$$

é uma v.a. (função) com valores reais obtida por uma aplicação de  $\Theta$  sobre a v.a.  $X(\cdot)$ . O esquema abaixo indica a sequência de aplicações:



Este conceito de função de v.a. é extremamente importante para a introdução de conceitos de médias e momentos de v.a.'s que são funções de v.a's.

Def.: Média ou Esperança de v.a.  $X(\cdot)$  contínua, quando existe, é dada por:

$$E[X] = \int_{-\infty}^{+\infty} z p_X(z) dz \triangleq \mu_X$$

$E[\cdot]$  é um operador linear no sentido de que:

$$E[zx_1 + b\bar{x}_2] = zE[\bar{x}_1] + bE[\bar{x}_2]$$

A relação acima é facilmente provada a partir da definição de esperança. O valor médio tem sua analogia mecânica no báculento ou centro de massa da distribuição de probabilidades.

Def.: Momento de ordem  $n$  de uma v.z.  $\bar{X}(\cdot)$  é dada por:

$$E[\bar{x}^n] = \int_{-\infty}^{\infty} z^n p_{\bar{x}}(z) dz$$

Nessa definição geral se encaram diversos casos de interesse como a média ( $n=1$ ) e o valor médio quadrático ( $n=2$ ) definido por:

$$E[\bar{x}^2] = \int_{-\infty}^{\infty} z^2 p_{\bar{x}}(z) dz$$

Def.: Momento central de ordem  $n$  de uma v.z.  $\bar{X}(\cdot)$  é dada por:

$$E[(\bar{x} - E[\bar{x}])^n] = \int_{-\infty}^{\infty} (\bar{z} - E[\bar{x}])^n p_{\bar{x}}(z) dz$$

Dentro desse definição geral o caso de maior importância ocorre para  $n=2$  quando obtemos a variança de v.z.  $\bar{X}(\cdot)$ .

$$E[(\bar{x} - E[\bar{x}])^2] = E[(\bar{x} - \mu_{\bar{x}})^2] = \int_{-\infty}^{\infty} (\bar{z} - \mu_{\bar{x}})^2 p_{\bar{x}}(z) dz = \sigma_x^2$$

A raiz quadrada da variança, dize-se o nome de  $\sigma_x$ , desvio padrão da v.z.  $\bar{X}$ .

Das definições acima conclui-se facilmente que:

$$\begin{aligned} E[(\bar{x} - \mu_{\bar{x}})^2] &= E[(\bar{x}^2 - 2\mu_{\bar{x}}\bar{x} + \mu_{\bar{x}}^2)] = \\ &= E[\bar{x}^2] - E[2\mu_{\bar{x}}\bar{x}] + E[\mu_{\bar{x}}^2] \end{aligned}$$

$$E[(\bar{X} - \mu_{\bar{X}})^2] = E[\bar{X}^2] - 2\mu_{\bar{X}} \cdot E[\bar{X}] + \mu_{\bar{X}}^2$$

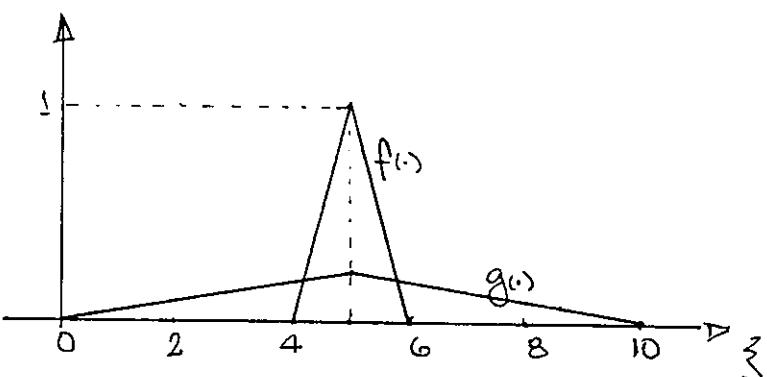
$$E[(\bar{X} - \mu_{\bar{X}})^2] = E[\bar{X}^2] - \mu_{\bar{X}}^2$$

Assim como a média de uma distribuição é uma indicação do centro de variação da v.z., análoga ao centro de massa, a variância de uma v.z. indica a dispersão da função distribuição em relação à média. A analogia mecânica da variação é feita com o momento de inércia de um corpo. ~~Por exemplo, se~~

Deseja-se notar, no entanto, que apenas média e variância não caracterizam completamente uma função distribuição ou função densidade, com algumas exceções, como distribuições gaussianas.

Por se tentar dar um significado mais claro sobre a caracterização da dispersão pela variância examinemos duas distribuições triangulares de suas médias:

$$f(\xi) = \begin{cases} 0 & ; \xi < 4 \\ \frac{\xi-4}{2} & ; 4 \leq \xi \leq 5 \\ -\frac{1}{2}(\xi-6) & ; 5 \leq \xi \leq 6 \\ 0 & ; \xi > 6 \end{cases} \quad \text{e} \quad g(\xi) = \begin{cases} 0 & ; \xi < 0 \\ \frac{3}{2}\xi & ; 0 \leq \xi \leq 5 \\ -\frac{1}{2}(\xi-10) & ; 5 < \xi \leq 10 \\ 0 & ; \xi > 10 \end{cases}$$



Observa-se do gráfico que  $f(\cdot)$  é bastante concentrada em torno da média enquanto  $g(\cdot)$  é bastante dispersa. Esse ideia é traduzida pelas variações que, calculadas através

As definições, nos fornecem:

$$\sigma_f^2 = \frac{1}{6} \quad ; \quad \sigma_g^2 = 25/6$$

Uma v.z. pode também ser especificada em termos de sua função característica.

Def.: A função característica da v.z.  $\bar{X}$ ,  $\bar{X}: \Omega \rightarrow \mathbb{R}$  é dada por:

$$\varphi_{\bar{X}}(u) \triangleq E[e^{iu\bar{X}}] \quad ; \quad i^2 = -1 \text{ (complexo)}$$

A importância da função característica reside no fato de que todos os momentos podem ser obtidos por diferenciação. Mostrase que:

$$E[\bar{X}^n] = \left(\frac{1}{i^n}\right) \frac{d^n}{du^n} \varphi_{\bar{X}}(u=0)$$

para todo  $n$ , desde  $\varphi_{\bar{X}}(u)$  seja uma função analítica. Não se garante que a função característica seja analítica, mas ela sempre existe.

Na definição verificamos que a função característica é a transformada de Fourier da função densidade, i.e.:

$$\varphi_{\bar{X}}(u) = \int_{-\infty}^{+\infty} e^{iuz} p_{\bar{X}}(z) dz$$

Por sua vez a função densidade de probabilidade é a transformada inversa de Fourier da função característica:

$$p_{\bar{X}}(z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuz} \varphi_{\bar{X}}(u) du$$

A função densidade ou a função distribuição ou ainda a função característica determinam completamente uma v.z. Com isto, e juntando de trabalharmos com  $(\Omega, \mathcal{F}, P)$ , usaremos uma um espaço transformado  $(\mathbb{R}^n, \mathcal{B}_F, p_{\bar{X}} \text{ ou } F_{\bar{X}} \text{ ou } \varphi_{\bar{X}})$  através da defini-

é conveniente da v.z. e de suas funções acima.

#### 4.5. V. ALEATÓRIA NORMALMENTE DISTRIBUIDA OU GAUSSIANA.

Uma v.a. é ditz normalmente distribuída ou gaussiana se sua função densidade de probabilidade é da forma:

$$f_X(z) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

onde  $\mu$  e  $\sigma$  são dois parâmetros constantes. É fácil verificar que

$$\mu_X = E[X] = \mu$$

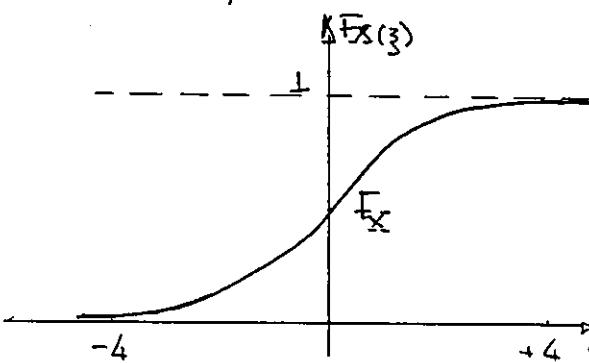
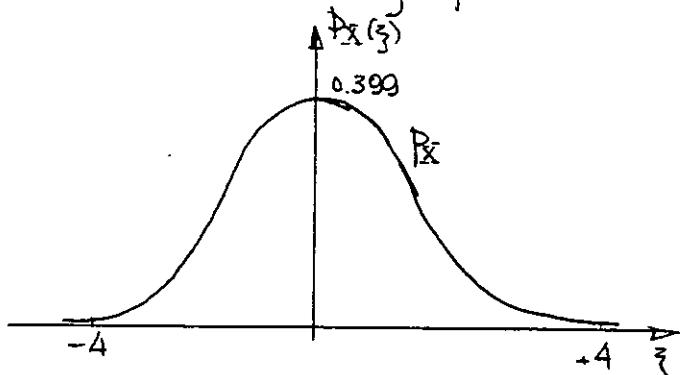
$$\sigma^2 = \text{Var}[X] = E[(X-\mu)^2]$$

e designaremos essa v.z. por  $X \sim N(\mu, \sigma^2)$ , i.e., uma v.z. de distribuição normal de média  $\mu$  e variância  $\sigma^2$ .

Para esse tipo de função densidade a função distribuição será:

$$F_X(z) = \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^z e^{-\frac{1}{2}\left(\frac{\eta-\mu}{\sigma}\right)^2} d\eta$$

Os gráficos abaixo ilustram a forma dessa v.z.:



Distribuição e densidade para  $N(0,1)$

Tomemos uma v.o.  $X$  gaussiana de esperança  $m$  e variancia  $\sigma^2$  e determinemos sua função característica:

$$\Phi_X(u) = \int_{-\infty}^{\infty} e^{iux} \varphi_X(z) dz = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{iu z - \frac{1}{2}(\frac{z-m}{\sigma})^2} dz.$$

Fazemos uma transformação de coordenadas  $y = \frac{(z-m)}{\sigma}$

$$\text{Então: } z = \sigma y + m \Rightarrow dz = \sigma dy$$

$$\Phi_X(u) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{iuy} e^{ium} e^{-\frac{1}{2}y^2} \sigma dy$$

$$\varphi_X(u) = \frac{1}{\sqrt{2\pi}} e^{ium} \int_{-\infty}^{\infty} e^{iuy - \frac{1}{2}y^2} dy$$

Verifiquemos o expoente do integrando. Como:

$$(y - iu\sigma)^2 = y^2 - 2iyu\sigma + i^2 u^2 \sigma^2, \quad \text{então}$$

$$iuy - \frac{1}{2}y^2 = -\frac{1}{2}[(y - iu\sigma)^2 + u^2 \sigma^2]$$

$$\varphi_X(u) = \frac{1}{\sqrt{2\pi}} e^{ium} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y - iu\sigma)^2} \cdot e^{-\frac{1}{2}u^2\sigma^2} dy$$

$$\varphi_X(u) = e^{ium} \cdot e^{-\frac{1}{2}u^2\sigma^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y - iu\sigma)^2} dy$$

Verifiquemos o último termo. Segundo Spiegel (Mathematical Handbook, Schurz):

$$\int_{-\infty}^{+\infty} e^{-(az^2 + bz + c)} dz = \sqrt{\frac{\pi}{a}} e^{\frac{(b^2 - 4ac)}{4a}}$$

e comparando com  $\int_{-\infty}^{\infty} e^{-\frac{1}{2}(y - iu\sigma)^2} dy = \int_{-\infty}^{\infty} e^{-(\frac{1}{2}y^2 - iuy - \frac{1}{2}u^2\sigma^2)} dy$ , temos:

$$a = \frac{1}{2} \quad ; \quad b = -iu\sigma \quad ; \quad c = -\frac{1}{2}u^2\sigma^2$$

e, portanto:

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(y-iu\sigma)^2} dy &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\pi}{1/2}} e^{(-u^2\sigma^2 + 4 \cdot \frac{1}{2} \cdot \frac{1}{2}u^2\sigma^2)/\frac{1}{2}} \\ &= \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} \cdot e^0 = 1 \end{aligned}$$

Com isso, conchusse que:

$$\phi_{\bar{X}}(u) = e^{[iu\bar{m} - \frac{1}{2}u^2\sigma^2]}$$

é a função característica de uma v.a.  $\bar{X} \sim N(\bar{m}, \sigma^2)$

Vamos verificar agora como ficam os momentos centrais de uma v.a. gaussiana, i.e., os momentos calculados em torno de média  $m$ . Para isso, calcularmos a função característica em torno de  $m$ , para obter:

$$\phi_{\bar{X}-m}(u) = E[e^{iu(\bar{z}-m)}] = \int_{-\infty}^{\infty} e^{iu(\bar{z}-m)} p_{\bar{X}}(\bar{z}) d\bar{z}$$

$$\phi_{\bar{X}-m}(u) = \int_{-\infty}^{\infty} e^{iu\bar{z}} e^{-ium} p_{\bar{X}}(\bar{z}) d\bar{z} = e^{-ium} \int_{-\infty}^{\infty} e^{iu\bar{z}} p_{\bar{X}}(\bar{z}) d\bar{z}$$

$$\phi_{\bar{X}-m}(u) = e^{-ium} \cdot \phi_{\bar{X}}(u) = e^{-ium} \cdot e^{[iu\bar{m} - \frac{1}{2}u^2\sigma^2]}$$

$$\phi_{\bar{X}-m}(u) = e^{-\frac{1}{2}u^2\sigma^2}$$

As derivadas da função característica são calculadas a seguir:

$$\phi_{\bar{X}-m}^{(0)}(u) = \phi_{\bar{X}-m}(u)$$

$$\phi_{\bar{X}-m}^{(1)}(u) = \frac{d}{du} (\phi_{\bar{X}-m}^{(0)}(u)) = \frac{d}{du} (e^{-\frac{1}{2}u^2\sigma^2})$$

$$\phi_{\bar{X}-m}^{(1)}(u) = -u\sigma^2 e^{-\frac{1}{2}u^2\sigma^2} = -u\sigma^2 \phi_{\bar{X}-m}^{(0)}(u)$$

$$\phi_{x-m}^{(2)}(u) = -\sigma^2 \phi_{x-m}^{(0)}(u) - u\sigma^2 (-u\sigma^2 e^{-\frac{1}{2}u^2\sigma^2})$$

$$\phi_{x-m}^{(2)}(u) = -u\sigma^2 \phi_{x-m}^{(1)}(u) - \sigma^2 \phi_{x-m}^{(0)}(u)$$

é mostrado, por indução, que:

$$\phi_{x-m}^{(n+2)}(u) = -(n+1)\sigma^2 \phi_{x-m}^{(n)}(u) - u\sigma^2 \phi_{x-m}^{(n+1)}(u)$$

$n=0, 1, \dots$

Avançando este último relação em  $u=0$ :

$$\phi_{x-m}^{(n+2)}(0) = -(n+1)\sigma^2 \phi_{x-m}^{(n)}(0) \quad (*)$$

já que o 2º termo se anula para  $u=0$ .

Notemos também que outras relações obtidas:

$$\phi_{x-m}^{(1)}(u=0) = 0$$

o que era esperado já que esse termo representa o momento em torno de média. Verifiquemos também que:

$$\phi_{x-m}^{(0)}(u=0) = \phi_{x-m}^{(0)}(u>0) = e^{-\frac{1}{2}\sigma^2 u^2} = 1$$

Os valores  $\phi_{x-m}^{(0)}$  e  $\phi_{x-m}^{(1)}$  são suficientes para gerar todos os momentos a partir das relações de recorrência (\*). Nesse caso, teremos:

- como  $\phi_{x-m}^{(1)}(u=0) = 0$ , segue-se por (\*) que:

$$\phi_{x-m}^{(n)}(0) = 0 \quad \text{para todo } n \text{ ímpar, } n \geq 1$$

- como  $\phi_{x-m}^{(0)}(0) = 1$ , segue-se por (\*) que:

$$\phi_{x-m}^{(n)}(0) = (i^n) 1.3.5. \dots (n-1) \sigma^n \quad \text{todo } n \text{ par, } n \geq 2.$$

Verifica-se, então, que a função característica tem sinais alternados. Portanto finalmente, como os momentos centrais são

dadas por:  $E[(X-m)^n] = \left(\frac{1}{i^n}\right) \frac{d^n}{du^n} \Phi_{X-m}(u=0)$

concluimos que uma distribuição gaussiana é tal que:

$$E[(X-m)^n] = \begin{cases} 0 & \text{todo ímpar } n \geq 1 \\ 1.35\dots(n-1)\sigma^n & \text{todo par } n \geq 2 \end{cases}$$

#### 4.6. VARIÁVEIS ALEATÓRIAS CONJUNTAMENTE DISTRIBUIDAS

Vamos agora estudar os conceitos para o caso em que temos mais de uma v.a. Consideremos um experimento  $\mathfrak{F} = (\Omega, \mathcal{P}, \mathbb{P})$  onde estes são definidas v.a. contínuas

$$X_i : \Omega \rightarrow \mathbb{R}^1 \quad , \quad i = 1, 2, \dots, n$$

Montamos o vetor de v.a.:

$$\underline{X} = [X_1, X_2, \dots, X_n]^T$$

As. v.a.  $X_1, X_2, \dots, X_n$  são ditas conjuntamente distribuídas se são definidas no mesmo experimento. Isso ressalva é muito importante: só tem sentido um vetor de v.a. quando todas as componentes são definidas no mesmo experimento ou seja, todas têm o mesmo  $(\Omega, \mathcal{P}, \mathbb{P})$ .

Isto posto, podemos definir a função distribuição conjunta do vetor aleatório  $\underline{X}$  pelo que se segue:

Def.: A função distribuição conjunta do vetor de v.a.  $\underline{X}$  é dado por

$$F_{\underline{X}}(\underline{\xi}) \triangleq \mathbb{P}(\{X_1(\omega) \leq \xi_1, X_2(\omega) \leq \xi_2, \dots, X_n(\omega) \leq \xi_n\})$$

O conjunto  $A = \{X_1(\omega) \leq \xi_1, \dots, X_n(\omega) \leq \xi_n\}$  é um evento do experimento  $\mathfrak{F}$ , i.e.,  $A \subset \Omega$  e  $A \in \mathcal{P}$ . O conjunto  $A$  pode também ser escrito na forma:

$$A = \bigcap_{i=1}^n \{\omega : X_i(\omega) \leq \xi_i\} = \{X_1(\omega) \leq \xi_1\} \cap \dots \cap \{X_n(\omega) \leq \xi_n\}$$

O vetor  $\vec{z} = [\vec{z}_1, \vec{z}_2, \dots, \vec{z}_n]^T$  é dito vetor realizatione do vetor aleatório  $\vec{X}(w)$ .

Podemos definir também uma função densidade conjunta para o vetor aleatório  $\vec{X}(w)$  como segue:

Def.: A função densidade conjunta  $p_{\vec{X}}(\vec{z}) \triangleq p_{\vec{X}_1, \dots, \vec{X}_n}(\vec{z}_1, \dots, \vec{z}_n)$ , quando existe, é dada por:

$$\begin{aligned} F_{\vec{X}}(\vec{z}) &= F_{\vec{X}_1, \dots, \vec{X}_n}(\vec{z}_1, \dots, \vec{z}_n) = \int_{-\infty}^{\vec{z}_1} \int_{-\infty}^{\vec{z}_2} \dots \int_{-\infty}^{\vec{z}_n} p_{\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n}(e_1, e_2, \dots, e_n) de_1 de_2 \dots de_n \\ &= \int_{-\infty}^{\vec{z}_1} \int_{-\infty}^{\vec{z}_2} \dots \int_{-\infty}^{\vec{z}_n} p_{\vec{X}}(e) de_1 de_2 \dots de_n \end{aligned}$$

Decorre, então, que se existir a derivada:

$$p_{\vec{X}}(\vec{z}) = p_{\vec{X}_1, \dots, \vec{X}_n}(\vec{z}_1, \dots, \vec{z}_n) = \frac{\partial^n}{\partial \vec{z}_1 \dots \partial \vec{z}_n} F_{\vec{X}_1, \dots, \vec{X}_n}(\vec{z}_1, \dots, \vec{z}_n)$$

Propriedades das funções distribuição e densidade:

$$\text{I)} \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_{\vec{X}}(\vec{z}) d\vec{z} = 1 \quad \text{onde } p_{\vec{X}}(\vec{z}) \geq 0$$

$$\text{II)} \quad F_{\vec{X}}(\vec{z}_1, \dots, -\infty, \vec{z}_{i+1}, \dots, \vec{z}_n) = \int_{-\infty}^{\vec{z}_1} \dots \int_{-\infty}^{\vec{z}_i} \int_{-\infty}^{\vec{z}_{i+1}} \dots \int_{-\infty}^{\vec{z}_n} p_{\vec{X}}(\vec{z}) d\vec{z} = 0$$

Valem também as outras propriedades enunciadas anteriormente desde que convenientemente estendidas para um vetor de variáveis aleatórias.

Um conceito muito importante para v.z. conjuntamente distribuídas é o de distribuição marginal. Vamos imaginar que estejamos interessados apenas em algumas das v.z.  $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n$ , mais especificamente nas  $m$  primeiras,  $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_m$ , com  $m > n$ .

Nesse caso não nos interessa qual a realização correspondente às v.z.  $X_{m+1}, \dots, X_n$  e definimos distribuição marginal através de:

$$F_{X_1, X_2, \dots, X_m}(\xi_1, \dots, \xi_m) = f_{\underline{X}}(\xi_1, \xi_2, \dots, \xi_m, \infty, \infty, \dots \infty) = \\ = P\left(\bigcap_{i=1}^m \{\omega : X_i(\omega) \leq \xi_i\}\right) \cap \left\{\bigcap_{i=m+1}^n \{\omega : X_i(\omega) \leq \infty\}\right\}$$

$$F_{\underline{X}}(\xi_1, \dots, \xi_m) = \int_{-\infty}^{\xi_1} \dots \int_{-\infty}^{\xi_m} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_{\underline{X}}(e) de_1 \dots de_m de_{m+1} \dots de_n \\ = \int_{-\infty}^{\xi_1} \dots \int_{-\infty}^{\xi_m} p_{X_1, \dots, X_m}(e_1, \dots, e_m) de_1 \dots de_m$$

Podemos também definir a função densidade marginal através de:

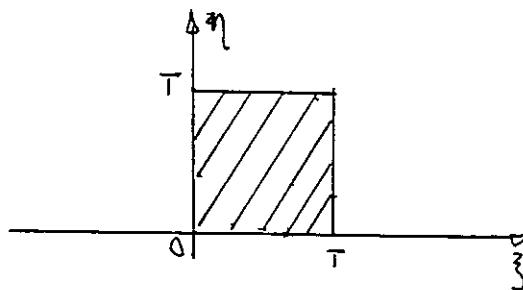
$$p_{X_1, \dots, X_m}(\xi_1, \dots, \xi_m) \triangleq \underbrace{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty}}_{(n-m) \text{ vozes}} p_{X_1, \dots, X_n}(e_1, \dots, e_n) de_{m+1} \dots de_n$$

Exemplo:

Sejam  $X$  e  $\bar{Y}$  duas v.z. conjuntamente distribuídas num retângulo  $[0, T] \times [0, T]$ . Admitzamos que a distribuição seja uniforme dentro do retângulo, i.e., podemos escrever a função densidade conjunta

$$p_{X, \bar{Y}}(\xi, \eta) = \begin{cases} 1/T^2 & \text{no retângulo } [0, T] \times [0, T] \\ 0 & \text{toda outra parte} \end{cases}$$

Podemos associar esse experimento com o de um lançamento de um dardo sobre um alvo quadrado. Os pontos inteiros no retângulo (quadrado) são os pontos de impacto. Os eventos de interesse, nesse caso, seriam subconjuntos do plano  $\xi - \eta$ .



As densidades marginais são facilmente computadas:

$$p_x(\bar{z}) = \int_{-\infty}^{\infty} p_{xz}(\bar{z}, \eta) d\eta = \int_0^T \left( \frac{1}{T^2} \right) d\eta = \frac{1}{T}, \quad \bar{z} \in [0, T]$$

$$p_x(\bar{z}) = 0 \quad , \quad \bar{z} \notin [0, T]$$

$$p_z(\eta) = \int_{-\infty}^{\infty} p_{xz}(\bar{z}, \eta) d\bar{z} = \int_0^T \left( \frac{1}{T^2} \right) d\bar{z} = \frac{1}{T}, \quad \eta \in [0, T]$$

$$p_z(\eta) = 0 \quad , \quad \eta \notin [0, T]$$

Note que nesse exemplo

$$p_{xz}(\bar{z}, \eta) = p_x(\bar{z}) p_z(\eta) = \frac{1}{T} \cdot \frac{1}{T} = \frac{1}{T^2}$$

Para um vetor de variáveis aleatórias conjuntamente distribuídas podemos definir, de maneira análoga ao caso de uma única v.z., os momentos, centrais ou não, de qualquer ordem.

Def.: O momento de ordem  $m$  da v.z.  $\underline{X}_k$  é dado por

$$E[\underline{X}_k^m] \triangleq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \bar{z}_k^m p_{xz}(\bar{z}_1, \dots, \bar{z}_k, \dots, \bar{z}_n) d\bar{z}_1 \cdots d\bar{z}_k \cdots d\bar{z}_n$$

Se  $m=1$ , temos o valor médio ou esperança da v.z.  $\bar{X}_k$ , componente do vetor aleatório  $\underline{X}$ :

$$E[\bar{X}_k] = \mu_k = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \bar{z}_k p_{xz}(\bar{z}_1, \dots, \bar{z}_k, \dots, \bar{z}_n) d\bar{z}_1 \cdots d\bar{z}_k \cdots d\bar{z}_n$$

Os momentos centrais de ordem  $m$  de uma v.z.  $\bar{X}_k$  do vetor  $\underline{X}$  são definidos de maneira semelhante. A forma geral do

momento central de ordem  $m$  é dado por:

$$E[(\bar{X}_k - \mu_k)^m] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\bar{z}_k - \mu_k)^m p_{\bar{X}}(\bar{z}_1, \dots, \bar{z}_k, \dots, \bar{z}_n) d\bar{z}_1 \dots d\bar{z}_n$$

Para  $m=2$  a relação acima temos a variação de  $\bar{X}_k$ :

$$\sigma_k^2 = E[(\bar{X}_k - \mu_k)^2] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\bar{z}_k - \mu_k)^2 p_{\bar{X}}(\bar{z}_1, \dots, \bar{z}_k, \dots, \bar{z}_n) d\bar{z}_1 \dots d\bar{z}_n$$

Para um vetor de v.z.  $\bar{X}$ , os momentos conjuntos de duas v.z.  $\bar{X}_k, \bar{X}_j$  se revestem de importância fundamental.

Def.: Momento conjunto de ordem  $m+l$  de duas v.z.  $\bar{X}_k$  e  $\bar{X}_j$  do mesmo vetor  $\bar{X}$  é dado por:

$$E[\bar{X}_k^m \bar{X}_j^l] \triangleq \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \bar{z}_k^m \bar{z}_j^l p_{\bar{X}}(\bar{z}_1, \dots, \bar{z}_k, \dots, \bar{z}_j, \dots, \bar{z}_n) d\bar{z}_1 \dots d\bar{z}_n$$

A relação acima pode ser colocada em função de densidade marginal de  $\bar{X}_k$  e  $\bar{X}_j$  e escrita como:

$$E[\bar{X}_k^m \bar{X}_j^l] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{z}_k^m \bar{z}_j^l p_{\bar{X}, \bar{X}_j}(\bar{z}_k, \bar{z}_j) d\bar{z}_k d\bar{z}_j$$

Note que para o caso de  $m=l=1$  e  $k=j$  a fórmula acima recaí na definição do valor médio quadrático da v.z.  $\bar{X}_k$ :

$$E[\bar{X}_k^2] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \bar{z}_k^2 p_{\bar{X}}(\bar{z}) d\bar{z} = \int_{-\infty}^{\infty} \bar{z}_k^2 p_{\bar{X}_k}(\bar{z}_k) d\bar{z}_k$$

Def.: Momentos centrais conjuntos de ordem  $m+l$  de duas v.z.  $\bar{X}_k, \bar{X}_j$  do mesmo vetor  $\bar{X}$  é dado por

$$E[(\bar{X}_k - \mu_k)^m (\bar{X}_j - \mu_j)^l] \triangleq \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\bar{z}_k - \mu_k)^m (\bar{z}_j - \mu_j)^l p_{\bar{X}}(\bar{z}) d\bar{z}$$

No caso particular de  $m=l=1$  e  $k=j$  a fórmula geral acima se reduz à variança  $\sigma_k^2 = \text{var}(\bar{X}_k)$

$$E[(\bar{X}_k - \mu_k)^2] \stackrel{\Delta}{=} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\bar{z}_k - \mu_k)^2 p_{\bar{X}_k, \bar{X}_j, \dots, \bar{X}_n}(\bar{z}_1, \dots, \bar{z}_n) d\bar{z}_1 \dots d\bar{z}_n$$

que pode ser reescrita na forma da densidade marginal de  $\bar{X}_k$  como:

$$\sigma_k^2 = E[(\bar{X}_k - \mu_k)^2] = \int_{-\infty}^{\infty} (\bar{z}_k - \mu_k)^2 p_{\bar{X}_k}(\bar{z}_k) d\bar{z}_k$$

O caso de maior interesse na definição dos momentos centrais conjuntos ocorre para  $m=l=1$  e  $k \neq j$  na forma geral. Esses valores geram a covariância dos v.z.  $\bar{X}_k$  e  $\bar{X}_j$ , notada através de  $\sigma_{kj}$  ou  $\text{cov}(\bar{X}_k, \bar{X}_j)$  e definida por

$$\begin{aligned} \sigma_{kj} &= \text{cov}(\bar{X}_k, \bar{X}_j) = E[(\bar{X}_k - \mu_k)(\bar{X}_j - \mu_j)] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\bar{z}_k - \mu_k)(\bar{z}_j - \mu_j) p_{\bar{X}_k, \dots, \bar{X}_n}(\bar{z}_1, \dots, \bar{z}_n) d\bar{z}_k d\bar{z}_j \end{aligned}$$

$$\sigma_{kj} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\bar{z}_k - \mu_k)(\bar{z}_j - \mu_j) p_{\bar{X}_k, \bar{X}_j}(\bar{z}_k, \bar{z}_j) d\bar{z}_k d\bar{z}_j$$

$$\text{Note que: } \sigma_{kj} = E[(\bar{X}_k - \mu_k)(\bar{X}_j - \mu_j)] = E[\bar{X}_k \bar{X}_j - \mu_k \bar{X}_j - \mu_j \bar{X}_k + \mu_k \mu_j]$$

$$\sigma_{kj} = E[\bar{X}_k \bar{X}_j] - E[\mu_j \bar{X}_k] - E[\mu_k \bar{X}_j] + E[\mu_k \mu_j]$$

$$= E[\bar{X}_k \bar{X}_j] - \mu_j E[\bar{X}_k] - \mu_k E[\bar{X}_j] + \mu_k \mu_j$$

$$= E[\bar{X}_j \bar{X}_k] - \mu_j \mu_k - \cancel{\mu_k \mu_j} + \cancel{\mu_k \mu_j}$$

$$\sigma_{kj} = E[\bar{X}_j \bar{X}_k] - \mu_j \mu_k$$

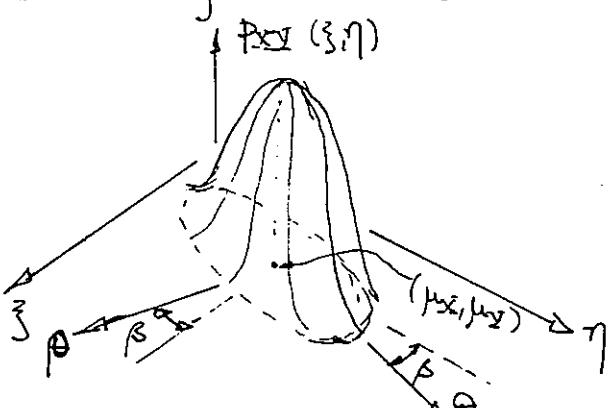
Obtida a covariância, que representa uma quantificação da influência de uma v.a.  $\bar{X}_k$  sobre a v.a.  $\bar{X}_j$ , usz-se representar essa influência por um coeficiente normalizado denominado coeficiente de correlação dado por:

$$\rho_{kj} = \rho(\bar{X}_k, \bar{X}_j) \triangleq \frac{\text{cov}(\bar{X}_k, \bar{X}_j)}{\sigma_k \sigma_j} = \frac{\sigma_{kj}}{\sigma_k \sigma_j}$$

onde  $\sigma_k = \sqrt{\sigma_k^2}$  e  $\sigma_j = \sqrt{\sigma_j^2}$ , são os ditos desvios padrões das v.a.  $\bar{X}_k, \bar{X}_j$ , obtidos da raiz quadrada das respectivas varianças.

Explorando ainda a idéia de bárcentros e momentos de inércia para representar uma analogia mecânica da média e das varianças, respectivamente, a correlação ou covariância indica uma analogia com os produtos de inércia. Então, da mesma forma que num sistema mecânico, dese existir uma transformação de coordenadas que leve a um setor eleitorio em que as componentes são não correlacionadas, um "sistema de eixos principais de inércia" onde os produtos de inércia são nulos. É o que veremos a seguir.

Consideramos uma função densidade conjunta  $f_{XY}(x, y)$  das v.a.  $X$  e  $Y$ , cujas médias são  $\mu_X$  e  $\mu_Y$ , respectivamente. Consideremos um novo sistema de coordenadas  $(\rho, \theta)$ , associado às v.a.'s  $\bar{U}$  e  $\bar{V}$ , cuja origem seja  $(\mu_X, \mu_Y)$  e admitirmos que a covariância entre  $X$  e  $Y$  seja não nula.



Mostraremos que existe um ângulo  $\beta$  entre as coordenadas  $(\bar{z}, \bar{\eta})$  e  $(\bar{\rho}, \bar{\theta})$  de forma que as novas coordenadas a covariâncias se anulam. No caso, temos:

$$\bar{\rho} = (\bar{z} - \mu_x) \cos \beta + (\bar{\eta} - \mu_y) \sin \beta$$

$$\bar{\theta} = -(\bar{z} - \mu_x) \sin \beta + (\bar{\eta} - \mu_y) \cos \beta$$

Em consequência:

$$\begin{aligned} E[\bar{U}] &= E[(\bar{x} - \mu_x) \cos \beta + (\bar{y} - \mu_y) \sin \beta] = \\ &= \cos \beta \{ E[\bar{x}] - \mu_x \} + \sin \beta \{ E[\bar{y}] - \mu_y \} = 0 \end{aligned}$$

$$E[\bar{V}] = 0$$

De mesma forma mostramos que  $E[\bar{V}] = 0$ . Calculamos a covariância para as novas v.z's  $\bar{U}$  e  $\bar{V}$ :

$$\sigma_{UV} = E[(\bar{U} - \mu_U)(\bar{V} - \mu_V)] - E[\bar{U}\bar{V}]$$

Admitimos que  $E[\bar{U}\bar{V}] = 0$ . Então:

$$E \{ [(\bar{x} - \mu_x) \cos \beta + (\bar{y} - \mu_y) \sin \beta] [-(\bar{x} - \mu_x) \sin \beta + (\bar{y} - \mu_y) \cos \beta] \} =$$

$$\begin{aligned} E \{ -(\bar{x} - \mu_x)^2 \cos \beta \sin \beta + (\bar{x} - \mu_x)(\bar{y} - \mu_y) \cos^2 \beta - (\bar{x} - \mu_x)(\bar{y} - \mu_y) \sin \beta + \\ + (\bar{y} - \mu_y)^2 \cos \beta \sin \beta \} = 0 \end{aligned}$$

$$\begin{aligned} -\cos \beta \sin \beta E[(\bar{x} - \mu_x)^2] + \cos^2 \beta E[(\bar{x} - \mu_x)(\bar{y} - \mu_y)] - \sin^2 \beta E[(\bar{x} - \mu_x)(\bar{y} - \mu_y)] \\ + \cos \beta \sin \beta E[(\bar{y} - \mu_y)^2] = 0 \end{aligned}$$

$$-\sigma_x^2 \cos \beta \sin \beta + \cos^2 \beta \sigma_{xy} - \sigma_{xy} \sin^2 \beta + \sigma_y^2 \cos \beta \sin \beta = 0$$

$$(\sigma_y^2 - \sigma_x^2) \cos \beta \sin \beta + \sigma_{xy} (\cos^2 \beta - \sin^2 \beta) = 0$$

$$\frac{1}{2} (\sigma_y^2 - \sigma_x^2) \sin 2\beta + \sigma_{xy} \cos 2\beta = 0$$

$$\tan 2\beta = \frac{2\sigma_{xy}}{\sigma_x^2 - \sigma_y^2}$$

Com a última relação fica determinada a transformação que leva duas v.z. correlacionadas a outras duas não correlacionadas, i.e., por uma transformação conveniente de coordenadas podemos eliminar a covariância. Note que esse resultado é válido ~~quaisquer~~ que seja a distribuição ou densidade conjunta de  $X$  e  $Y$ .

É possível provar, também, que a transformação proposta leva os valores das novas covariâncias  $\sigma_x^2$  e  $\sigma_y^2$  que são ~~maximums~~ ou ~~minimums~~. Esse resultado é obtido provando-se que

$$\frac{d}{d\beta} (\sigma_u^2) = 0 \quad \text{ou} \quad \frac{d}{d\beta} (\sigma_v^2) = 0$$

é obtido através de  $\operatorname{tg} 2\beta = \frac{2\sigma_{xy}}{\sigma_x^2 - \sigma_y^2}$ , mesmo resultado obtido anteriormente.

Por um vetor de v.z.  $\bar{X}$  é comum representar as variâncias e covariâncias na forma de uma matriz, chamada matriz de covariâncias.

Def.: Consideremos um vetor  $\bar{V}$  de duas v.z.  $X$  e  $Y$ . A matriz de covariâncias de  $\bar{V}$  é dada por:

$$\operatorname{Cov}(\bar{V}) = E[(\bar{V} - \mu_V)(\bar{V} - \mu_V)^T] = E \begin{bmatrix} [X - \mu_X] & [Y - \mu_Y] \end{bmatrix} \begin{bmatrix} [X - \mu_X] \\ [Y - \mu_Y] \end{bmatrix}^T = E \begin{bmatrix} [X - \mu_X] & [Y - \mu_Y] \end{bmatrix} \begin{bmatrix} [X - \mu_X] \\ [Y - \mu_Y] \end{bmatrix}$$

$$\operatorname{Cov}(\bar{V}) = E \begin{bmatrix} (X - \mu_X)^2 & (X - \mu_X)(Y - \mu_Y) \\ (Y - \mu_Y)(X - \mu_X) & (Y - \mu_Y)^2 \end{bmatrix} = \begin{bmatrix} E[(X - \mu_X)^2] & E[(X - \mu_X)(Y - \mu_Y)] \\ E[(Y - \mu_Y)(X - \mu_X)] & E[(Y - \mu_Y)^2] \end{bmatrix}$$

$$\operatorname{Cov}(\bar{V}) = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$$

Essa definição pode ser estendida para o caso de dois vetores aleatórios  $\bar{X}^T = (X_1, X_2, \dots, X_n)$  e  $\bar{Y}^T = (Y_1, Y_2, \dots, Y_n)$ .

$$\text{Cov}(\underline{x}, \underline{x}) = \begin{bmatrix} \sigma_{x_1 x_1} & \sigma_{x_1 x_2} & \dots & \sigma_{x_1 x_n} \\ \sigma_{x_2 x_1} & \sigma_{x_2 x_2} & \dots & \sigma_{x_2 x_n} \\ \vdots & \vdots & & \vdots \\ \sigma_{x_n x_1} & \sigma_{x_n x_2} & \dots & \sigma_{x_n x_n} \end{bmatrix}$$

Um caso particular da definição acima é a matriz de covariância do vetor estatístico  $\bar{x}$ :

$$P = \text{Cov}(\bar{x}, \bar{x}) = \text{Cov}(\bar{x}) = \begin{bmatrix} \sigma_{\bar{x}_1}^2 & \sigma_{\bar{x}_1 \bar{x}_2} & \dots & \sigma_{\bar{x}_1 \bar{x}_n} \\ \sigma_{\bar{x}_2 \bar{x}_1} & \sigma_{\bar{x}_2}^2 & \dots & \sigma_{\bar{x}_2 \bar{x}_n} \\ \vdots & \vdots & & \vdots \\ \sigma_{\bar{x}_n \bar{x}_1} & \sigma_{\bar{x}_n \bar{x}_2} & \dots & \sigma_{\bar{x}_n}^2 \end{bmatrix}$$

Note que, como  $\sigma_{x_i x_j} = \sigma_{x_j x_i}$ ,  $i, j = 1, \dots, n$ , a matriz  $\text{Cov}(\bar{x}, \bar{x})$  é uma matriz simétrica.

As matrizes de covariância acima são definidas em termos de momentos centrais. Podemos também definir matrizes com momentos não centrais. Essas matrizes são chamadas matrizes de segunda ordem e são tais que:

$$1) \quad C_x = E[\underline{x}\underline{x}^T] \quad \text{é simétrica}$$

$$2) \quad \text{Cov}(\bar{x}) = C_x - \mu_x \mu_x^T = P$$

$$3) \quad C_{xy} = E[\underline{x}\underline{y}^T] \quad \text{é não simétrica}$$

#### 4.7. VARIÁVEIS ALEATÓRIAS INDEPENDENTES

Antes de caracterizar v.z. independentes, estudaremos a independência de eventos. Com isso, a independência de v.z.'s

é facilmente obtida. As seguintes afirmativas se aplicam ao caso de eventos:

1)  $A \in \mathcal{C}$  são eventos independentes se:

$$P(A \cap C) = P(A) \cdot P(C)$$

2)  $A_1, A_2, A_3$  são eventos independentes se:

$$P(A_1 \cap A_2) = P(A_1)P(A_2)$$

$$P(A_1 \cap A_3) = P(A_1)P(A_3)$$

$$P(A_2 \cap A_3) = P(A_2)P(A_3)$$

$$\stackrel{e}{=} P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$$

3) apesar as 3 principais regras são válidas entre  $A_1, A_2, A_3$  são independentes ou mutuamente independentes.

3)  $A_1, A_2, \dots, A_n$  são eventos independentes se com  $k_1, k_2, k_3, \dots, k_r$  qualquer conjunto de inteiros gerados por  $\{1, \dots, n\}$

$$P(A_{k_1} \cap A_{k_2} \cap \dots \cap A_{k_r}) = P(A_{k_1})P(A_{k_2}) \dots P(A_{k_r})$$

Existem então  $2^n - (n+1)$  relações a serem verificadas.

Entendidas essas assertivas, colocamos agora as definições correspondentes às v.e's.

Def.: Duas v.e. conjuntamente distribuídas são ditas independentes se qualquer das condições equivalentes abaixo for verificada:

$$1.1) F_{X_1 X_2}(z_1, z_2) = F_{X_1}(z_1) \cdot F_{X_2}(z_2)$$

$$1.2) f_{X_1 X_2}(z_1, z_2) = f_{X_1}(z_1) f_{X_2}(z_2)$$

$$1.3) E[f(X_1) g(X_2)] = E[f(X_1)] \cdot E[g(X_2)], \text{ para todas as funções fixas (não escalares) } f(\cdot), g(\cdot), \text{ admitindo que essas expectativas existem.}$$

Note que essa definição é equivalente apenas à definição 1) para eventos e que as definições 2) e 3) dos eventos po-

devem ser estendidas facilmente para duas ou  $n$  variáveis aleatórias. No entanto, vamos nos restringir a apenas duas outras definições de interesse.

Def.: Dizemos que  $X_1, X_2, \dots, X_n$  são v.z's mutuamente independentes, se  $X_1, \dots, X_n$  componentes de  $\bar{X}$ , se:

$$p_{\bar{X}}(\bar{z}) = p_{X_1, X_2, \dots, X_n}(z_1, z_2, \dots, z_n) = p_{X_1}(z_1) p_{X_2}(z_2) \dots p_{X_n}(z_n)$$

Def. Dizemos que  $X_1, \dots, X_n$  são v.z.'s conjuntamente independentes de  $Y_1, \dots, Y_m$  se:

$$p_{X_1, \dots, X_n, Y_1, \dots, Y_m}(z_1, \dots, z_n; \eta_1, \dots, \eta_m) = p_{X_1}(z_1) \dots p_{X_n}(z_n) p_{Y_1}(\eta_1) \dots p_{Y_m}(\eta_m)$$

Note que independência total entre v.z.'s implica em independência mutua e conjunta para qualquer número de v.z's que for considerado.

Def.: Duas v.z.  $\bar{X}_k, \bar{X}_e$  conjuntamente distribuídas são não correlacionadas se:

$$\text{cov}(\bar{X}_k, \bar{X}_e) = 0 \quad \text{ou} \quad \rho_{ke} = 0$$

A independência entre duas v.z's implica na sua não correlação. Essa afirmativa é facilmente provada:

$$\begin{aligned} \text{cov}(\bar{X}_k, \bar{X}_e) &= E[(\bar{X}_k - \mu_k)(\bar{X}_e - \mu_e)] = E[\bar{X}_k \bar{X}_e] - \mu_k \mu_e = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{z}_k \bar{z}_e p_{\bar{X}_k \bar{X}_e}(\bar{z}_k, \bar{z}_e) d\bar{z}_k d\bar{z}_e - \mu_k \mu_e = \\ &= \left( \int_{-\infty}^{\infty} \bar{z}_k p_{\bar{X}_k}(\bar{z}_k) d\bar{z}_k \right) \left( \int_{-\infty}^{\infty} \bar{z}_e p_{\bar{X}_e}(\bar{z}_e) d\bar{z}_e \right) - \mu_k \mu_e = \\ &= \mu_k \cdot \mu_e - \mu_k \mu_e = 0 \end{aligned}$$

Note que a recíproco não é verdadeira já que se  $\text{cov}(X_1, X_2) = 0$  não necessariamente  $P_{X_1, X_2}(\xi_1, \xi_2) = P_{X_1}(\xi_1) P_{X_2}(\xi_2)$ .  
 Existe uma exceção muito importante para esse recíproco, que pode ser colocada na forma de um teorema (ver Jeżwinski, 1970):

Se as v.z. conjuntamente distribuídas gaussianas  $X_1, \dots, X_n$  são não correlacionadas em pares, então elas são mutuamente independentes.

Uma última definição para o caso de um vetor de v.z.'s  $\mathbf{X}$  é a que se refere à função característica. Nesse caso:

Def.: A função característica das v.z. conjuntamente distribuídas  $X_1, X_2, \dots, X_n$  é dada por:

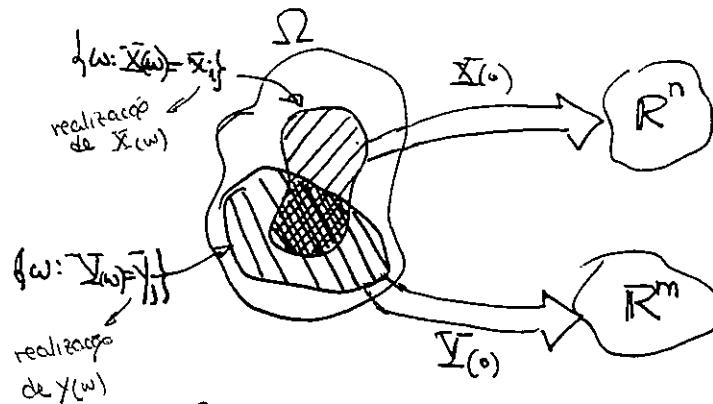
$$\varphi_{X_1, \dots, X_n}(u_1, \dots, u_n) \triangleq E \left[ e^{i \sum_{j=1}^n u_j X_j} \right]$$

As mesmas propriedades da função característica para uma única v.z. descritas anteriormente podem ser estendidas para a função característica de várias v.z.'s.

#### 4.7. PROBABILIDADE CONDICIONADA

Vamos adquirir duas v.z.  $X$  e  $Y$  aplicadas sobre um espaço amostral  $\Omega$  e levando valores em  $\mathbb{R}^n$  e  $\mathbb{R}^m$ , respectivamente. Admitimos, também, que as variáveis assim sejam discretas. Se soubermos que  $Y$  tem uma valor particular realização  $\bar{y}_j$ , esse conhecimento, em geral, irá afetar a determinação da probabilidade que  $X(w) = \bar{x}_i$ , para uma dada realização  $\bar{x}_i$ . A figura abaixo ilustra os conjuntos de interesse em  $\Omega$ , que são  $\{w : X(w) = \bar{x}_i\}$  e  $\{w : Y(w) = \bar{y}_j\}$ .

$(\Omega, \mathcal{B}, P)$



Se não soubermos o valor que  $\mathbb{Y}$  toma, então, simplesmente avlizirímos  $P(\{\omega: X(\omega) = \bar{x}_i\})$  usando a função probabilidade definida para os eventos  $A, A \subset \Omega$  e  $A \in \mathcal{B}$ . Entretanto, se soubermos que  $\mathbb{Y}(\omega) = \bar{y}_j$ , poderíamos restringir nossa atenção ao conjunto  $\{\omega: \mathbb{Y}(\omega) = \bar{y}_j\} \subset \Omega$  ao invés de considerarmos todo  $\Omega$ . Dentro desse conjunto de eventos elementares (seus)  $\omega$  queremos determinar a probabilidade do conjunto de  $\omega$  tal que  $X(\omega) = \bar{x}_i$ , isto é, a probabilidade de  $\{\omega: X(\omega) = \bar{x}_i \text{ e } \mathbb{Y}(\omega) = \bar{y}_j\}$  (ver figura), relativa ao conjunto  $\{\omega: \mathbb{Y}(\omega) = \bar{y}_j\}$ . A essa nova função probabilidade diz-se o nome de probabilidade condicionada, isto é, a probabilidade de que  $X(\omega) = \bar{x}_i$  condicionada a que  $\mathbb{Y}(\omega) = \bar{y}_j$ . e define-se:

$$P(X(\omega) = \bar{x}_i | \mathbb{Y}(\omega) = \bar{y}_j) = \frac{P(\{X(\omega) = \bar{x}_i \text{ e } \mathbb{Y}(\omega) = \bar{y}_j\})}{P(\{\mathbb{Y}(\omega) = \bar{y}_j\})}$$

espaço amostral  
total para este evento (restrições: espaço da  $\mathbb{Y}_d$ )

Note que, nesse caso, a nova função probabilidade só precisa ser definida para conjuntos  $A \subset \{\omega: \mathbb{Y}(\omega) = \bar{y}_j\} \subset \Omega$ ,  $A \in \mathcal{B} \subset \mathcal{F}$ , uma  $\sigma$ -álgebra menor que para o caso de probabilidade não condicionada. (menor que  $\mathcal{B}$ )

Para generalizar esse conceito, dado um experimento  $\mathfrak{F} \triangleq (\Omega, \mathcal{B}, P)$  e dados dois eventos  $A \in \mathcal{C}$ , a função de probabilidade condicionada  $P(A|C)$  do evento  $A$ , dado o evento  $C$ , é:

$P(A C) \triangleq \frac{P(A \cap C)}{P(C)}$	com $P(C) > 0$
--	----------------

Exemplo: Consideremos o experimento do lançamento de um dado e temos:

$$\Omega = \{1, 2, \dots, 6\}$$

$\mathcal{P}$  = classe de todos subconjuntos de  $\Omega$   $\leftarrow \underline{\underline{*}}$

$$P(i) = \frac{1}{6}, 1 \leq i \leq 6$$

Consideremos o evento  $C = \{\text{sair no par}\} = \{2, 4, 6\}$ . Construimos uma nova função de probabilidade condicionada a  $C$  por:

$$P(1|C) = P(3|C) = P(5|C) = 0$$

$$P(2|C) = P(4|C) = P(6|C) = \frac{1}{3}$$

Dado o experimento  $\mathfrak{F} = (\Omega, \mathcal{P}, P)$  podemos construir um novo experimento  $(\Omega, \mathcal{P}, P(\cdot|C))$ . Pode-se mostrar que  $P(\cdot|C)$  satisfaz todos os axiomas da Teoria da Probabilidade.

Se  $A$  e  $B$  são eventos independentes, então:

$$P(A \cap B) = P(A) P(B)$$

$$P(A|B) = \frac{P(A) P(B)}{P(B)} = P(A)$$

e se verifica que qualquer que seja a ocorrência em  $B$ , o resultado em  $A$  não se modifica.

Vemos então que no caso de v.a. discretas, a definição conhecida para probabilidade condicional é totalmente coerente. No entanto, como a definição exige  $P(C) > 0$  no denominador, este falha para o caso de v.a. contínuas. A extensão para o caso de v.a. contínuas não é, em consequência, trivial, exigindo um formalismo matemático muito difícil (ver Janiński, 1970; Maybeck, 1979).

Para v.a. contínuas a definição de probabilidade condicional é feita através de funções de densidade de probabilidade

condicionada. Para isso, consideremos  $X$  e  $Y$  duas v.a. contínuas conjuntamente distribuídas. Dadas a distribuição e a densidade conjuntas de  $X$  e  $Y$ , pode-se estabelecer a função densidade da v.a.  $X$  dada uma realização  $\eta$  da v.a.  $Y$  através de:

$$p_{X|Y}(z|\eta) = \frac{p_{XY}(z,\eta)}{p_Y(\eta)} = \frac{p_{XY}(z,\eta)}{\int_{-\infty}^{\infty} p_{XY}(z,\eta) dz}$$

Note que o denominador nesse caso é a densidade marginal de  $Y$ . ( $p_Y(\eta)$ )

Se trocarmos  $X$  com  $Y$  na relação acima:

$$p_{Y|X}(\eta|z) = \frac{p_{XY}(z,\eta)}{p_X(z)} \rightarrow \text{densidade conjunta}$$

de onde:  $p_{XY}(z,\eta) = p_{Y|X}(\eta|z) p_X(z)$

Substituindo o valor de  $p_{XY}(\cdot, \cdot)$  no primeiro caso, chegamos à chamada Regra de Bayes, de fundamental importância em Teoria de Estimação:

$$p_{X|Y}(z|\eta) = \frac{p_{Y|X}(\eta|z) p_X(z)}{p_Y(\eta)}$$

Caracterizada a função densidade condicionada para v.a.  $X, Y$  muito bem definidas, as mesmas considerações feitas para densidades não condicionadas podem ser feitas, isto é, pode-se definir momentos, centrais ou não, formular funções características, estudar para o caso de vetores de v.a's, etc. Os passos a serem seguidos são exatamente os mesmos.

Como exemplo, caracterizemos a esperança condicionada de uma v.a.  $X$  em relação a uma v.a.  $Y$ , isto é, o mo-

mento de 1º ordem condicionado. Este pode ser obtido de:

$$E[X|\Sigma] \triangleq \int_{-\infty}^{\infty} z p_{X|\Sigma}(z|\eta) dz$$

Finalmente, a partir das considerações já feitas, as seguintes propriedades da função densidade condicionada podem ser provadas:

$$1) p_{X|\Sigma}(z|\eta) \geq 0$$

$$2) \int_{-\infty}^{\infty} p_{X|\Sigma}(z|\eta) dz = 1$$

$$3) p_X(z) = E[p_{X|\Sigma}(z|\eta)] = \int_{-\infty}^{\infty} p_{X|\Sigma}(z|\eta) p_{\Sigma}(\eta) d\eta$$

$$4) \text{ Se } X \text{ e } \Sigma \text{ são independentes } p_{X|\Sigma}(z|\eta) = p_X(z)$$

## CAP. 5 - TEORIA DE ESTIMAÇÃO: SISTEMAS ESTÁTICOS

Com os conceitos obtidos através da Teoria de Probabilidade é possível começar a desenvolver os fundamentos da Teoria de Estimação, aplicando-a a sistemas estáticos, ou seja, invariáveis no tempo. Embora a literatura sobre o assunto consagre também o nome de Estimação de Parâmetros, julgamos conveniente evitá-lo para não confundir com outros itens que serão desenvolvidos no futuro.

Isto posto, a formulação geral do problema de estimação de sistemas estáticos pode ser colocado como a seguir:

- \*  $X$  é um vetor cujas componentes devem ser estimadas;
- \*\*  $\Theta$  é um conjunto, colocado na forma de um vetor, de observações feitas.
- \*\*\* Admite-se, qualquer que seja  $\Theta$ , que existe uma relação funcional conhecida entre  $X$  e  $\Theta$ , isto é:

$$\Theta = \Theta(X)$$

O vetor  $\Theta$  é um vetor de dimensão  $m$ :

$$\Theta^T = [\Theta_1, \Theta_2, \dots, \Theta_m]$$

enquanto  $X$  é um vetor de dimensão  $n$ :

$$X^T = [X_1, X_2, \dots, X_n]$$

O problema de estimação consiste em, dadas as observações  $\Theta$ , estimar o vetor  $X$ .

Tomemos como exemplo o caso de um voo espacial. Podemos caracterizar  $X$  como o vetor cujas componentes são a posição e a velocidade da sonda ou do satélite em um dado instante  $t_0$  e mais outras constantes físicas de interesse.  $\Theta$  é um

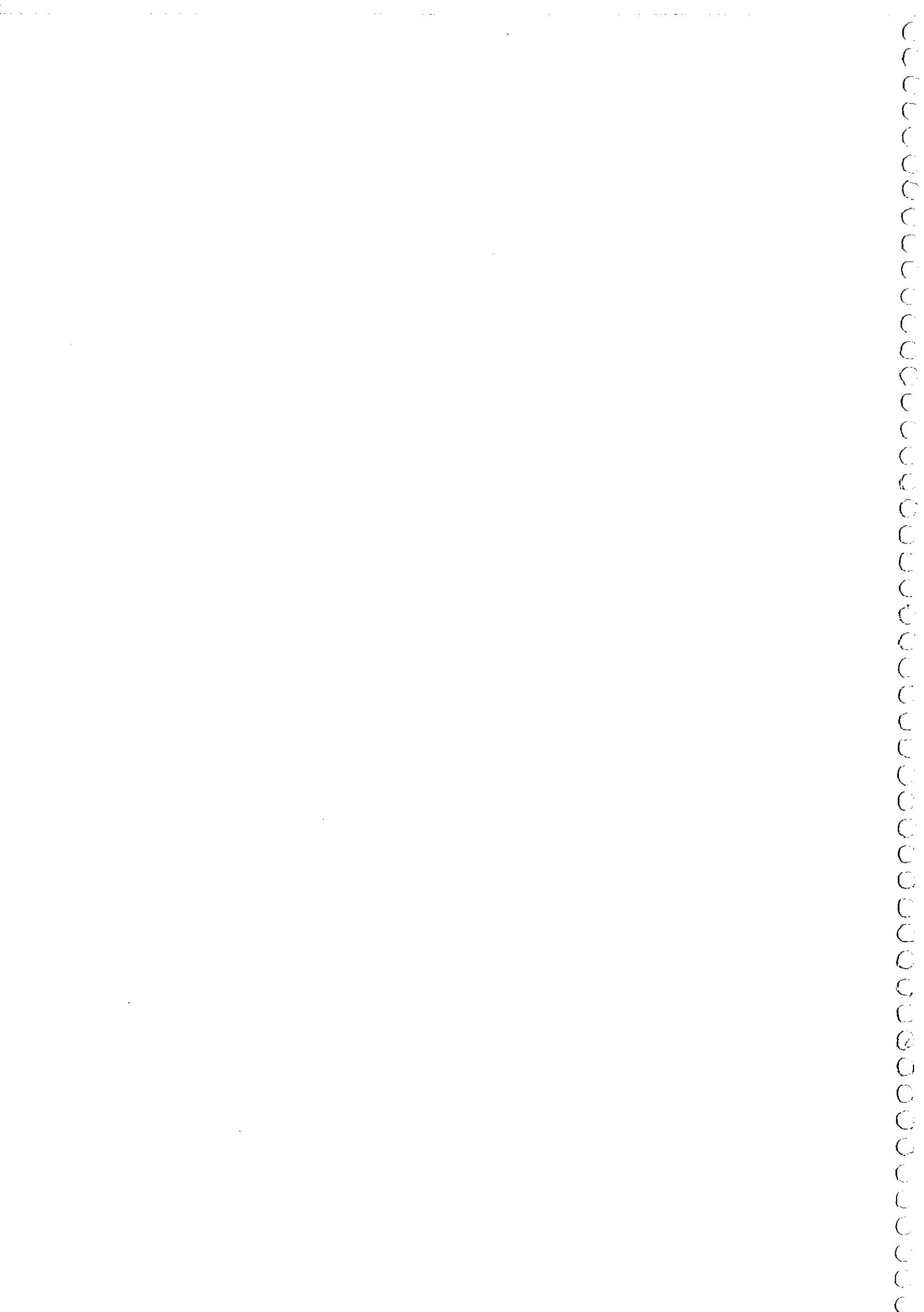
vetor cujas componentes são centrais de medidas obtidas por radar do alcance, de variação do alcance e ângulos relativos ao radar.

O problema, como descrito, lógicamente não tem solução única, visto que podemos combinar de diversas maneiras as observações, produzindo diversas estimativas de  $\bar{X}$ , dígua em dante designado vetor de estado. Isto nos leva a considerar ou procurar uma estimativa "melhor" ou "ótima" em relação a outras possíveis estimativas, isto é, somas induzidas a estabelecer um critério de otimização para gerar a estimativa

### 5.1. CRITÉRIOS DE OTIMIZAÇÃO

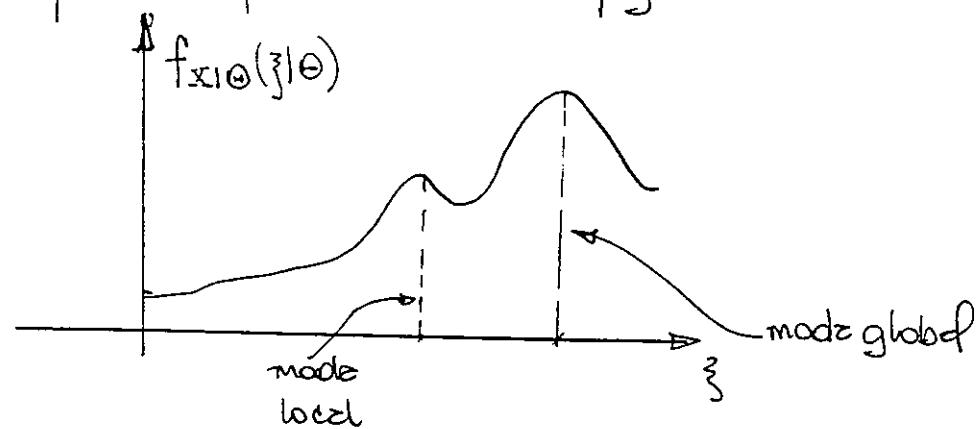
Em breves palavras, o problema de estimação como colocado tanto, dado o nível de incerteza no vetor de estado e as observações, determinar uma estimativa a posteriori (em relação às observações) do vetor de estado que envolve a menor incerteza. Um estimador é, então, segundo Schweppe (1973), uma função que aplica-se sobre o vetor de observações produz uma estimativa ótima segundo algum critério. Lógicamente, a "função estimador" está ligada ao critério de otimização escolhido.

A escolha do critério de otimização leva a uma série enorme de procedimentos de estimação. Um primeiro classificação, bastante comum, é a de se separar os estimadores em Bayesianos e não Bayesianos. Independentemente, zindz, do critério escolhido de otimização, um estimador Bayesiano é aquele que se preocupa em obter, para qualquer instante de tempo, a função densidade de probabilidade do vetor de estado condicionado às medidas feitas. Estimadores não Bayesianos são aqueles em que não há a necessidade de se obter a função densidade condicionada. Um exemplo clássico de Estimador não Bayesiano é o método



dos mínimos quadrados, seja na sua variação padronizada, seja na forma direta de Gauss.

Este curso pretende fazer uma abordagem a um tipo especial de estimador Bayesiano, o Filtro de Kalman. O filtro de Kalman é um estimador Bayesiano não (<sup>viciado</sup>) de menor variação do erro. Isso significa que, em condições ideais, o filtro de Kalman produz um estimativo que em média deve antecipar o valor verdadeiro do vetor de estados e que tem a menor dispersão do erro em torno do valor verdadeiro. Outro tipo de estimador Bayesiano são os filtros de máxima verossimilhança (maximum likelihood). A diferença entre o Filtro de Kalman e o de máxima verossimilhança reside exatamente no critério de otimização adotado. Enquanto no filtro de Kalman o critério é minimizar os termos da diagonal da matriz de covariância dos erros, no filtro de máxima verossimilhança o critério passa a ser o valor de maior probabilidade, já estima-se o modo ou o pico da função densidade, o valor mais provável em toda distribuição. Critérios de máxima verossimilhança são largamente empregados em processos de identificação de parâmetros, embora possam carregar convergência para máximas locais, conforme esquematizado na figura.



Com exceção das textos específicos sobre identificação

(Astrom, Eykoff, 1974) a grande maioria das textos em estimação trabalha sobre o conceito de filtro de mínima variação do erro (Jazwinski, 1970; Maybeck, 1979, etc.). Deve também ser ressaltado que no caso linear gaussiano esses dois tipos de filtros Bayesianos produzem as mesmas estimativas. Então, a escolha da abordagem de mínima variação do erro é normalmente justificada pelo seguinte:

- \* facilmente obtível e compreensível
- \* o desenvolvimento de uma teoria linear permite o estudo de problemas não lineares com boas aproximações
- \* para distribuições gaussianas dos erros, o estimador de mínima variação empata com vários outros tipos de erros
- \* traz a grande vantagem de permitir aplicação a problemas com pouco ou nenhum conhecimento das funções densidade das v.a.'s envolvidas.

Os critérios usados para montagem de um estimador segundo a mínima variação do erro não visando obter as seguintes condições:

1º condição) Se  $\hat{x}$  for o valor estimado, então é conveniente que:

$$E[\hat{x}] = \bar{x}$$

onde  $\bar{x}$  é o valor verdadeiro. Isto significa que a média da estimativa é o valor verdadeiro, não conhecido, do vetor de estado  $x$ .

Sóksfazer a relação acima implica em se ter um estimador não tendencioso ou não viciado (unbiased) já que na ausência de erros de medida e de incerteza no procedimento de estimativa

$$\hat{x} = E[\hat{x}] = \bar{x}$$

2º condição) Desde que qualquer estrutura, em geral,

contém erros em relação ao valor verdadeiro, é natural que se tente minimizar esse erro. Definindo o erro como a diferença entre a estimativa e o valor verdadeiro:

$$e = \hat{X} - X$$

Verificamos que o erro  $e$  é uma variável aleatória que pode ser positiva ou negativa, já que  $\hat{X}$  e  $X$  são aleatórias. Não é possível, então, minimizar diretamente o erro  $e$  devemos escolher uma função conveniente de  $e$  para trabalhar. Intuitivamente, uma boa escolha seria minimizar o erro médio quadrático das componentes de  $e$ . Então

$$C_e = E[e e^T] = E[(\hat{X} - X)(\hat{X} - X)^T]$$

$$C_e = \{c_{ij}\}, c_{ij} = E[(\hat{x}_i - x_i)(\hat{x}_j - x_j)]$$

e os termos da diagonal,  $c_{ii} = E[(\hat{x}_i - x_i)^2]$  são erros médios quadráticos.

Quando essa 2ª condição é satisfeita temos um estimador com mínimo erro médio quadrático.

3ª condição) Se juntarmos agora as 2 primeiras condições num único estimador, teremos:

$$E[\hat{X}] = X$$

$$C_e = E[(\hat{X} - X)(\hat{X} - X)^T] = E[(\hat{X} - E[\hat{X}])(\hat{X} - E[\hat{X}])^T]$$

e, com isso, a matriz  $C_e$  passa a ser a matriz de covariâncias da estimativa, onde os termos da diagonal são as variações em cada uma das componentes.

Um estimador em que a matriz  $C_e$  na forma apresentada é minimizada é um estimador não viciado de mínima variancia. Esse tipo de estimador engloba os critérios de otimização que serão por nós utilizados e, desde já, se coloca que uma de suas

grandes vantagens é que o processo matemático para sua obtenção em cossos práticos é muito mais simples que no caso de outros estimadores.

Fazendo um resumo, a esta altura já temos quatro das cinco componentes fundamentais do problema de estimação, a saber:

- 1) as variáveis a serem estimadas, vetor  $\mathbf{X}$ ;
- 2) as medidas ou observações disponíveis, vetor  $\boldsymbol{\Theta}$ ; cjb de numeros
- 3) a relação funcional entre observações e variáveis,  
 $\boldsymbol{\Theta} = \boldsymbol{\Theta}(\mathbf{X})$ ;
- 4) o critério de otimização

A única componente que deve ser fornecida ao problema de estimação é o modelo matemático das incertezas presentes no sistema, ou seja, uma descrição das incertezas que afetam o vetor de estados e o vetor de observações. Sobre esse último ponto os comentários aparecerão à medida que se for desenvolvendo o curso.

#### → 5.2. TEOREMA FUNDAMENTAL DE GAUSS-MARKOV

O estimador não viésado de mínima variação pode ser conhecido apenas para poucas situações. Uma dessas situações acontece quando supomos que a estimativa pode ser obtida como combinação linear de um conjunto de dados (medidas) observados em instantes discretos de tempo.

Se  $\boldsymbol{\Theta}$  é o vetor de dados cujos m componentes  $\theta_i$  são vários tipos de dados obtidos em diversos instantes  $t_i$  e se  $\mathbf{X}$  é o vetor cujas n componentes  $x_j$  estão sendo estimadas, então, uma estimativa linear  $\hat{\mathbf{X}}$  de  $\mathbf{X}$  tem a forma:

$$\hat{\mathbf{X}} = \mathbf{A} \boldsymbol{\Theta}$$

com  $\boldsymbol{\Theta} = \boldsymbol{\Theta}(\mathbf{X})$  e  $\mathbf{X}$  eletrórios e  $\mathbf{A}$  matriz ( $n \times m$ ) de constantes cujos elementos devem ser achados, então podemos formular o segui-

te teorema:

TEOREMA FUNDAMENTAL DE GAUSS - MARKOV

Se  $X \in \Theta$  são vetores aleatórios com matrizes de momentos de 2ª ordem conhecidas

$$C_X = E[XX^T] \quad ; \quad C_{X\theta} = E[X\theta^T] \quad e \quad C_\theta = E[\theta\theta^T]$$

com  $C_\theta$  não singular, então a estimativa linear de mínimo erro médio quadrático  $\hat{X}$  de  $X$  é:

$$\hat{X} = C_{X\theta} C_\theta^{-1} \theta \quad , \text{ isto é, } A = C_{X\theta} C_\theta^{-1}$$

A matriz de erros associada é dada por:

$$C_e = C_X - C_{X\theta} C_\theta^{-1} C_{X\theta}^T \quad (A)$$

Se, além disso, for verdade que

$$X = C_{X\theta} C_\theta^{-1} E(\theta)$$

então  $\hat{X}$  como acima é uma estimativa não viésada já que:

$$E[\hat{X}] = E[C_{X\theta} C_\theta^{-1} \theta] = C_{X\theta} C_\theta^{-1} E[\theta]$$

e teremos um estimador não viésado de menor variação linear. Nesse caso,  $C_e$  é a própria matriz de covariância de  $\hat{X}$ .

A prova do teorema é feita rapidamente partindo de:

$$\hat{X} = A\theta$$

$$C_e = E[ee^T] = E[(\hat{X}-X)(\hat{X}-X)^T]$$

$$C_e = E[(A\theta - X)(A\theta - X)^T] = E[(A\theta - X)(\theta^T A^T - X^T)]$$

$$C_e = E[A\theta\theta^T A^T] - E[\bar{x}\theta^T A^T] - E[A\theta\bar{x}^T] + E[\bar{x}\bar{x}^T]$$

$$C_e = AE[\theta\theta^T]A^T - E[\bar{x}\theta^T]A^T - AE[\theta\bar{x}^T] + E[\bar{x}\bar{x}^T]$$

$$C_e = AC_0 A^T - C_{x\theta} A^T - AC_{\theta x} + C_x$$

Como  $C_{\theta x} = E[\theta\bar{x}^T] = E[(\bar{x}\theta^T)^T] = C_{\bar{x}\theta}^T$

$$C_e = AC_0 A^T - C_{x\theta} A^T - AC_{\theta x}^T + C_x$$

Usamos agora a seguinte identidade de matrizes (ver capítulo)

$$(A - BC^{-1})C(A - BC^{-1})^T - BC^{-1}B^T = ACAT - 3A^T(BA^T)^T$$

e identificando, nos 3 primeiros termos à direita da igualdade:

$$A \Rightarrow A ; \quad B \Rightarrow BC_{x\theta} , \quad C \Rightarrow C_0$$

resulta:

$$\rightarrow C_e = (A - C_{x\theta} C_0^{-1})C_0 (A - C_{x\theta} C_0^{-1})^T - C_{x\theta} C_0^{-1} C_{x\theta}^T + C_x$$

Como as matrizes  $C_0$  e  $C_0^{-1}$  são não negativas definidas, então:

$(A - C_{x\theta} C_0^{-1})C_0 (A - C_{x\theta} C_0^{-1})^T$  e  $C_{x\theta} C_0^{-1} C_{x\theta}^T$  necessariamente tem na diagonal principal elementos positivos. Deve-se que  $C_e$  terá valor mínimo apenas na hipótese de o primeiro termo se anular.

Logo:

$$A = C_{x\theta} C_0^{-1}$$

O que completa a prova do teorema.

O teorema de Gauss-Markov, embora seja a base de todo

o desenvolvimento a seguir, é de aplicação extremamente restrita, visto exigir o conhecimento de  $C_0$ ,  $C_{x0}$  e  $C_x$ , i.e., de todas as estatísticas envolvidas no problema. Além disso, existe o problema de se ter que inverter a matriz  $C_0$ , de dimensão  $(m \times m)$ , o que, em geral, não é muito comodo. A importância do Teorema de Gauss-Markov reside na sua imediata aplicação aos casos particulares que veremos a seguir.

### 5.3. APLICAÇÃO A CASOS PARTICULARES: OBSERVAÇÕES LINEARES

O caso de maior interesse na aplicação do Teorema Fundamental de Gauss-Markov é aquele que envolve uma relação linear entre as observações e o vetor de estado, com todas as suas decorrências.

Admitimos, então, que o vetor de observações  $\Theta$  está relacionado com o vetor  $X$  por:

$$\boxed{\Theta = B X + v}$$

onde  $B$  é uma matriz  $(m \times n)$  de valores constantes e conhecidos e  $v$  é um vetor  $(m \times 1)$  que representa a incerteza nas medidas, o erro aleatório ou ruído nas observações  $\Theta$ .

Verifiquemos que cada componente de  $\Theta$  é tal que:

$$\Theta_1 = B_{11} X_1 + v_1$$

$$\Theta_2 = B_{21} X_1 + v_2$$

$$\vdots \quad \vdots$$

$$\Theta_m = B_{m1} X_1 + v_m$$

onde cada  $B_i = \{B_{i1}; B_{i2}, \dots, B_{in}\}$  e  $B = \begin{bmatrix} B_{11} \\ B_{21} \\ \vdots \\ B_{m1} \end{bmatrix}$

Nesse caso:

$$\begin{aligned} C_{x\theta} &= E[X\theta^T] = E[X(BX + V)^T] = E[X(X^T B^T + V^T)] \\ &= E[XX^T]B^T + E[XV^T] \end{aligned}$$

$$C_{x\theta} = C_x B^T + C_{xv}$$

$$\begin{aligned} C_\theta &= E[\theta\theta^T] = E[(BX + V)(BX + V)^T] = E[(BX + V)(X^T B^T + V^T)] \\ &= BE[XX^T]B^T + E[VX^T]B^T + BE[XV^T] + E[VV^T] \end{aligned}$$

$$C_\theta = BC_x B^T + C_{xv}^T B^T + BC_{xv} + C_v$$

Substituindo em  $\hat{X} = C_{x\theta} C_\theta^{-1} \theta$  e em  $C_e = C_x - C_{x\theta} C_\theta^{-1} C_{x\theta}^T$ , obtemos a nova forma do estimador, dada por:

estimador linear com observações lineares

$$\begin{cases} \hat{X} = [C_x B^T + C_{xv}] [BC_x B^T + (BC_{xv})^T + BC_v + C_v]^{-1} \theta \\ C_e = C_x - (C_x B^T + C_{xv}) [BC_x B^T + (BC_{xv})^T + BC_{xv} + C_v]^{-1} (C_x B^T + C_{xv})^T \end{cases}$$

Observe que a estimativa ótima exige conhecimento numérico das matrizes  $C_x$ ,  $C_{xv}$  e  $C_v$ , o que não é muito simples, principalmente em termos de se estabelecer a correlação entre o estado  $X$  e a incerteza nas medidas,  $V$ , através de matriz  $C_{xv}$ .

Um caso particular das observações lineares no estimador linear de Gauss-Markov que torna o estimador bastante simples ocorre quando podemos afirmar que o vetor de estado  $X$  é não correlacionado com o ruído  $V$ . Então:

$$C_{xv} = 0$$

e as expressões acima se reduzem a:

Estimador linear  
com observações  
lineares e não-correladas  
relação estado-ruido

$$\hat{\mathbf{x}} = \mathbf{C}_x \mathbf{B}^T [\mathbf{B} \mathbf{C}_x \mathbf{B}^T + \mathbf{C}_v]^{-1} \theta$$

$$\mathbf{C}_e = \mathbf{C}_x - \mathbf{C}_x \mathbf{B}^T [\mathbf{B} \mathbf{C}_x \mathbf{B}^T + \mathbf{C}_v]^{-1} (\mathbf{C}_x \mathbf{B}^T)^T$$

Note que o estimador na forma proposta, exige a inversão de uma matriz,  $[\mathbf{B} \mathbf{C}_x \mathbf{B}^T + \mathbf{C}_v]^{-1}$ , de dimensão  $(m \times m)$ ; o que pode não ser um bom acréscimo quando o número de observações,  $m$ , for muito maior que o número de componentes do vetor de estado,  $n$ . Uma forma equivalente pode ser obtida através de identidades de matrizes (capítulo 1).

De fato, usando a identidade:

$$\Delta \mathbf{B}^T (\mathbf{C} + \mathbf{B} \mathbf{A} \mathbf{B}^T)^{-1} = (\mathbf{A}^{-1} + \mathbf{B}^T \mathbf{C}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{C}^{-1}$$

e identificando, na 1ª forma do estimador:

$$\mathbf{C}_x = \mathbf{A} ; \quad \mathbf{B} = \mathbf{B} ; \quad \mathbf{C}_v = \mathbf{C}_e$$

teremos:

$$\hat{\mathbf{x}} = [\mathbf{C}_x^{-1} + \mathbf{B}^T \mathbf{C}_v^{-1} \mathbf{B}]^{-1} \mathbf{B}^T \mathbf{C}_v^{-1} \theta$$

Usando a 2ª identidade de matrizes:

$$(\mathbf{C}^{-1} + \mathbf{B}^T \Delta^{-1} \mathbf{B})^{-1} = \mathbf{C} - \mathbf{C} \mathbf{B}^T (\mathbf{B} \mathbf{C} \mathbf{B}^T + \mathbf{A})^{-1} \mathbf{B} \mathbf{C}$$

e identificando, na equação da matriz de covariância dos erros:

$$\mathbf{C}_v = \mathbf{A} ; \quad \mathbf{B} = \mathbf{B} ; \quad \mathbf{C}_x = \mathbf{C}$$

chegamos a:

$$\mathbf{C}_e = (\mathbf{C}_x^{-1} + \mathbf{B}^T \mathbf{C}_v^{-1} \mathbf{B})^{-1}$$

Resumindo, a nova forma do estimador é:

Estimador linear  
com observações li-  
neares e não-correladas  
estados-ruido

$$\hat{\mathbf{x}} = [\mathbf{C}_x^{-1} + \mathbf{B}^T \mathbf{C}_v^{-1} \mathbf{B}]^{-1} \mathbf{B}^T \mathbf{C}_v^{-1} \theta$$

$$\mathbf{C}_e = (\mathbf{C}_x^{-1} + \mathbf{B}^T \mathbf{C}_v^{-1} \mathbf{B})^{-1}$$

A inversão zogiz é de matrizes ( $n \times n$ ), o menor ole matriz  $C_x$  que é um elado de perda para o estimador. A segundo formulão, portanto, é mais conveniente no caso de o número de observações ser maior que a dimensão do vetor de estado.

Finalmente, ainda no caso de observações lineares, em muitos problemas de estimação muito pouco ou nada se conhece sobre o estado verdadeiro  $X$ . Mesmo quando o estado  $X$  e os ruídos são não correlacionados, o estimador apresentado exige o conhecimento da matriz de momentos de segundo ordem  $C_x$ , que, no caso de não se conhecer muito sobre  $X$ , deve ser muito difícil de estabelecer. Então, no caso de pouca se conhecer sobre  $X$ , é razoável admitir que a matriz de momentos de  $X$  tenha, na sua diagonal, termos que tendam para infinito, e, com isso:

$$\text{Cov}(X) \rightarrow \infty$$

Lembrando, zogiz, que:

$$\text{Cov}(X) = C_x - \mu_X \mu_X^T$$

ou

$$C_x = \text{Cov}(X) + \mu_X \mu_X^T$$

então,  $C_x \rightarrow \infty$  quando  $\text{Cov}(X) \rightarrow \infty$ , já que o vetor  $\mu_X$  é um vetor de constantes mesas no caso em que o estado é desconhecido. Basicamente, com isso estamos admitindo incerteza infinita ou ilimitada sobre o conhecimento do estado  $X$ .

Mas, se  $C_x \rightarrow \infty$  então é também verdade que  $C_x^{-1} \rightarrow 0$  e a estimativa de  $X$  nas formas obtidas acima se simplifica bastante, passando a:

Estimador linear com  
observações lineares,  
não-correlação estado-ruído  
e desconhecimento  
do estado

$$\hat{X} = (B^T C_V^{-1} B)^{-1} B^T C_V^{-1} \theta$$

$$C_e = (B^T C_V^{-1} B)^{-1}$$

Esse critério de avaliação é muito importante e forma o base de muitos procedimentos de estimação ótima. Recunbremos as hipóteses usadas para sua obtenção:

- a.  $C_v$  pode ser calculado;
- b.  $C_v$  existe;
- c. Nada se conhece sobre o estado verdadeiro  $X$ , de modo que  $C_x \rightarrow \infty$ ;
- d. Não há correlações entre o estado e o ruído, i.e.,  $C_{xv} = 0$ ;
- e. A estimativa ótima de  $X$ ,  $\hat{X}$ , é o estimador linear que minimiza os elementos da diagonal de  $C_e = C_{\hat{X}} - X$ ;
- f. Vede a relação  $\Theta = BX + V$ .

#### → Exemplo:

Neste exemplo traz-se do problema de detecção (identificação) de uma frequência conhecida, onde se aproveita para mostrar como resolver um problema não linear variando no tempo à estimativa linear de um sistema estético. logicamente essa transformação não sempre é possível de ser realizada, mas vede a advertência de que antes de se tentar a solução para exemplares do problema devem ser examinados todos os caminhos que levam a simplificações.

Nosso problema é, então, detectar um sinal periódico com frequência conhecida  $f$  na presença de ruídos ou incertezas. Admitimos que as medidas  $\Theta_i$  do sinal tem a forma:

$$\Theta_i = a \operatorname{sen}(wt_i + \alpha) + v_i \quad w = 2\pi f$$

onde  $a$  é a amplitude desconhecida do sinal,  $\alpha$  é o ângulo de fase desconhecido,  $t_i$  representa os diversos instantes de medida e os  $v_i$  são as componentes de ruído nas observações. A velocidade angular  $w = 2\pi f$  é conhecida e devemos estimar a amplitude e o ângulo de fase.

O problema é não linear já que  $\alpha$ , a ser estimado, é o argumento do seno. Definimos variáveis auxiliares para contornar esse problema:

$$x(t) = a \operatorname{sen}(wt + \alpha)$$

$$y(t) = \frac{d}{dt} x(t) = aw \cos(wt + \alpha)$$

Logo:

$$\dot{y}(t) = \frac{d}{dt} y(t) = -aw^2 \operatorname{sen}(wt + \alpha) = -w^2 x$$

e podemos colocar as variáveis auxiliares na forma:

$$\frac{d}{dt} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -w^2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

É possível mostrar que a matriz de transição para esse sistema tem a forma (ver Liebelt, 1971):

$$\Phi(t, s) = \frac{1}{\omega} \begin{bmatrix} w \cos \omega t & \operatorname{sen} \omega t \\ -w^2 \operatorname{sen} \omega t & w \cos \omega t \end{bmatrix}$$

e podemos escrever:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \Phi(t, t_0) \begin{bmatrix} x(t_0) \\ y(t_0) \end{bmatrix} = \frac{1}{\omega} \begin{bmatrix} w \cos \omega t & \operatorname{sen} \omega t \\ -w^2 \operatorname{sen} \omega t & w \cos \omega t \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

para todo instante  $t$ .

Notemos agora que, para qualquer  $t$ :

$$(wx)^2 + y^2 = z\omega^2$$

$$\frac{wx}{y} = \operatorname{tg}(wt + \alpha)$$

Em particular, para  $t=0$ ,  $x(t_0) = x_0$  e  $y(t_0) = y_0$

$$\omega = \left[ \frac{\omega^2 x_0^2 + y_0^2}{\omega^2} \right]^{1/2} = \left[ x_0^2 + \left( \frac{y_0}{\omega} \right)^2 \right]^{1/2}$$

$$\frac{\omega x_0}{y_0} = \operatorname{tg}(0+\alpha) \Rightarrow \alpha = \arctg\left(\frac{\omega x_0}{y_0}\right)$$

Agora, se tivermos estimativas ótimas de  $x_0$  e de  $y_0$ , então as estimativas ótimas de  $\omega$  e  $\alpha$  podem ser obtidas. Precisamos, contudo, escrever o vetor de observações  $\theta$  em termos de  $x_0$  e  $y_0$ . Como:

$$\theta_i = \omega \operatorname{sen}(\omega t_i + \alpha) + v_i$$

$$\theta_i = x_i + y_i \quad \text{ou}$$

$$\theta_i = [1 \ ; \ 0] \begin{bmatrix} x_i \\ y_i \end{bmatrix} + v_i$$

Usando a matriz de transição já obtida:

$$\theta_i = \frac{1}{\omega} [1 \ ; \ 0] \begin{bmatrix} \omega \cos \omega t_i & \operatorname{sen} \omega t_i \\ -\omega^2 \operatorname{sen} \omega t_i & \omega \cos \omega t_i \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + v_i$$

$$\theta_i = \frac{1}{\omega} [\omega \cos \omega t_i \ \ \operatorname{sen} \omega t_i] \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + v_i$$

Colocando na forma matricial:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_m \end{bmatrix} = \frac{1}{\omega} \begin{bmatrix} \omega \cos \omega t_1 & \operatorname{sen} \omega t_1 \\ \omega \cos \omega t_2 & \operatorname{sen} \omega t_2 \\ \vdots & \vdots \\ \omega \cos \omega t_m & \operatorname{sen} \omega t_m \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

$$\theta = B X + v$$

que, fuzinante, tem a forma linear e podemos escrever o estimador:

$$\hat{\underline{x}} = \begin{bmatrix} \hat{x}_0 \\ \hat{y}_0 \end{bmatrix} = (\underline{B}^T \underline{C}_V^{-1} \underline{B})^{-1} \underline{B}^T \underline{C}_V^{-1} \underline{\theta}$$

$$\underline{C}_e = (\underline{B}^T \underline{C}_V^{-1} \underline{B})^{-1}$$

As estimativas obtidas de  $\omega$  e  $\alpha$  ficam:

$$\hat{\omega} = [\hat{x}_0^2 + (\hat{y}_0/\omega)^2]^{1/2} \quad ; \quad \hat{\alpha} = \arctg\left(\frac{\hat{y}_0}{\hat{x}_0}\right)$$

com matriz de covariância de erros, para  $\omega$  e  $\alpha$ , dado por:

$$P_{\omega\alpha} = H (\underline{B}^T \underline{C}_V^{-1} \underline{B})^{-1} H^T$$

onde

$$H = \frac{\partial(\hat{\underline{x}}, \alpha)}{\partial(x_0, y_0)} = \begin{bmatrix} \hat{x}_0/\hat{\omega} & \hat{y}_0/\hat{\omega}^2 \hat{\alpha} \\ w \cos^2 \hat{\alpha} & -w \hat{x}_0 \cos^2 \hat{\alpha} \end{bmatrix}$$

#### 5.4. ESTIMAÇÃO POR MÍNIMOS QUADRADOS

A teoria de mínimos quadrados, como já atado anteriormente, tem suas origens nas trabalhos de Gauß, no começo do século XIX. Atualmente é uma técnica largamente empregada em Engenharia principalmente por ser uma abordagem determinística, ou não probabilística, o que leva a estar baseada em conceitos muito mais simples que os estimadores baseados em princípios probabilísticos.

Para se obter o estimador de mínimos quadrados admite que a variável  $\theta_i$  esteja linearmente relacionada com o

conjunto de  $n$  variáveis  $\underline{X} = (x_1, \dots, x_n)$  a ser estimado através de funções de regressão do tipo:

$$\theta_i = b_{i1} x_1 + b_{i2} x_2 + \dots + b_{in} x_n$$

Fazendo  $i = 1, 2, \dots, m$ , isto é, usando  $m$  observações  $\theta$  para gerar o vetor  $\underline{X}$ , admitindo conhecidos os coeficientes  $b_{ij}$ , podemos colocar a relação linear entre  $\theta$  e  $\underline{X}$  na forma matricial:

$$\theta = B \underline{X}$$

Note que estamos admitindo até aqui que as medidas são perfeitas, i.e., não existe nenhum componente aleatório na definição da relação entre  $\theta$  e  $\underline{X}$ .

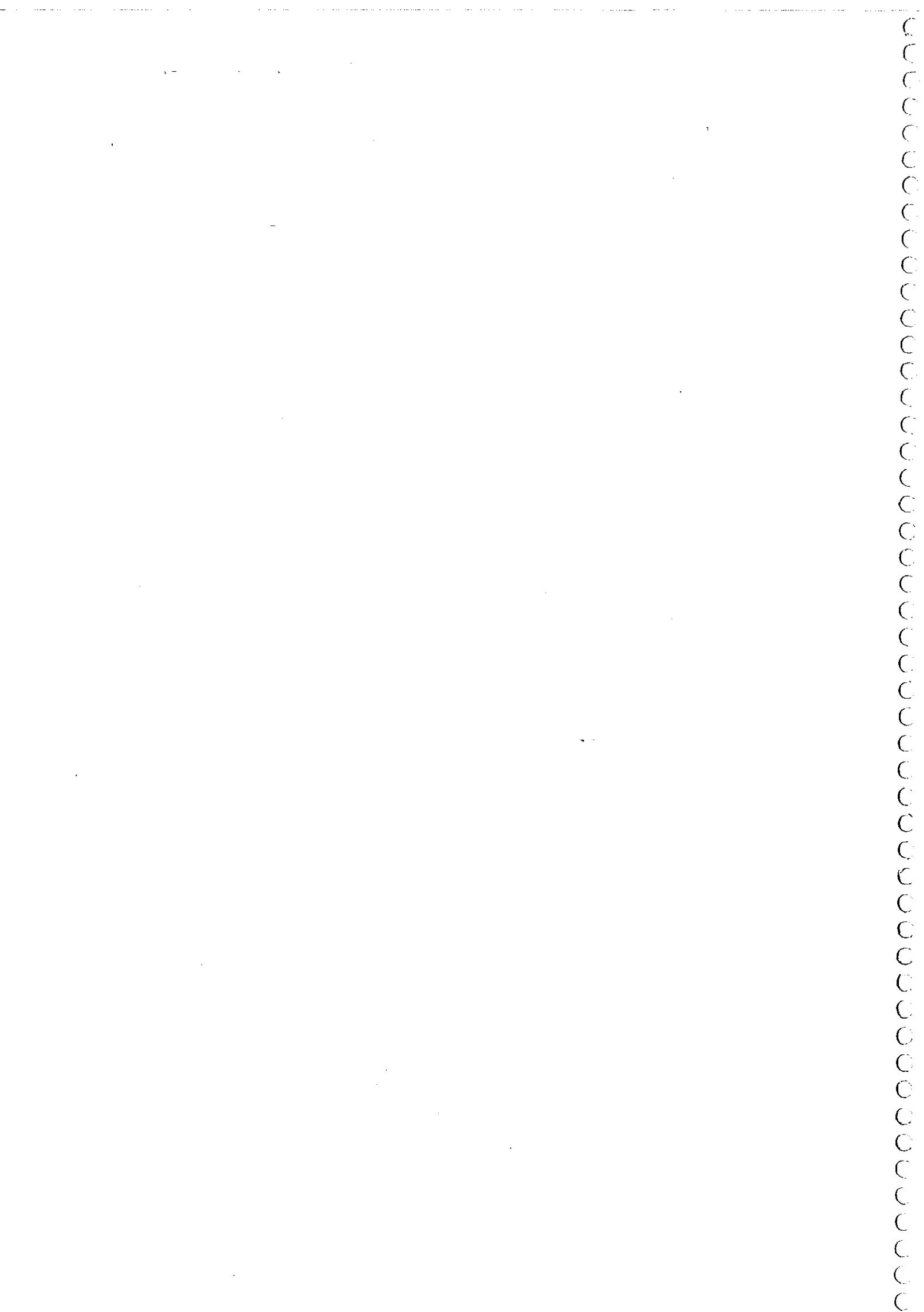
Lógicamente, obviamente, se infere que se  $B$  for inversa, supondo  $\theta$  e  $\underline{X}$  de mesmas dimensões, o problema estaria resolvido. Como, em geral,  $m \neq n$ , então define-se um vetor de erro, de dimensão  $m$ :

$$e = \theta - B \underline{X} ; e \stackrel{\Delta}{=} \text{erro ou resíduo}$$

e procura-se uma estimativa de  $\underline{X}$  baseada nos minimizados do vetor de erro  $e$ . A escolha mais óbvia para o estimador é procurar minimizar o erro quadrático e defina-se um critério de otimização determinístico a ser minimizado:

$$J = \sum_{i=1}^m e_i^2 = e^T e^* = (\theta - B \underline{X})^T (\theta - B \underline{X})$$

Fazendo:



$$\frac{\partial J}{\partial \hat{x}} \Big|_{\hat{x}=\hat{x}} = -2B^T\theta + 2B^TB\hat{x} = 0$$

obtemos a chamada equação normal dos resíduos

$$B^TB\hat{x} = B^T\theta$$

de onde sai o estimador de mínimos quadrados ordinários na forma:

$$\boxed{\hat{x} = (B^TB)^{-1}B^T\theta}$$

Poderíamos ter usado como critério minimizar uma função ponderada dos resíduos ou erros. Este último é a forma mais comum de uso nas metodologias de mínimos quadrados e corresponde a minimizar:

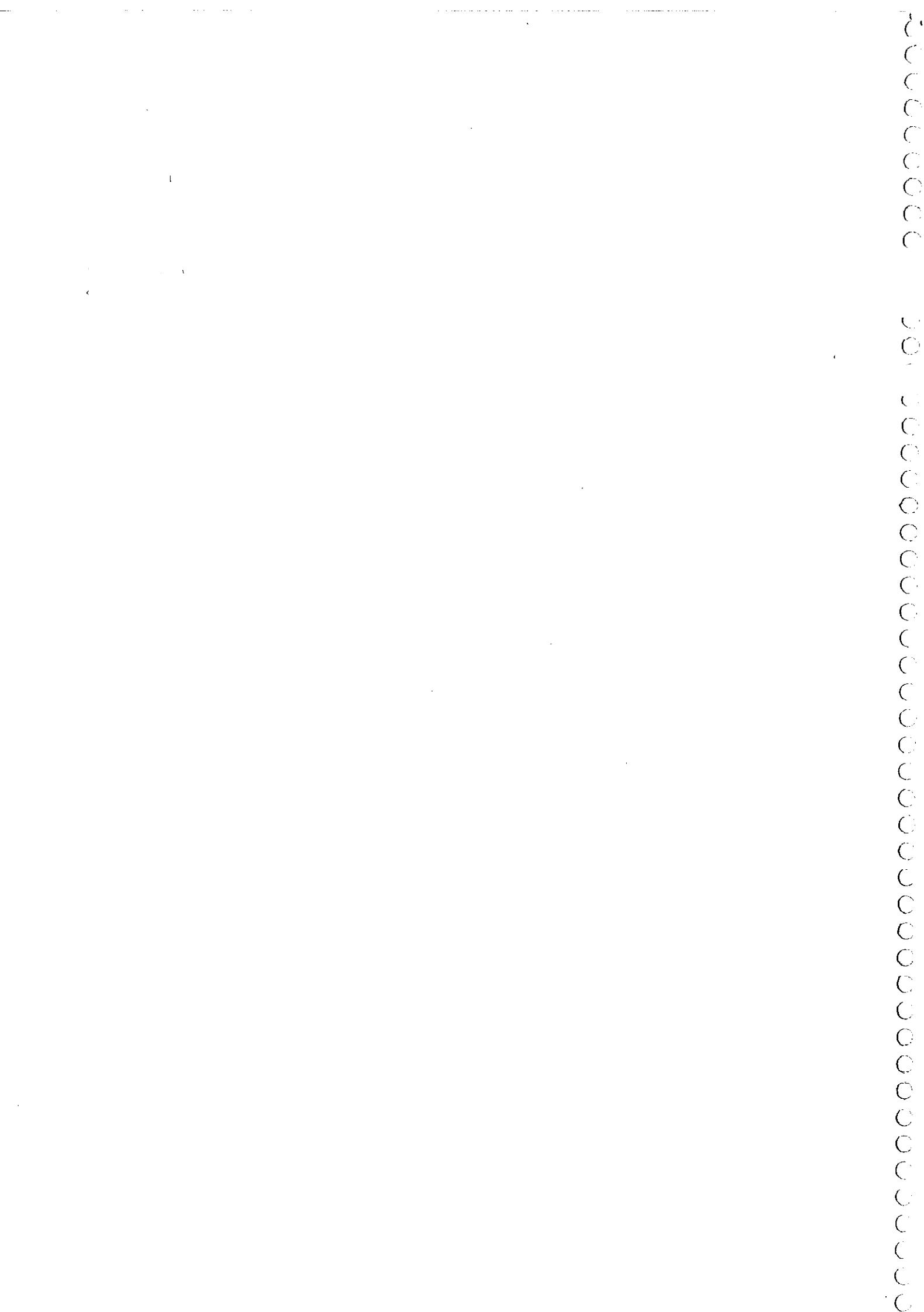
$$J_W = e^T W e = (\theta - B\hat{x})^T W (\theta - B\hat{x})$$

com  $\otimes$   $W$  matriz simétrica e definida positiva. Nesse caso, o estimador obtido é dito de mínimos quadrados ponderados e tem a forma:

$$\boxed{\hat{x}_W = (B^T W B)^{-1} B^T W \theta}$$

 \* Muito importante é observar que se  $W = C_V^{-1}$ , o estimador acima é exatamente o mesmo que o estimador linear com observações lineares, não correlação estado-ruído e desconhecimento do estado obtido no item anterior.

Aproveitando agora para introduzir algumas conceitos que auxiliem na análise do desempenho de um estimador. A per de mostrar que, na hipótese de  $C_V^{-1} = W$ , o estimador de mínimos quadrados ponderados é não-viesado e de menor variância, é também possível mostrar que esse estimador é con-



sistente e eficiente. Os conceitos de consistência e eficiência são extremamente importantes no projeto de um estimador e pode ser provado que os estimadores probabilísticos obtidos no item anterior são consistentes e eficientes.

Consideremos, então, o modelo de observações

$$\theta = BX + v$$

onde  $v$  é um ruído estacionário de média nula,  $E[v] = 0$ , e não correlacionado com  $X$ , tanto para o estimador de mínimos quadrados ordinários como para o estimador ponderado.

Prove-se, então:

1)  $\hat{X}$  como acima, é um estimador não viciado, i.e.,

$$E[\hat{X}] = X.$$

Partindo de  $\hat{X} = (B^T B)^{-1} B^T \theta$  e substituindo  $\theta$  da relação  $\theta = BX + v$  I

$$\hat{X} = (B^T B)^{-1} B^T BX + (B^T B)^{-1} B^T v$$

$$\hat{X} = X + (B^T B)^{-1} B^T v$$

$$E[\hat{X}] = E[X] + E[(B^T B)^{-1} B^T v] = E[X] + (B^T B)^{-1} B^T E[v]$$

$$E[\hat{X}] = E[X] = X - \text{valor real, um número}$$

2) O estimador de mínimos quadrados ponderados é estimador de menor variação se  $W = C_v^{-1}$

Definindo a matriz de covariância do erro ou resíduo  $e = \hat{X} - X$ :

$$C_e = E\{(\hat{X} - X)(\hat{X} - X)^T\}$$

e sendo  $C_v = E\{vv^T\}$ , já que o ruído é de média nula

$$C_e = (B^T B)^{-1} B^T C_v B (B^T B)^{-1}$$

para o estimador de mínimos quadrados ordinários e que

$$C_{e_W} = (B^T B)^{-1} B^T W C_v W^T B (B^T B)^{-1}$$

para o estimador de mínimos quadrados ponderados.

É possível então mostrar que se  $w = C^{-1}v$

$$\boxed{C_{ew=C^{-1}v} \leq C_{ew}}$$

onde a 2<sup>a</sup> matriz se refere a matriz de erro para qualquer outra possível escolha de w.

3) Prova-se, também, que  $\hat{x}$  na forma dada é um estimador eficiente, no sentido de que se os ruidos  $v$  são <sup>(independentes)</sup> idênticamente distribuídos e i.i.d., i.e.:

$$C_v = E[vv^T] = \sigma^2 I$$

e, de médio zero e mesma variança, então:

$$C_e = C_{ew=C^{-1}v} = C_{e_{\hat{x}w}}$$

i.e., a matriz de erros quadráticos é a própria matriz de médio zero e mesma variança.

4) Finalmente, o estimador de mínimos quadrados, nas hipóteses do item 3, é também um estimador consistente, isto é:

$$\lim_{m \rightarrow \infty} C_e = 0$$

onde m é o número de observações disponíveis.

Formas bastante comuns dos estimadores de mínimos quadrados são as que usam a matriz  $C_v$  diagonal e as que usam formas polinomiais de ajuste.

Usar a matriz  $C_v$  diagonal implica admitir a hipótese de que não há correlação entre as observações, ou seja, nenhuma das observações influir sobre qualquer uma das outras. Então:

$$C_V = \begin{bmatrix} E[v_1^2] & 0 & \dots & 0 \\ 0 & E[v_2^2] & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & E[v_m^2] \end{bmatrix} \Rightarrow C_V^{-1} = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_m \end{bmatrix}$$

com  $w_i = \frac{1}{E[v_i^2]}, i = 1, 2, \dots, m$

Nesse caso, teremos:

$$B^T C_V^{-1} \theta = (B_1^T, B_2^T, \dots, B_m^T) \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_m \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_m \end{bmatrix} = \sum_{i=1}^m B_i^T w_i \theta_i$$

com  $B_i$ , vetor lento tal que:

$$B_i = (b_{i1}, b_{i2}, \dots, b_{in})$$

de forma análoga

$$B^T C_V^{-1} B = \sum_{i=1}^m B_i^T w_i B_i$$

Usando as relações acima nas equações do estimador linear com observações lentas, não correlação estado-ruido e desconhecimento do estado, essas últimas são transformadas em:

Estimador de mínimos quadrados com  $C_V$  diagonal

$$\hat{x} = \left( \sum_{i=1}^m B_i^T w_i B_i \right)^{-1} \left( \sum_{i=1}^m B_i^T w_i \theta_i \right) = (B^T C_V^{-1} B)^{-1} B^T C_V^{-1} \theta$$

$$C_e = \left( \sum_{i=1}^m B_i^T w_i B_i \right)^{-1} = (B^T C_V^{-1} B)^{-1}$$

Para obtermos a forma polinomial juntaremos as medidas redimensionadas as variáveis a serem estimadas pelo modelo polinomial

$$\theta(t) = x_1 + x_2 t + \dots + x_n t^{n-1}$$

e, considerando as incertezas, as observações ficam:

$$\theta_i = x_1 + x_2 t_i + \dots + x_n t_i^{n-1} + v_i$$

$$\begin{aligned}\Theta_1 &= x_1 + x_2 t_2 + \dots + x_n t_2^{n-1} + v_1 \\ &\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ \Theta_m &= x_1 + x_2 t_m + \dots + x_n t_m^{n-1} + v_m\end{aligned}$$

ou, em forma matricial:

$$\begin{bmatrix} \Theta_1 \\ \Theta_2 \\ \vdots \\ \Theta_m \end{bmatrix} = \begin{bmatrix} 1 & t_1 & \dots & t_1^{n-1} \\ 1 & t_2 & \dots & t_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_m & \dots & t_m^{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

$$\boldsymbol{\Theta} = \mathbf{B} \mathbf{x} + \mathbf{v}$$

Se admitirmos, além de todas as hipóteses já feitas, que  $\mathbf{v}$  é diagonal e que os  $w_i$  e, em consequência, as covariâncias, são unitárias, i.e:

$$E[v_i^2] = w_i = 1, \quad i = 1, 2, \dots, m$$

então

$$S = \sum_{i=1}^m \mathbf{B}_i^T \mathbf{w}_i \mathbf{B}_i = \begin{bmatrix} \sum_{i=1}^m 1 & \sum_{i=1}^m t_i \dots \sum_{i=1}^m t_i^{n-1} \\ \sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2 \dots \sum_{i=1}^m t_i^n \\ \vdots & \vdots \\ \sum_{i=1}^m t_i^{n-1} & \sum_{i=1}^m t_i^n \dots \sum_{i=1}^m t_i^{2(n-1)} \end{bmatrix}$$

$$R = \sum_{i=1}^m \mathbf{B}_i w_i \Theta_i = \begin{bmatrix} \sum_{i=1}^m \Theta_i \\ \sum_{i=1}^m t_i \Theta_i \\ \vdots \\ \sum_{i=1}^m t_i^{n-1} \Theta_i \end{bmatrix}$$

Com isso, as relações do estatudor se reduzem a:

Estimador de minimos quadrados no formato matricial

$$\boxed{\begin{aligned}\hat{\mathbf{x}} &= S^{-1} R \\ C_e &= S^{-1}\end{aligned}}$$

Para finalizar este tópico sobre mínimos quadrados, fizemos um resumo sobre o que foi apresentado. Deve-se destacar que a abordagem por mínimos quadrados é essencialmente determinística já que não é necessário nenhum conhecimento estatístico sobre o problema em mãos. Quando se usa um díodo estatístico, no caso de mínimos quadrados ponderados, fazendo a matriz de pesos  $W = C_V^{-1}$ , então, o algoritmo de mínimos quadrados compete com o estimador linear com observações knewres não viésadas e de número variância do erro, obtido por abordagens probabilísticas. Nesse caso, os estimadores equivalentes são também consistentes e eficientes, propriedades altamente desejadas para um estimador. Finalmente, as formas mais usadas para o estimador de mínimos quadrados envolvem a matriz  $W$  de ponderação diagonal e ajuste por formas polinomiais. Quanto a este último deve-se chamar a atenção para a forma polinomial escolhida pois a forma do polinomio pode não ser coerente com os dados nem que as observações sejam perfeitas, ou seja, isentas de erros. Como consequência disso, o vetor de erro  $\epsilon$  pode não evoluir frequentemente o erro total de proximidade no estimador. Além disso, existem entre um erro adicional pelo fato de que a forma polinomial dificilmente se ajusta de modo exato ao sistema físico que produz os dados de observação.

## 5.5. INFORMAÇÕES A PRIORI COM OBSERVAÇÕES LINEARES

Os estimadores obtidos nos itens anteriores são estimadores do tipo fornecido ("batch"), isto é, dado um conjunto de  $m$  observações, processamos essas  $m$  observações para ter uma estimativa do vetor de estados. No caso de as observações chegarem em instantes diferentes do tempo, um estimador do tipo "batch" far-

cará a que se estre a chegada de todas as informações para que então se dé a partida no estimador. Um estimador do tipo recursivo evita que se tenha de esperar por todas as informações. Desse maneira, a partida é dada com um estimador tipo "batch" e à medida que chegam as outras informações estes serão acrescentados no algoritmo para gerar uma melhor estimativa através de um estimador recursivo.

Imaginemos, então, a seguinte situação:

- recebemos  $r$  observações do fenômeno e computamos com essas  $\theta^r$  observações uma estimativa  $\hat{X}^r$ ;
- encerrada essa parte recebemos mais  $s$  observações de forma que  $r + s = m$

A pergunta é como usar essas novas informações de modo a melhorar a qualidade da nossa estimativa?

A partir de  $\theta^r$ , conjunto das  $r$  primeiras medidas, e de  $\hat{X}^r$ , estimativa a partir de  $\theta^r$ , admitimos que as observações são bensortes, i.e.:

$$\theta^r = B^r X + v^r$$

onde  $v^r$  é o ruído nas  $r$  primeiras observações

Nessas hipóteses englobam os fatos de que:

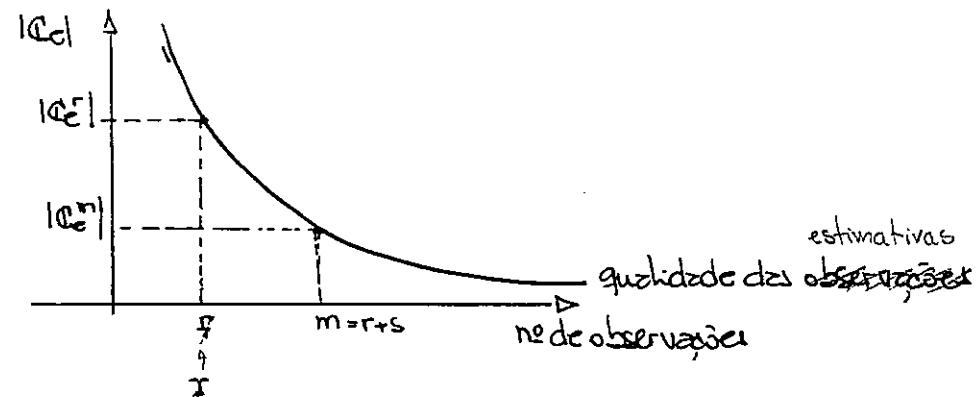
- $C_v^{-1}$  existe, qualquer que seja a dimensão de  $v$ ;
- O ruído é não correlacionado com o estado:  $C_{xv}^r = C_{vv}^r = 0$ ;
- Não se conhece sobre o estado verdadeiro  $X$ :  $C_x^{-1} = 0$

o que nos remete ao último estimador obtido no item 5.3:

$$\begin{cases} \hat{X}^r = [B^r (C_v^r)^{-1} B^r]^{-1} B^r (C_v^r)^{-1} \theta^r \\ C_e^r = C_{x-x}^r = [B^r (C_v^r)^{-1} B^r]^{-1} \end{cases}$$

Recebendo mais  $s$  observações desejamos melhorar a qu-

lidade das nossas estimativas o que pode ser traduzido pelo gráfico abaixo onde  $|C_{\text{rel}}|$  é mais ou menos conveniente para a matriz  $C_e$ .



Para as novas informações adquiridas também teremos de:

$$\theta^s = B^s \bar{x} + v^s$$

e não correlação estado - ruído :  $C_{xv}^s = 0$ .

A quarta hipótese fundamental é que

4) as novas observações  $\theta^s$  o ruído é não correlacionado com o ruído nas  $r$  observações já processadas, isto é:

$$C_{vv}^{rs} = 0$$

Combinação as  $m$  observações com:

$$\theta^m = \begin{bmatrix} \theta^r \\ \theta^s \end{bmatrix} ; \quad B^m = \begin{bmatrix} B^r \\ B^s \end{bmatrix} ; \quad v^m = \begin{bmatrix} v^r \\ v^s \end{bmatrix}$$

$$\begin{bmatrix} \theta^r \\ \theta^s \end{bmatrix} = \begin{bmatrix} B^r \\ B^s \end{bmatrix} \bar{x} + \begin{bmatrix} v^r \\ v^s \end{bmatrix}$$

Expressemos  $C_v^m$  em termos de  $C_v^r$  e  $C_v^s$ :

$$C_v^m = E[v^m v^{mT}] = E \left[ \begin{bmatrix} v^r \\ v^s \end{bmatrix} \begin{bmatrix} v^{rT} & v^{sT} \end{bmatrix} \right].$$

$$\mathbb{C}_V^m = E \begin{bmatrix} V^r V^{rT} & V^r V^{sT} \\ \hline V^s V^{rT} & V^s V^{sT} \end{bmatrix} = \begin{bmatrix} \mathbb{C}_V^r & \mathbb{C}_{V^r V^s}^r \\ \hline \mathbb{C}_{V^s V^r}^r & \mathbb{C}_V^s \end{bmatrix}$$

$$\mathbb{C}_V^m = \begin{bmatrix} \mathbb{C}_V^r & 0 \\ \hline 0 & \mathbb{C}_V^s \end{bmatrix}$$

Como a estimativa  $\hat{\mathbf{x}}^m$  baseada em m informações é dada por:

$$\hat{\mathbf{x}}^m = [B^{rT} (\mathbb{C}_V^m)^{-1} B^r]^{-1} B^{mT} (\mathbb{C}_V^m)^{-1} \theta^m$$

$$\hat{\mathbf{x}}^m = \left[ [B^r, B^s] \begin{bmatrix} \mathbb{C}_V^r & 0 \\ 0 & \mathbb{C}_V^s \end{bmatrix}^{-1} [B^r] \right]^{-1} [B^r, B^s] \begin{bmatrix} \mathbb{C}_V^r & 0 \\ 0 & \mathbb{C}_V^s \end{bmatrix}^{-1} \begin{bmatrix} \theta^r \\ \theta^s \end{bmatrix}$$

$$\hat{\mathbf{x}}^m = [B^{rT} (\mathbb{C}_V^r)^{-1} B^r + B^{sT} (\mathbb{C}_V^s)^{-1} B^s]^{-1} [B^{rT} (\mathbb{C}_V^r)^{-1} \theta^r + B^{sT} (\mathbb{C}_V^s)^{-1} \theta^s]$$

Usando agora as expressões de  $\hat{\mathbf{x}}^r$  e  $\mathbb{C}_e^r$  apresentadas no inicio deste item, a expressão do estimador se reduz a:

estimador linear com informações a priori

$$\hat{\mathbf{x}}^m = \mathbb{C}_e^r \left[ (\mathbb{C}_e^r)^{-1} \hat{\mathbf{x}}^r + \underbrace{B^{sT} (\mathbb{C}_V^s)^{-1} \theta^s}_{\substack{\text{informações novas para serem processadas}}} \right]$$

$$\mathbb{C}_e^m = \left[ (\mathbb{C}_e^r)^{-1} + B^{sT} (\mathbb{C}_V^s)^{-1} B^s \right]^{-1}$$

$\hookrightarrow$  matriz de covariância de erros das m medidas.

O estimador acima nos mostra como combinar a estimativa  $\hat{\mathbf{x}}^r$ , a matriz de erro antigo  $\mathbb{C}_e^r$  e as novas observações  $\theta^s$  para gerar, recursivamente, uma nova estimativa  $\hat{\mathbf{x}}^m$  que contenha os termos de variação de  $\mathbb{C}_e^m$ , com a única hipótese adicional de que  $\mathbb{C}_{V^s} = 0$ .

Uma outra consequência importante é que faz provédo que, nas hipóteses de  $\mathbb{C}_{V^s} = 0$ ,  $\mathbb{C}_x^{-1} = 0$  e  $\mathbb{C}_{V^s}^{rs} = 0$ , não importa

é ordem em que as observações são processadas e nem é necessário agrupá-las em cada processamento intermediário.

O procedimento nos sugere, então, uma forma razoável de tratar um problema com número m de observações muito grande, através do computador.

## 5.6. A FORMA DE KALMAN

Um dos passos fundamentais de um filtro de Kalman, chamado passo de atualização, corresponde, num filtro com observações discretas, a um problema de estimacão de um sistema estatístico com informação a priori. A chamada forma de Kalman pode ser obtida diretamente do estimador linear com informações a priori desenvolvido no item anterior.

Reescrevemos esse estimador que tem a forma:

$$\hat{X}^m = C_e^m [C_e^r]^{-1} \hat{X}^e + B^s T [C_V^s]^{-1} \Theta^s$$

$$C_e^m = [(C_e^r)^{-1} + B^s T [C_V^s]^{-1} B^s]^{-1}$$

Definimos uma matriz (nxs) dita matriz de ganhos de Kalman por:

$$\bar{K}^s = C_e^m B^s T [C_V^s]^{-1} = [(C_e^r)^{-1} + B^s T [C_V^s]^{-1} B^s]^{-1} B^s T [C_V^s]^{-1}$$

Usamos a identidade de matrizes:

$$AB^T (C + B A B^T)^{-1} = (A^{-1} + B^T C^{-1} B)^{-1} B^T C^{-1}$$

identificando:

$$(C_e^r)^{-1} \rightarrow A^{-1}; B^e \rightarrow B; (C_V^s)^{-1} \rightarrow C^{-1}$$

para obter a matriz de ganhos como:

$$K^s = C_e^r B^{sT} (C_V^s + B^s C_e^r B^{sT})^{-1}$$

Para obter uma expressão de  $C_e^m$  e  $K$ , usamos uma outra identidade de matrizes. De fato, da definição de  $C_e^m$ :

$$C_e^m = [(C_e^r)^{-1} + B^{sT} (C_V^s)^{-1} B^s]^{-1}$$

através da identidade:

$$(C^{-1} + B^T A^{-1} B)^{-1} = C - C B^T (B C B^T + A)^{-1} B C$$

identificando

$$(C_e^r)^{-1} \rightarrow C^r, B^{sT} \rightarrow B, (C_V^s)^{-1} \rightarrow A^{-1}$$

chegamos a:

$$C_e^m = C_e^r - C_e^r B^{sT} (B^s C_e^r B^{sT} + C_V^s)^{-1} B^s C_e^r$$

Comparando com a forma de Kalman:

$$C_e^m = C_e^r - K B^s C_e^r \Rightarrow C_e^m = [I - K B^s] C_e^r$$

Substituindo, agora os valores de  $C_e^m$  e  $K$  na expressão de  $\hat{X}^m$ :

$$\hat{X}^m = (C_e^r - K B^s C_e^r) (C_e^r)^{-1} \hat{X}^r + K \theta^s$$

$$\hat{X}^m = \hat{X}^r - K B^s \hat{X}^r + K \theta^s$$

$$\hat{X}^m = \hat{X}^r + K (\theta^s - B^s \hat{X}^r)$$

Nesse estimador de informações a priori foi, então, transformado num estimador de Kalman, que consiste das três seguintes passos:

a) Cálculo da Matriz de Azulhos de Kalman:

$$K = C_e^r B^{sT} (C_V^s + B^s C_e^r B^{sT})^{-1}$$

2) Cálculo do vetor média  $\hat{X}^m$ :

$$\hat{X}^m = \hat{X}^r + K[\theta^s - B^s \hat{X}^r]$$

3) Cálculo da matriz de erros:

$$\mathbb{C}_e^m = \mathbb{C}_e^r - K B^s \mathbb{C}_e^r$$

A matriz de ganhos  $K$  pode ser vista como realmente um fator de ganho para a nova estimativa a partir de  $\hat{X}^r$  e  $\theta^s$ , dando correndo o nome ganho de kalman. Note que a matriz  $K$  contém todas as informações sobre as s observações  $\theta^s$  a serem acrescentadas.

Note, também, que a diferença entre a forma de kalmán e o estimador do item anterior é de sentido inverso. se a dimensão do estado  $\mathbf{x}$ ,  $n$ , for maior que o número de observações  $s$ , do vetor  $\theta^s$ , então a forma de kalmán é preferível por implicar na inversão de uma matriz de ordem maior, correspondente a  $(\mathbb{C}_e^s + B^s \mathbb{C}_e^r B^{sr})^{-1}$ , o que sugere até se arescavar as observações para a unir, desde que possível. Caso contrário, pela mesma razão de inversão de matrizes, é preferível usar o estimador na forma obtida no item anterior.

## CAPÍTULO 6 - PROCESSOS ESTOCÁSTICOS

Se o fenômeno físico que pretendemos estudar evolui no tempo com alguma incerteza, ou seja, existem componentes para os quais não dispomos de leis físicas adequadas, somos obrigados a longo prazo da modelagem por processos estocásticos.

A teoria de processos estocásticos é o último requisito teórico necessário para a construção de filtros, estimadores para sistemas dinâmicos e a analogia que se pode fazer é a de que processos estocásticos representam para a teoria probabilística o mesmo que o cálculo diferencial e integral para os sistemas determinísticos.

Para conceituar o que é um processo estocástico, consideremos dada um experimento  $\mathfrak{F}$ , especificado por seu espaço amostral  $\Omega$ , cujos elementos são designados por  $w$ ; pela  $\sigma$ -álgebra formada por subconjuntos de  $\Omega$ , os eventos; e pela função de probabilidade associada a esses eventos. Dado, então, o experimento  $(\Omega, \mathcal{B}, P)$  vamos definir uma família de variáveis aleatórias.

Para cada resultado  $\xi$  ( $X(\cdot, w) \leq \xi$ ) designamos, de acordo com certas regras, uma função do tempo  $X(t, \xi)$  real ou complexa. Dessa maneira criamos uma família de funções, uma para cada resultado  $\xi$ . Para caracterizar os membros deste família usamos um ou mais índices, que se constituem em parâmetros suficientes, pois, uma família parametrizada de variáveis aleatórias. No nosso caso específico, será usado apenas um parâmetro, gerando uma família

$$\{X_t(\xi) : t \in \mathbb{T}, \xi \in \Omega\}$$

onde  $\mathbb{T}$  é o conjunto de todos os parâmetros  $t$ .

## 6.1. DEFINIÇÃO DE PROCESSO ESTOCÁSTICO

Um processo estocástico é uma família de variáveis aleatórias conjuntamente distribuídas num mesmo experimento  $(\Omega, \mathcal{F}, P)$ , dada por  $\{X_t(\xi) : t \in T, \xi \in \Omega\}$ .

tempo  $X(\cdot, \cdot)$  definido em  $T \times \Omega \rightarrow \mathbb{R}$  é um processo estocástico se todo o conjunto A:

$$A = \{\omega \in \Omega, t \in T : X(t, \omega) \leq z\} \in \sigma \mathcal{B}(T\text{-elge-}brz)$$

Um processo estocástico nas condições acima será notado indistintamente por  $X_t(\xi)$  ou  $X(\xi, t)$  ou  $X(t, \xi)$  ou, se não houver risco de confusão por  $X(t)$ .

Dependendo de como se comportam  $t$  e  $\xi$ , um processo aleatório  $X(t, \xi)$  pode representar 4 coisas distintas:

a) Para um resultado específico  $\xi_i$  (amostra), o que são  $X(t, \xi_i)$  representa uma função simples do tempo. Logo, para  $\xi$  fixo e variável,  $X(t, \xi)$  é uma função no espaço de tempo;

b) Porc. um instante específico  $t_i$ ,  $X(t_i, \xi)$  é uma quantidade dependente de  $\xi$ , portanto, uma variável aleatória;

c)  $X(t, \xi)$  com  $t$  e  $\xi$  variáveis representa uma função de funções do tempo;

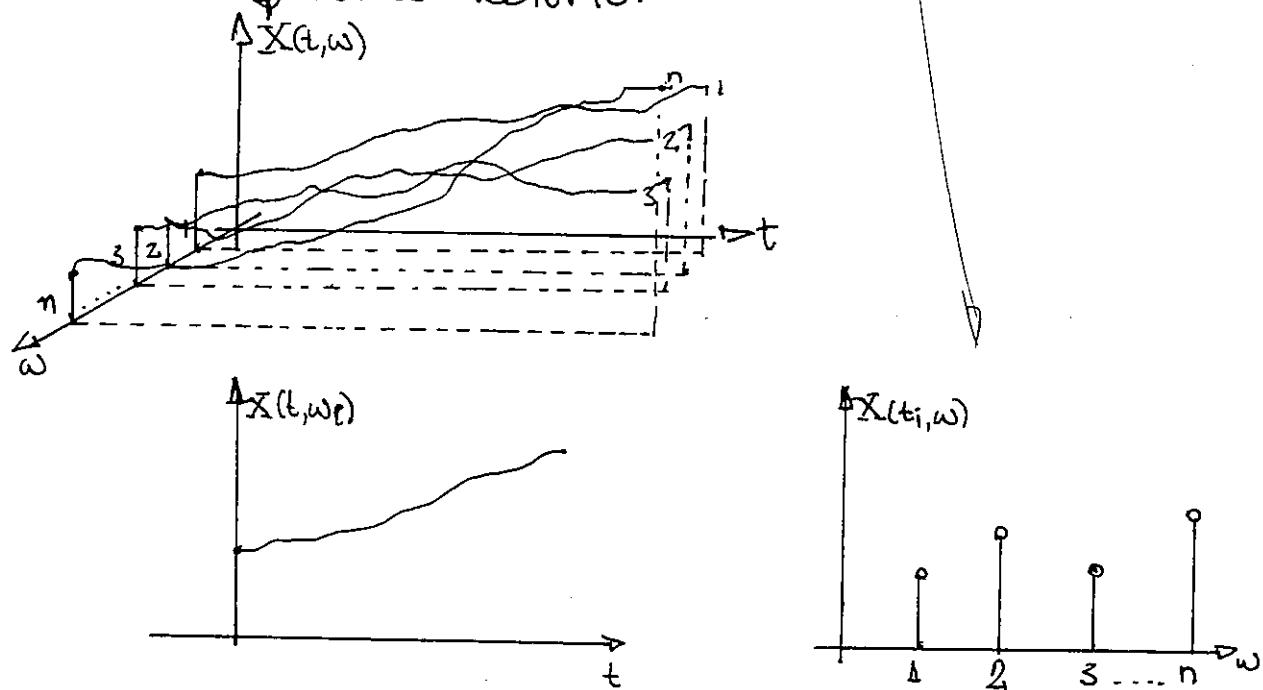
d)  $X(t_i, \xi_i)$  com  $t_i$  e  $\xi_i$  fixos representa um número.

Alguns exemplos de processos aleatórios podem ser desenvolvidos facilmente:

1) Consideremos como nosso espaço amostral o conjunto de todas as crizas nascidas num determinado dia. As associamos à cada crizal a sua altura respetiva como variável de interesse e começamos a contar o tempo a partir da data do nascimento.

A altura de cada criança varia no tempo segundo uma lei que pode ser estabelecida se for feito um acompanhamento. logicamente, a lei de evolução temporal varia de criança para criança, isto é, no processo eleatório, cada amostra segue uma certa função do tempo. Por outro lado, fixando um determinado instante do tempo, as alturas das crianças formam um fenômeno que pode ser descrito por uma variável aleatória, já que essas alturas dependem das amostras que estão sendo consideradas, sendo governada pelo acaso.

Os gráficos abaixo representam uma tentativa de visualizar tal processo eleatório.



Note que, nesse exemplo, o espaço amostral é discreto já que o número de crianças nascidas num dia é um conjunto enumerável e que o espaço de parâmetros,  $\Omega$ , é contínuo.

2) Um processo eleatório extremamente simples pode ser obtido a partir do conhecimento de umas moedas. Define-se, então:

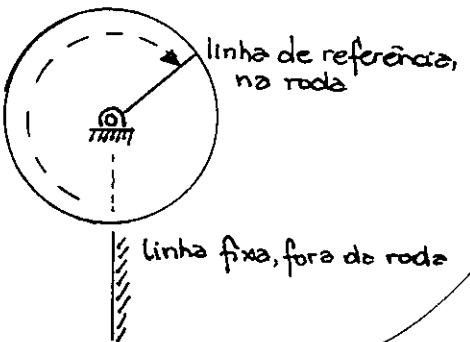
$$\begin{cases} \bar{X}(t, \omega) = \sin t & \text{se } \omega = \cos \alpha \\ \bar{X}(t, \omega) = 2t & \text{se } \omega = \cos \beta \end{cases}$$

$\bar{X}(t, \omega)$  como acima tem, como o exemplo anterior, este é um processo discreto e espaço de parâmetros contínuos. Tal processo é constituído por 2 curvas bastante regulares, o que mostra que funções complicadíssimas não são requisito para um processo estocástico.

3) Consideremos um satélite em órbita periódica em torno da Terra. Nem todas as forças atuantes sobre o satélite podem ser modeladas. A altitude do satélite ao longo do tempo,  $h_t$ , é um processo estocástico onde a variável aleatória é contínua e o parâmetro  $t$  é contínuo.

Definimos, então, um processo estocástico  $\{h(t, \zeta) = h_t(\zeta), t \in [0, T], \zeta \in \Omega\}$  onde  $T$  é o período da órbita.

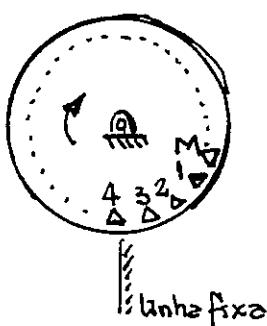
4) Tomamos agora o exemplo de um tombola.



Num instante  $t_i$ , gira-se uma roda balançada e deixa-se cair pelo chão. Um resultado típico é a distância  $s$ , medida sobre o perímetro, no sentido horário, entre a linha de referência na roda e a linha fixa, fora da roda.

Forma-se, então, um processo  $S(t_i, \zeta) = S(t_i)$  onde  $t_i$  é um parâmetro discreto e  $S$  uma v.z. contínua.

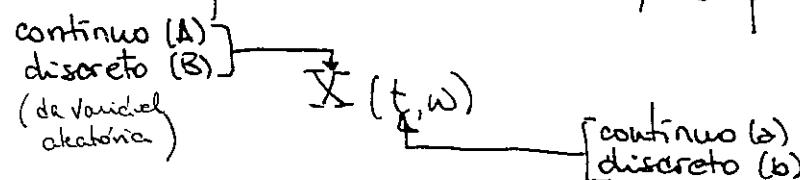
5) Como último exemplo, consideremos novamente o exemplo da tombola, apenas um pouco modificado. Agora, divide-se a roda em  $M$  intervalos igualmente espaçados e repete-se o procedimento anterior de gerar a roda em um instante  $t_i$  e deixá-la parar por efeitos de atrito.



Gerar-se, então, um processo  $S(t_i, z_i)$  =  $S(t_i)$  onde  
 $\{S(t_i, z_i) \mid t_i \in \{t_i, i=1, 2, \dots\}, z_i \text{ inteiro,}$   
 $z_i = \{1, 2, \dots, M\}\}$

Nesse caso,  $t_i$  e  $z_i$  são discretos.

Os exemplos dados acima permitem fazer a classificação dos processos estocásticos segundo a forma de variação de seus parâmetros. Assim, o processo:



pode ser:

- \* Caso (A,a) - v.z. contínua, parâmetro contínuo (ex.3)  
**PROCESSO ESTOCÁSTICO**
- \*\* Caso (A,b) - v.z. contínua, parâmetro discreto (ex.4)  
**SEQUÊNCIA ESTOCÁSTICA OU PROCESSO ESTOCÁSTICO A PARÂMETRO DISCRETO**
- \*\*\* Caso (B,b) - v.z. discreta, parâmetro discreto (ex.5)  
**CADEIA ESTOCÁSTICA**

\*\*\*\* Caso (3,2) - v.z. discreto, parâmetro contínuo (ex. 1e)

### PROCESSO ESTOCÁSTICO DISCRETO

Para aplicações no filtro de Kalman nosso interesse se resume aos 3 primeiros casos, processos, sequências ou cadeias estocásticas. O quarto caso é de pouco interesse para a obtenção do filtro mas de grande aplicação em determinados campos do controle, como controle estocástico de filhos.

## 6.2. CARACTERIZAÇÃO PROBABILÍSTICA DE UM PROCESSO ESTOCÁSTICO.

Seja  $\bar{X}_t(z) = \{x_t, t \in T\}$  um processo estocástico, de o conjunto de parâmetros  $T$  pode ser discreto ou contínuo. Para qualquer conjunto finito  $\{t_1, t_2, \dots, t_n\} \triangleq \{t_i\} \in T$  a função distribuição conjunta do vetor de variáveis aleatórias  $x_{t_1}, x_{t_2}, \dots, x_{t_n}$  é chamada Função Distribuição de Dimensão Finita.

Afirm-se, então, que um processo estocástico pode ser caracterizado probabilisticamente, ou seja, é estatisticamente determinado, pelas especificações da distribuição de dimensão finita

$$F(x_{t_1}, \dots, x_{t_n}) = F(x_{t_1}(w), \dots, x_{t_n}(w)) = F_{x_{t_1}, \dots, x_{t_n}}(x_{t_1}, \dots, x_{t_n})$$

para todos os conjuntos finitos  $\{t_i\} \in T$ .

Isto significa que tendo todas as funções distribuição acima pode-se responder todas as questões probabilísticas sobre o processo. Usaremos daqui para a frente a notação

$$F_{x_{t_1}, \dots, x_{t_n}}(\xi_1, \dots, \xi_n) \triangleq P(\{\omega: \bar{X}(t_i, \omega) \leq \xi_1, \dots, \bar{X}(t_n, \omega) \leq \xi_n\})$$

para designar a função distribuição conjunta.

Um processo estocástico pode ser também car-

terizado pela especificação da função densidade conjunta

$$p(x_{t_1}, \dots, x_{t_N}) = p_{x_{t_1} \dots x_{t_N}}(\xi_1, \dots, \xi_N) = \frac{\partial^N F_{x_{t_1} \dots x_{t_N}}(\xi_1, \dots, \xi_N)}{\partial \xi_1 \dots \partial \xi_{t_1} \dots \partial \xi_{t_N} \dots \partial \xi_N}$$

↗ conjunto de tempos  
 ↗ conjunto de variáveis aleatórias

ou pela especificação da função característica conjunta

$$\varphi_{x_{t_1} \dots x_{t_N}}(u_1, \dots, u_N)$$

para todos os conjuntos finitos  $\{t_i\} \in \mathbb{T}$ .

Note, no caso da função densidade conjunta, o efeito vetorial do processo  $X_{\mathbb{T}}(\omega)$ .

Tendo qualquer uma das três funções acima para todos os conjuntos de parâmetros  $\{t_i\} \in \mathbb{T}$  é conhecida a lei de probabilidade do processo estocástico.

Embora sójam suficientes para fornecer a lei de probabilidades do processo, as densidades  $p_{x_t}(\xi_t) = p(\xi_t)$  e  $p_{x_t x_\sigma}(\xi_t, \xi_\sigma)$  para  $t, \sigma \in \mathbb{T}$  desempenham papel muito importante na teoria dos processos estocásticos. Observe que se tivermos a lei de probabilidades do processo, por exemplo dada por  $p(x_{t_1}, \dots, x_{t_N})$  as duas densidades acima podem ser obtidas como densidades marginais.

$p_{x_t}(\xi_t)$  é dita densidade de primeira ordem do processo enquanto  $p_{x_t x_\sigma}(\xi_t, \xi_\sigma)$  é dita densidade de segunda ordem. Caracterizadas essas duas funções podemos definir uma função densidade condicional através de:

$$p_{x_t | x_\sigma}(\xi_t | \xi_\sigma) = p(x_t | x_\sigma) = \frac{p_{x_t x_\sigma}(\xi_t, \xi_\sigma)}{p_{x_\sigma}(\xi_\sigma)}$$

A importância das densidades de probabilidade reside no fato de que para duas importantsíssimas classes de processos estocásticos, os processos Gaussianos e os processos Markovianos, elas respondem todas as questões probabilísticas, ou seja, caracterizam totalmente esses dois tipos de processos.

### 6.3. MOMENTOS DE UM PROCESSO ESTOCÁSTICO

Seja  $\{x_t, t \in T\}$  um processo escalar o parâmetros contínuos ou discretos.

Def.: A função valor médio ou função esperança do processo  $X_t(z)$  é dada por

$$\mu_x(t) \triangleq E[X_t] = \int_{-\infty}^{\infty} z p_{X_t}(z, t) dz$$

que é, claramente, uma função do parâmetro  $t$ .

Def.: A função de autocorrelação ou função de correlação, simplesmente, do processo escalar  $X_t(z)$  é dada pelo momento de segunda ordem definido por:

$$f_x(t, \tau) \triangleq E[X_t X_{t+\tau}] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_t z_{t+\tau} p_{X_t X_{t+\tau}}(z_t, z_{t+\tau}, t, \tau) dz_t dz_{t+\tau}$$

A função de autocorrelação é função de  $t$  e de  $\tau$ .

Def.: A função de auto covariância ou função de covariância, simplesmente, do processo escalar  $X_t(z)$  é dada pelo momento de segunda ordem definido por:

$$C_x(t, \tau) \triangleq E\left[\bar{X}_t - \mu_{X(t)}\right]\left[\bar{X}_{\tau} - \mu_{X(\tau)}\right] =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\bar{z}_t - \mu_{X(t)}) (\bar{z}_{\tau} - \mu_{X(\tau)}) p_{\bar{X}_t \bar{X}_{\tau}}(\bar{z}_t, \bar{z}_{\tau}) d\bar{z}_t d\bar{z}_{\tau}$$

que, novamente, é função de  $t$  e de  $\tau$ .

Das duas últimas definições demonstra-se que:

$$C_x(t, \tau) = E\left[\bar{X}_t \bar{X}_{\tau} - \mu_{X(t)} \bar{X}_t - \mu_{X(\tau)} \bar{X}_{\tau} + \mu_{X(t)} \mu_{X(\tau)}\right]$$

$$= E[\bar{X}_t \bar{X}_{\tau}] - \mu_{X(\tau)} E[\bar{X}_t] - \mu_{X(t)} E[\bar{X}_{\tau}] + \mu_{X(t)} \mu_{X(\tau)}$$

$$C_x(t, \tau) = \delta_x(t, \tau) - \mu_{X(\tau)} \mu_{X(t)}. \leftarrow$$

\* Exemplo:

Considere o processo estocástico  $\{X_t, t \geq 0\}$ , definido por uma equação do tipo:

$$X_t = a + bt$$

onde  $a$  e  $b$  são v.r. de distribuições conhecidas.

Este processo tem função esperanças:

$$\mu_X(t) = E[\bar{X}_t] = E[a + bt]$$

$$\mu_X(t) = E[a] + E[b].t$$

Sua função de correlação é dada por:

$$\delta_X(t, \tau) = E[(a+bt)(a+b\tau)] = E[a^2 + ab(t+\tau) + b^2 t\tau]$$

$$\delta_X(t, \tau) = E[a^2] + E[ab](t+\tau) + E[b^2]t\tau$$

Finalmente, sua função de covariância é obtida

através de:

$$\begin{aligned}
 C_x(t, \tau) &= E[(X_t - \mu_{X(t)})(X_\tau - \mu_{X(\tau)})] = \delta_x(t, \tau) - \mu_{X(t)}\mu_{X(\tau)} \\
 &= E[z^2] + E[zb](t+\tau) + E[b^2]\tau - (E[z] + E[b])t(E[z] + E[b]) \\
 &= E[z^2] + E[zb](t+\tau) + E[b^2]\tau - E^2[z] - E[z]E[b](t+\tau) \\
 &\quad + E^2[b]t\tau \\
 C_x(t, \tau) &= \{E[z^2] - E^2[z]\} + \{E[zb] - E[z]E[b]\}(t+\tau) + \\
 &\quad + \{E[b^2] - E^2[b]\}t\tau \\
 C_x(t, \tau) &= \text{Var}[z] + \text{cov}[zb](t+\tau) + \text{Var}[b]t\tau
 \end{aligned}$$

Para fazer a extensão para o caso vetorial, tomemos, agora, um processo  $\{X_t, t \in \mathbb{T}\}$  vetorial, de dimensão  $n$ . Temos:

Def. Função Vetorial Valor Médio ou Esperança do processo  $X_t(z)$  é dada por

$$\mu_X(t) = E[X(t)] , \quad \forall t \in \mathbb{T}.$$

Def.: A Matriz de Covariâncias do processo  $X_t(z)$  é um momento de segunda ordem, função de um único parâmetro, definida por:

$$P_{XX}(t) = E[\{X(t) - \mu_X(t)\}\{X(t) - \mu_X(t)\}^T] , \quad \forall t \in \mathbb{T}$$

A noção de matriz de covariâncias pode ser generalizada para se puxar informações sobre quanto rapidamente os valores do amostra  $X(t)$  variam com o tempo. Então:

Def.: A Matriz (ou kernel) de Autocovariâncias do processo

so  $\mathbb{X}_t(\zeta)$  é um momento de segunda ordem, central como a matriz de covariâncias, função de dois parâmetros e dada por:

$$P_{xx}(t_1, t_2) = E[\{\mathbb{X}(t_1) - \mu_x(t_1)\} \{\mathbb{X}(t_2) - \mu_x(t_2)\}^T]$$

para qualquer  $t_1, t_2 \in \Gamma$ .

Note que, das duas últimas definições:

$$P_{xx}(t) = P_{xx}(t, t)$$

i.e., a matriz de covariância pode ser definida a partir da matriz de autocovariância com os 2 argumentos no mesmo instante. Not, também, que os termos da diagonal de  $P_{xx}(t_1, t_2)$  são as funções de covariância dos componentes  $(\mathbb{X}_t)_i$ ,  $i=1, \dots, n$  e os termos fora da diagonal são funções de covariância cruzada de  $(\mathbb{X}_t)_i$  e  $(\mathbb{X}_t)_j$ ,  $i \neq j$ ,  $i, j = 1, \dots, n$ .

→ Os momentos não centrais de 2ª ordem análogos aos definidos acima são a matriz de correlação e a matriz ou kernel de autocorrelação:

Def.: Matriz de correlação do processo  $\mathbb{X}_t(\zeta)$  é dada por:

$$\Psi_{xx}(t) \stackrel{\Delta}{=} E[\bar{\mathbb{X}}(t) \bar{\mathbb{X}}(t)^T]$$

Def.: Matriz (ou Kernel) de Autocorrelação do processo vetorial  $\bar{\mathbb{X}}_t(\zeta)$  é:

$$\Psi_{xx}(t_1, t_2) = E[\bar{\mathbb{X}}(t_1) \bar{\mathbb{X}}(t_2)^T]$$

Claramente  $\Psi_{xx}(t) = \Psi_{xx}(t, t)$  e os termos da diagonal de  $\Psi_{xx}(t_1, t_2)$  são funções de correlação dos componentes  $(\mathbb{X}_t)_i$ ,  $i=1, 2, \dots, n$ , enquanto os termos fora da diagonal são funções de cor-

relação cruzada das componentes  $(\bar{X}_t)_i$  e  $(\bar{X}_t)_j$ ,  $i \neq j$ ,  $i, j = 1, \dots, n$ .

É possível também demonstrar que

$$\Psi_{xx}(t_1, t_2) = P_{xx}(t_1, t_2) + \mu_x(t_1) \mu_x^T(t_2)$$

Portanto, é possível caracterizar a interrelação entre dois processos vetoriais  $\bar{X}_t(\eta)$  e  $\bar{Y}_t(\eta)$ , definindo:

- Matriz ou kernel de Covariâncias Cruzadas

$$P_{xy}(t_1, t_2) = E[\{\bar{X}(t_1) - \mu_x(t_1)\} \{\bar{Y}(t_2) - \mu_y(t_2)\}^T]$$

- Matriz de Covariâncias Cruzadas

$$P_{xy}(t) = P_{xy}(t, t)$$

- Kernel ou Matriz de Correlações Cruzadas

$$\Psi_{xy}(t_1, t_2) = E[\bar{X}(t_1) \bar{Y}(t_2)^T]$$

Demonstra-se, também, que:

$$\Psi_{xy}(t_1, t_2) = P_{xy}(t_1, t_2) + \mu_x(t_1) \mu_y^T(t_2)$$

#### 6.4. INDEPENDÊNCIA E CORRELAÇÃO

O conceito de independência é fundamental no estudo de processos estocásticos. Muita atenção deve ser dada ao fato de que, sendo o processo estocástico especificado por dois parâmetros (ou mais), a independência pode ser definida para um único processo ao longo do tempo ou para dois processos para todos os instantes do tempo. Os conceitos obtidos para independência entre processos estocásticos podem ser rapidamente reduzidos para tratar o aspecto de correlação. Como esperamos que

que é claro ao final deste item, independência implica na separação de todos as funções densidade e a correlação implica na separação apenas dos momentos de segunda ordem.

Def.: Um único processo vetorial  $X_t(\zeta) = \bar{X}(t, \zeta)$  é um processo independente no tempo (ou branco), se para qualquer  $t_1, \dots, t_N \in T$ ,  $\bar{X}_{t_1}(\zeta_1), \dots, \bar{X}_{t_N}(\zeta_N)$  é um conjunto de vetores estocásticos independentes, isto é:

= a saída em um instante não influencia a saída no instante posterior

$$P(\{\omega: \bar{X}(t_1, \omega) \leq \xi_1, \dots, \bar{X}(t_N, \omega) \leq \xi_N\}) = \prod_{i=1}^N P(\{\omega: \bar{X}(t_i, \omega) \leq \xi_i\})$$

onde o símbolo  $\prod$  implica produtório. Equivalentemente podemos escrever a condição acima para funções distribuição:

$$F_{\bar{X}_{t_1}, \dots, \bar{X}_{t_N}}(\xi_1, \dots, \xi_N) = \prod_{i=1}^N F_{\bar{X}_{t_i}}(\xi_i) = F_{\bar{X}_{t_1}}(\xi_1) \cdot F_{\bar{X}_{t_2}}(\xi_2) \cdots F_{\bar{X}_{t_N}}(\xi_N)$$

ou, ainda, para as funções densidade:

$$p_{\bar{X}_{t_1}, \dots, \bar{X}_{t_N}}(\xi_1, \dots, \xi_N) = \prod_{i=1}^N p_{\bar{X}_{t_i}}(\xi_i)$$

processos estocásticos  
(funções)

Def.: Dois processos  $\bar{X}_t(\cdot)$  e  $\bar{Y}_t(\cdot)$  são independentes (mutuamente ou um do outro) se para quaisquer  $t_1, \dots, t_N \in T$ , temos:

$$\begin{aligned} & P(\{\omega: \bar{X}(t_1, \omega) \leq \xi_1, \dots, \bar{X}(t_N, \omega) \leq \xi_N; \bar{Y}(t_1, \omega) \leq \eta_1, \dots, \bar{Y}(t_N, \omega) \leq \eta_N\}) = \\ & = P(\{\omega: \bar{X}(t_1, \omega) \leq \xi_1, \dots, \bar{X}(t_N, \omega) \leq \xi_N\}) \underbrace{P(\{\omega: \bar{Y}(t_1, \omega) \leq \eta_1, \dots, \bar{Y}(t_N, \omega) \leq \eta_N\})}_{\text{produto}} \end{aligned}$$

Def.: Um único processo vetorial  $X_t(\zeta) = \bar{X}(t, \zeta)$  é não correlacionado no tempo se para todos os  $t_1, t_2 \in T$ , exceto  $t_1 = t_2$  ocorre:

Definição de processo não-correlacionado

$$\Psi_{xx}(t_1, t_2) = E[X(t_1)X^T(t_2)] = E[X(t_1)] \cdot E[X^T(t_2)]$$

ou seja:

$$P_{xx}(t_1, t_2) = 0$$

jé que:

$$P_{xx}(t_1, t_2) = \Psi_{xx}(t_1, t_2) - \mu_x(t_1) \mu_x^T(t_2)$$

Def.: Dois processos  $X_t(\cdot)$  e  $\bar{Y}_t(\cdot)$  são não correlacionados entre si se para todo  $t_1, t_2 \in T$ , incluindo  $t_1 = t_2$  ocorre

$$\underbrace{\Psi_{xy}(t_1, t_2)}_{\text{correlação}} = E[X(t_1)\bar{Y}^T(t_2)] = E[X(t_1)]E[\bar{Y}^T(t_2)]$$

ou

$$P_{xy}(t_1, t_2) = 0$$

Note que independência, seja no tempo ou entre dois processos, implica sempre em não correlação. A reciprocidade é verdadeira, isto é, a não correlação, no tempo ou entre dois processos, não implica em independência. A exceção da reciprocidade é por conta dos processos gaussianos, ou seja, se o processo (ou os processos) é não correlacionado no tempo (não correlacionados entre si) então o processo é independente no tempo (independentes entre si), desde que sejam também gaussianos. Voltaremos, mais adiante, aos processos Gaussianos.

## 6.5. PROCESSOS ESTOCÁSTICOS ESTACIONÁRIOS

A estacionariedade de um processo estocástico é uma caracterização importante, que permite uma série de simplificações. De mesma forma que independência e correlação, a estacionariedade ou não de um processo estocástico pode ser definida envolvendo toda a função densidade ou apenas as suas meias.

tos de primeira e de segunda ordem. Dessa maneira, diz-se que a estacionaridade tem sentido amplo (wide-sense) se está relacionada apenas com os dois primeiros momentos e que a estacionaridade tem sentido estrito se está relacionada com todas as momentos, de todas as ordens. Define-se:

a)  $\bar{X}(\cdot, \cdot)$  é um processo estocástico estritamente estacionário se para todos os conjuntos  $t_1, \dots, t_N \in \mathbb{T}$  e qualquer  $\tau \in \mathbb{T}$ , supondo  $(t_i + \tau) \in \mathbb{T}$ , a distribuição conjunta de  $X(t_1 + \tau), \dots, X(t_N + \tau)$  não depende do deslocamento  $\tau$ , i.e.:

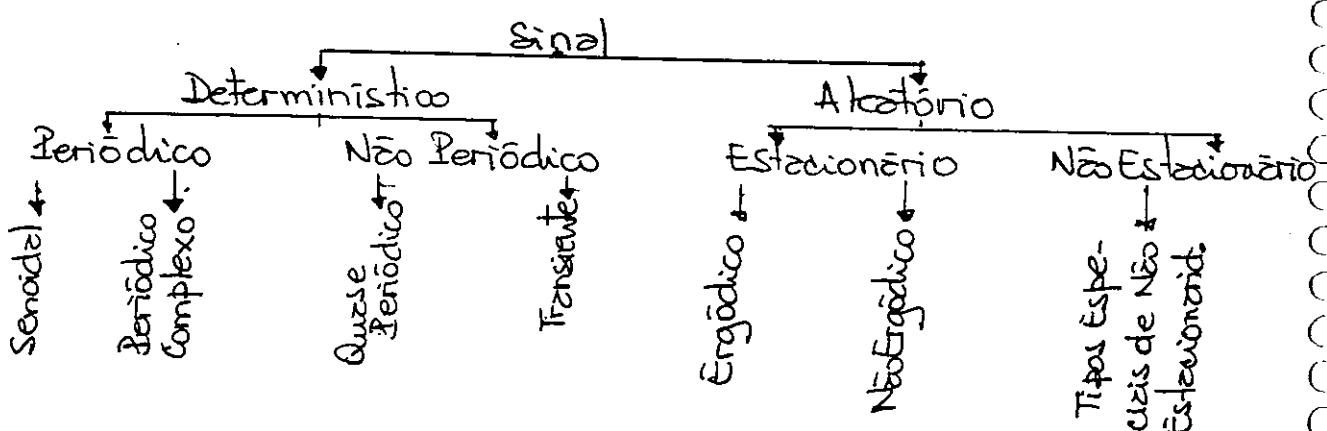
$$\begin{aligned} P(\{\omega : X(t_1 + \tau, \omega) \leq \xi_1, \dots, X(t_N + \tau, \omega) \leq \xi_N\}) &= \\ &= P(\{\omega : X(t_1, \omega) \leq \xi_1, \dots, X(t_N, \omega) \leq \xi_N\}) \end{aligned}$$

b)  $\bar{X}(\cdot, \cdot)$  é um processo estocástico estacionário no sentido amplo se:

- b. 1)  $E[X(t) X(t)]$  é finito;
- b. 2)  $E[\bar{X}(t)]$  é constante;
- b. 3)  $E[\{X(t) - \mu_X\} \{X(t+\tau) - \mu_X\}^T]$  depende apenas de  $\tau$ , diferença de tempo, e, com isso,  $\bar{\Psi}_{XX}(t)$  e  $\bar{P}_{XX}(t) = \bar{P}_{XX}(t, t)$  são constantes.

Das duas definições acima imediatamente se conclui que um processo estritamente estacionário é estacionário no sentido amplo se e somente se tem momentos de 2ª ordem finitos e que a estacionaridade no sentido amplo não implica estacionaridade estrita. A exceção, para essa segunda afirmação, é, novamente, para os processos Gaussianos já que seus dois primeiros momentos são suficientes para caracterizar todo a função densidade. Assim, se um processo Gaussiano for estacionário no sentido amplo, será também estritamente estacionário.

A importância da estacionaridade ou não de um processo estocástico está ligada à análise que deve ser feita para caracterizar o processo. Dessa maneira, a maioria dos analisadores de sinal costuma, em primeiro lugar, tentar separar a parte determinística da aleatoriedade, em seguida, analisando a parte randomica, separar componentes estacionárias e não estacionárias. O esquema abaixo, sugerido por Bendat e Piersol (1971), mostra os caminhos seguidos num procedimento típico.



Os aspectos de estacionaridade podem ser analisados a partir da densidade espectral de potência do processo, definida no item seguinte.

### 6.5. DENSIDADE ESPECTRAL DE POTÊNCIAS

A densidade espectral de potências de um processo aleatório é uma descrição da composição geral das frequências envolvidas no processo em termos da densidade espectral do seu valor médio quadrático. Para gerar as densidades espetrais de potência é necessário o uso das Transformadas de Fourier.

Se  $y(t)$  for uma função escalar e  $\tilde{Y}(w)$  sua transformada de Fourier, então:

$$\bar{y}(\omega) = a \int_{-\infty}^{\infty} y(t) e^{-i\omega t} dt = F[y(t)] - \text{transf. Fourier}$$

$$y(t) = b \int_{-\infty}^{\infty} \bar{y}(\omega) e^{i\omega t} d\omega = \mathcal{F}[\bar{y}(\omega)] - \text{antitransf. Fourier}$$

com  $ab = 1/2\pi$ , podendo ser definidas as transformadas convenientes. São propriedades da Transformação de Fourier:

- $\bar{y}(\omega)$  é, em geral, complexa;
- A parte real de  $\bar{y}(\cdot)$  é função par; a parte imaginária é ímpar em  $\omega$ ;
- Embora  $\delta(t)$ , (delta de Dirac), não seja uma função no sentido comum, define-se convenientemente sua transformada através de:

$$\bar{\delta}(\omega) = a \int_{-\infty}^{\infty} \delta(t) e^{-i\omega t} dt = a$$

$$\delta(t) = b \int_{-\infty}^{\infty} a e^{i\omega t} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} dt$$

média e covariância constantes

Def.: A Densidade Espectral de Potência de um processo escalar estacionário no sentido amplo é a transformada de Fourier das funções de correlação  $\bar{\Phi}_{xx}(\omega) \triangleq \Phi_{xx}(t+\tau)$  (note que  $\Phi_{xx}(\cdot)$  é usado agora para um processo escalar).

Como existem diversas maneiras de definir os coeficientes  $a$  e  $b$  da transformação, usaremos a mais comum:

$$\bar{\Phi}_{xx}(\omega) = \int_{-\infty}^{\infty} \Phi_{xx}(\tau) e^{-i\omega\tau} d\tau , \quad a=1$$

para a densidade espectral. A antitransformada é dada por:

$$\Phi_{xx}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{\Phi}_{xx}(\omega) e^{+i\omega\tau} d\omega ; \quad b = \frac{1}{2\pi}$$

Se lembrarmos que:

$$\Phi_{xx}(0) = \Psi_{xx}(t, t) = E[X(t)^2]$$

$$\omega = 2\pi f$$

então:

$$E[X(t)^2] = \Psi_{xx}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{\Psi}_{xx}(\omega) d\omega = \int_{-\infty}^{\infty} \bar{\Psi}_{xx}(\omega) d\omega$$

$$\bar{\Psi}_{xx}(\omega) = \int_{-\infty}^{\infty} \Psi_{xx}(\tau) [\cos \omega \tau - i \sin \omega \tau] d\tau$$

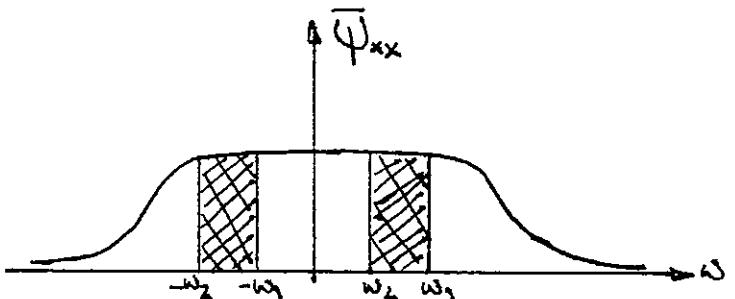
Agora, desde que  $\Psi_{xx}(\tau)$  e  $\cos \omega \tau$  são funções pares de  $\tau$  e  $\sin \omega \tau$  é ímpar em  $\tau$ :

$$\bar{\Psi}_{xx}(\omega) = \int_{-\infty}^{\infty} \Psi_{xx}(\tau) \cos \omega \tau d\tau = 2 \int_0^{\infty} \Psi_{xx}(\tau) \cos \omega \tau d\tau$$

Logo,  $\bar{\Psi}_{xx}$  é uma função real e par de  $\omega$ , como a transformada dada por

$$\Psi_{xx}(\omega) = \frac{1}{\pi} \int_0^{\infty} \bar{\Psi}_{xx}(\omega') \cos \omega' \omega d\omega'$$

O nome densidade espectral de potências é, então, motivado no sentido de valores médios quadráticos estatísticos de um amostrado. A integração de  $\bar{\Psi}_{xx}$  sobre a faixa de frequências  $\omega_1$  até  $\omega_2$  leva ao valor médio quadrático do processo que consiste apenas de aqueles componentes harmônicos de  $X(t)$  que estão entre  $\omega_1$  e  $\omega_2$ , conforme representado abaixo.



Nesse caso, o valor médio quadrático de todo o processo  $X(t)$  é dado através de uma integração de  $\bar{\Psi}_{xx}(\omega)$  sobre todo o domínio das possíveis frequências  $\omega$ . Isso pode ser feito se definirmos a densidade espectral de potência de uma forma equivalente:

$$\bar{\Psi}'_{xx}(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \Psi_{xx}(t) e^{-i\omega t} dt \quad a = 1/\pi$$

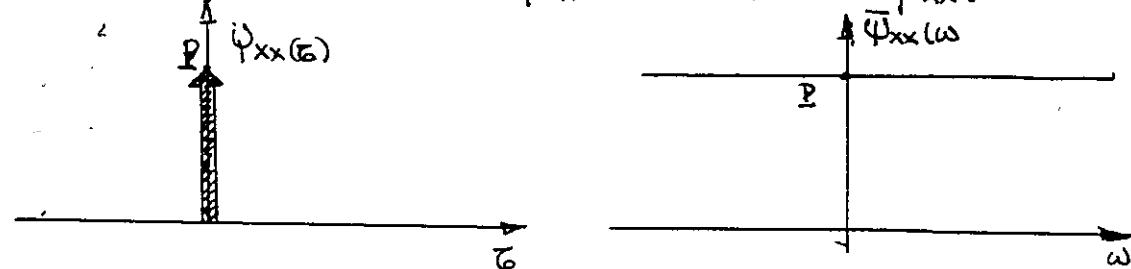
$$\Psi_{xx}(t) = \frac{1}{2} \int_{-\infty}^{\infty} \bar{\Psi}'_{xx}(\omega) e^{i\omega t} d\omega \quad b = 1/2$$

Nesse caso, se mostra que:

$$E[X^2(t)] = \int_0^{\infty} \bar{\Psi}'_{xx}(\omega) d\omega$$

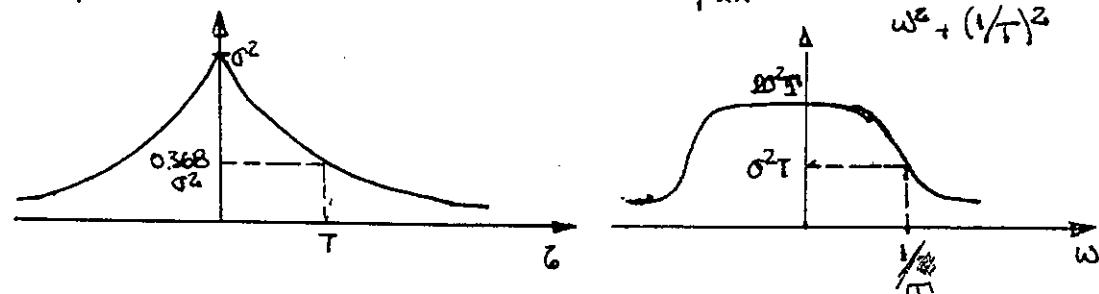
Exemplos:

a) Processo Branco:  $\Psi_{xx}(t) = P \delta(t) \Rightarrow \bar{\Psi}'_{xx}(\omega) = P$



b) Processo Exponencialmente Correlacionado ao tempo com intervalo de correlação 'T':

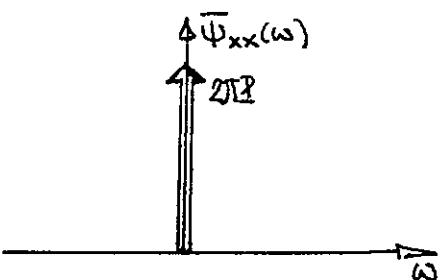
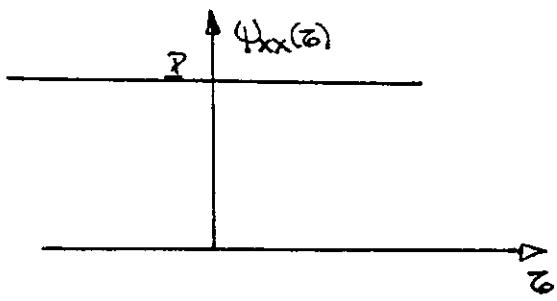
$$\Psi_{xx}(t) = \sigma^2 e^{-|t|/T} \Rightarrow \bar{\Psi}'_{xx}(\omega) = \frac{2\sigma^2/T}{\omega^2 + (1/T)^2}$$



c) "Bis" Atentório (Erro Sistêmico):

$$\Psi_{xx}(t) = P \Rightarrow \bar{\Psi}'_{xx}(\omega) = 2\pi P \delta(\omega)$$

↓  
Impulso



Pode-se definir, agora, a densidade espectral cruzada de potências de 2 processos escalares estacionários no sentido amplio  $\tilde{X}(\cdot, \cdot)$  e  $\tilde{Y}(\cdot, \cdot)$  como sendo a transformada de Fourier da função correlação cruzada

$$\bar{\Psi}_{xy}(w) = \int_{-\infty}^{\infty} \Psi_{xy}(\tau) e^{-i\omega\tau} d\tau$$

De definição acima, poderíamos ter propriedades semelhantes ao caso escalar desenvolvido, e, semelhantemente, estender as definições para casos vetoriais, etc, etc.

## 6.6. PROCESSOS ERGÓDICOS

Processos ergódicos formam uma importante sub-classe dos processos estacionários em sentido amplo.

Um processo é ergódico se qualquer propriedade estatística (média, covariâncias, etc) calculada pela média sobre todo o conjunto de amostras num instante  $t$  fixo pode ser também calculada por uma média no tempo de um único número representativo do conjunto. A hipótese de ergodicidade era fundamental na teoria de filtragem de Wiener, hipótese essa que pode ser abandonada quando se trabalha com filtros de Kalman.

Não todo processo estocástico é ergódico. Um

exemplo clássico usado para provar essa afirmação é o dos processos que envolvem funções constantes: o cálculo da média sobre uma única amostra fornece a constante, enquanto a média sobre todo conjunto num instante t fixo fornece a média das constantes.

• Não há regras gerais para testar a ergodicidade de um processo. No entanto, é possível estabelecer um teste para ergodicidade se o processo for gaussiano. No caso escalar,  $X(t, \cdot)$  é gaussiano em  $\mathbb{T} = (-\infty, +\infty)$ . A condição suficiente para que  $X(\cdot, \cdot)$  como acima, seja <sup>ergódico</sup> gaussiano é que:

$$\int_{-\infty}^{+\infty} |C_{xx}(z)| dz \quad \text{finita}$$

## 6.7. PROCESSOS E SEQUÊNCIAS DE MARKOV

De todo a discussão feita até aqui fica ressaltada a dificuldade de caracterizar probabilisticamente um processo estocástico. Uma grande simplificação poderia ser introduzida se pudermos admitir, que o processo em questão é Markoviano, ou seja, que o nosso modelo em estudo tem propriedades de Markov. Com isso, a caracterização do processo tornar-se-á bem mais fácil.

Def.: Dada uma sequência de variáveis aleatórias  $X(k), k = 0, 1, 2, \dots, N, \dots$  essa sequência é dita Markoviana se:

$$\begin{aligned} & p(x_{k+1}) | x_0, x_1, x_2, \dots, x_k) = p(x_{k+1}) | x_k, \dots, x_0) = \\ & \text{densidade de probabilidade} \\ & = p(x_{k+1}) | x_k) \quad \text{para todo } k. \end{aligned}$$

↳ densidade de probabilidade

A propriedade de Markov estabelece, então, que é feito de

probabilidades do processo no futuro, sendo conhecido no presente, não depende de como o processo chegou ao atual estado (presente). Em outras palavras, num processo de Markov, o presente separa o passado do futuro.

Sob esse aspecto, a propriedade de Markov para processos ou sequências estocásticas é conceitualmente análoga à colocação de estados do sistema por processos determinísticos.

Voltando à sequência estocástica colocada, note-se que:

Teorema de Bayes

$$p(\xi(t+1) | \xi(t), \xi(t-1), \dots, \xi(1), \xi(0)) = \frac{p(\xi(t+1), \xi(t), \dots, \xi(1), \xi(0))}{p(\xi(t), \xi(t-1), \dots, \xi(0))}$$

$$p(\xi(t+1), \xi(t), \dots, \xi(0)) = p(\xi(t+1) | \xi(t), \xi(t-1), \dots, \xi(1), \xi(0)) \cdot p(\xi(t), \xi(t-1), \dots, \xi(0))$$

Pela propriedade markoviana:

$$p(\xi(t+1), \xi(t), \dots, \xi(0)) = p(\xi(t+1) | \xi(t)) \cdot p(\xi(t), \xi(t-1), \dots, \xi(0))$$

Usando o mesmo desenvolvimento para  $\xi(t)$ :

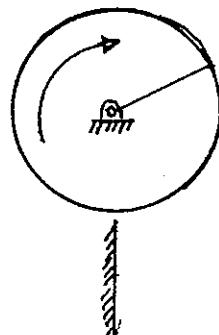
$$p(\xi(t), \xi(t-1), \dots, \xi(0)) = p(\xi(t) | \xi(t-1)) \cdot p(\xi(t-1), \xi(t-2), \dots, \xi(0))$$

substituindo na relação acima e repetindo o procedimento, chegar-se-á:

$$p(\xi(t+1), \xi(t), \dots, \xi(1), \xi(0)) = p(\xi(t+1) | \xi(t)) \cdot p(\xi(t) | \xi(t-1)), \dots, p(\xi(1) | \xi(0)) \cdot p(\xi(0))$$

Exemplo:

O caso da tombola, analisado na seção 6.1., é o exemplo de sequências markovianas. Definindo o resultado no instante  $t_i$  como  $S(t_i, \omega) \triangleq S(t_i)$ , para impulsos totalmente aleatórios so-



brs a roda podemos escrever:

$$p(S(t_k) | S(t_{k-1})) = p(S(t_k)) \quad \text{para todos } k$$

Em particular:

$$p(S(t_k) | S(t_{k-1}), \dots, S(0)) = p(S_k | S_{k-1}) = p(S_k) = p(S(t_k))$$

Como a densidade de probabilidade no instante  $t_k$  não depende das densidades nos instantes anteriores, temos, como já visto, uma sequência branca em puramente aleatória, definida no item 6.4. Note que uma sequência branca é um caso particular de sequência de Markov.

Ainda com referência ao exemplo da tombola, admitindo que ela seja ideal podemos colocar a densidade de probabilidades como sendo uniforme, isto é:

$$p(S_k) = \frac{1}{2\pi R}, \quad \forall k$$

O que nos fornece uma função distribuição na forma

$$F(S_k, \xi) = \int_{-\infty}^{\xi} p(S_k, \xi) d\xi = \begin{cases} \frac{\xi}{2\pi R} & 0 \leq \xi < 2\pi R \\ 0 & \text{no resto} \end{cases}$$

O valor acumulado, no exemplo da tombola, gera uma sequência Markoviana não branca, pois o valor acumulado depende do valor anterior. Sendo  $\bar{Y}(t_k)$  o valor acumulado em:

$$t_0 - S(0) - \bar{Y}(0)$$

$$t_1 - S(1) - \bar{Y}(1), \quad \text{etc}$$

Conclusão:

$$\bar{Y}(k) = \bar{Y}(k-1) + S(k-1) = \sum_{i=0}^{k-1} S(i)$$

com a condição inicial  $X(0) = 0$ .

A sequência  $\{X(t)\}$  é, então, Markoviana, mas não permanece estocástica (branca).

Def.: Um processo estocástico  $X(t)$  é dito Markoviano se para todo  $\forall N$ ,  $t_1 < t_2 < \dots < t_N$ :

$$P\{\bar{X}(t_N) \leq \bar{z}_N | X(t_{N-1}), \dots, X(t_1)\} = P\{\bar{X}(t_N) \leq \bar{z}_N | X(t_N)\}$$

Para um processo Markoviano valem as mesmas conclusões estabelecidas para os processos Markovianos.

Um último e importante fato relacionado com as sequências e processos de Markov decorre da regra:

$$p(z_{(k+1)}, z_{(k)}, \dots, z_{(1)}, z_{(0)}) = p(z_{(k+1)} | z_{(k)}) p(z_{(k)} | z_{(k-1)}) \dots p(z_{(1)} | z_{(0)}) p(z_{(0)})$$

Temos, então, que a lei de probabilidades do processo de Markov pode ser obtida conhecendo-se  $p(z_{(0)})$  e  $p(z(t) | z_{(0)})$  para todo  $t > 0 \in \mathbb{T}$ , conjunto de parâmetros. As densidades condicionais  $p(z(t) | z_{(0)})$  recebem o nome, bastante sugestivo, de densidades de transição de probabilidades do processo de Markov. Essas densidades guardam uma importante analogia com as matrizes de transição dos sistemas determinísticos.

## 6.8. PROCESSOS GAUSSIANOS

Def.: Um processo estocástico  $\{X_t, t \in \mathbb{T}\}$  é um processo gaussiano ou normal se a sua lei de probabilidades obedece à função distribuição  $F(X_{t_1}, \dots, X_{t_N})$  ou pelas suas funções densidade  $f(x_{t_1}, \dots, x_{t_N})$  ou ainda pelas suas funções características  $\phi_{x_{t_1}, \dots, x_{t_N}}(u_1, \dots, u_N)$  for normal.

Por exemplo, se  $X(\cdot, \cdot)$  é gaussiana e as densidades apropriadas existem, então, para qualquer escolha de  $t_1, t_2 \in \mathbb{T}$

$$F_{X_{t_1} X_{t_2}}(z) = [(2\pi)^n |P|^{1/2}]^{-1} \exp \left\{ -\frac{1}{2} (z - \mu)^T P^{-1} (z - \mu) \right\}$$

onde

$$\mu = \begin{bmatrix} \mu_X(t_1) \\ \mu_X(t_2) \end{bmatrix} = \begin{bmatrix} E[X(t_1)] \\ E[X(t_2)] \end{bmatrix}$$

$$P = \begin{bmatrix} E[X_{t_1} X_{t_1}^T] - \mu_{X(t_1)} \mu_{X(t_1)}^T & E[X_{t_1} X_{t_2}^T] - \mu_{X(t_1)} \mu_{X(t_2)}^T \\ E[X_{t_2} X_{t_1}^T] - \mu_{X(t_2)} \mu_{X(t_1)}^T & E[X_{t_2} X_{t_2}^T] - \mu_{X(t_2)} \mu_{X(t_2)}^T \end{bmatrix}$$

$$P = \begin{bmatrix} E[X_{t_1} X_{t_1}^T] & E[X_{t_1} X_{t_2}^T] \\ E[X_{t_2} X_{t_1}^T] & E[X_{t_2} X_{t_2}^T] \end{bmatrix} - \mu \mu^T$$

Exemplo:

Consideremos o processo escalar

$$X_t = a + bt$$

e imaginemos  $a$  e  $b$  como v.a. gaussianas, não correlacionadas de médias nulas e varianças unitárias.

E importante observar que  $X_t = a + bt$  é uma combinação linear de v.a.s gaussianas e, com isso,  $X_t$  é também gaussiana. Para o caso do exemplo, já havíamos calculado na seção 6.3:

$$\mu_X(t) = E[a] + E[b]t$$

o que nos deu:

$$\mu_X(t) = 0 \quad \forall t \in \mathbb{T}$$

de mesma forma, da seção 6.3:

$$C_X(t, \bar{t}) = \text{var}[a] + \text{cov}[ab](t + \bar{t}) + \text{var}[b]t\bar{t}$$

e, sendo  $a$  e  $b$  não correlacionados, de varianças unitárias:

autocorrelação

$$\text{cov}[zb] = 0 \quad (\text{não correlacionadas})$$

$$C_x(t_1, t_2) = 1 + t_2$$

Para quaisquer  $t_1, t_2$ , teremos, então:

$$\mu = \begin{bmatrix} \mu_x(t_1) \\ \mu_x(t_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$P = \begin{bmatrix} C_x(t_1, t_1) & C_x(t_1, t_2) \\ C_x(t_2, t_1) & C_x(t_2, t_2) \end{bmatrix} = \mu \mu^T$$

$$P = \begin{bmatrix} 1 + t_1^2 & 1 + t_1 t_2 \\ 1 + t_1 t_2 & 1 + t_2^2 \end{bmatrix}$$

### 6.9. RUIDO BRANCO

No modelagem sistemática de qualquer processo em dispositivo físico, é comum usar-se um método iterativo até atingir o ponto onde não adianta mais tentar refinar o modelo, não ocorrerão mudanças significativas. As causas para esse fato são flutuações imprevisíveis para as quais não existem, em princípio, teléscopos disponíveis ou são erros aleatórios nos instrumentos de medida que não se consegue detectar. Nesse caso, dizemos que a modelagem atingiu o "nível de ruído".

Um bom modelo para incluir esses efeitos no todo do fenômeno é dado pelo chamado ruído branco. Embora já tenhamos feito menção a esse tipo de ruído ou processo no texto, vamos analisá-lo agora com um pouco mais de detalhe.

Def.: Uma sequência aleatória branca  $\{x_n, n=1, 2, \dots\}$

é uma sequência de Markov para a qual

$$P(X_k | X_l) = f(X_k) \quad k > l$$

ié, todos os  $X_k$  são mutuamente independentes e conhecer uma realização de  $X_l$  não ajuda em nada para prever como será  $X_k$ .  
Uma sequência branca é, portanto, completamente aleatória automaticamente imprévivel.

Se os  $X_k$  são todos normalmente distribuídos, a sequência  $\{X_k\}$  é dita sequência branca Gaussiana. O grande interesse nesse tipo de sequência está baseado no Teorema do Limite Central que mostra que o ruído devido à superposição de um grande número de pequenos efeitos aleatórios independentes pode ser aproximado de maneira excelente por uma distribuição Gaussiana, ~~é~~ desde que suas distribuições de probabilidades dos efeitos individuamente. Papoulis (1965) mostra um exemplo em que porzenas 3 efeitos superpostos a aproximação por distribuições Gaussianas já é muito boa.

Uma sequência branca Gaussiana fica totalmente caracterizada, em sua lei de probabilidades, pelo média

$$\mu_{X_N} = E[X_N] \quad \text{todo } N \geq 1$$

e pela matriz de covariâncias

$$R_{xx}(t_N, t_M) = Q_N S_{MN}$$

onde  $S_{MN}$  é o delta de Kronecker

$$S_{MN} = \begin{cases} 1 & M=N \\ 0 & M \neq N \end{cases}$$

e  $Q_N$  é uma matriz semi-definida positiva.

A extensão para o caso contínuo, isto é, para um processo branco é feita, formalmente, por cálculo, através de

$$\left. \begin{array}{l} \text{Processo} \\ \text{Branco} \end{array} \right\} \begin{aligned} \mu_{\bar{x}}(t) &= E[\bar{x}(t)] \quad t \in T \\ e \quad P_{xx}(t, \tau) &= t \left[ (\bar{x}_t - \mu_{\bar{x}}(t)) (\bar{x}_{\tau} - \mu_{\bar{x}}(\tau))^T \right] = Q(t) \delta(t-\tau) \end{aligned}$$

onde  $Q(t)$  é também uma matriz semi-definida positiva e  $\delta(t-\tau)$  é a função delta de Dirac, uma função não definida no sentido clássico.

Claramente, esse extensão é formal pois, como já visto, um processo branco exigiria potências infinitas para sua realização.

#### 6.10. REPRESENTAÇÃO PARA EQUAÇÕES DIFERENCIAIS ESTOCÁSTICAS

Num problema típico de Teoria de Estimação generalizadas de estudar a evolução de um sistema dinâmico linear na forma:

$$\dot{\bar{x}}(t) = f(t)x(t) + B(t)u(t) + G(t)w(t) \quad (*)$$

onde  $x(t)$  é o nosso vetor de estado;  $u(t)$  é uma entrada determinística ou controle a ser aplicada ao sistema;  $w(t)$  é um ruído eletrônico adicionado à dinâmica de movimento para levar em conta a nossa incerteza sobre o modelo e  $f(t), B(t), G(t)$  são matrizes de funções não eletrônicas com as dimensões apropriadas.

As informações disponíveis sobre o sistema são dadas por

$$z(t_i) = H(t_i)x(t_i) + v(t_i) \quad (**)$$

onde  $z(t_i)$  é um vetor de medidas;  $v(t_i)$  representa o ruído eletrônico nas medidas e  $H(t_i)$  é uma matriz de funções não fixas.

Note que o modelo traduzido pelas relações (\*)

e (\*\*), se, a menos das componentes de ruído, exatamente as mesmas usadas no caso do observador determinístico desenvolvido no Capítulo 3. Por pura analogia com aquele item, poderíamos tentar escrever a solução de (\*) na forma:

$$x(t) = \underline{\Phi}(t, t_0) x(t_0) + \int_{t_0}^t \underline{\Phi}(t, \tau) B(\tau) u(\tau) d\tau + \int_{t_0}^t \underline{\Phi}(t, \tau) G(\tau) w(\tau) d\tau$$

onde  $\underline{\Phi}(t, t_0)$  é a matriz de transição de estados do nosso sistema linear. A pergunta crucial é se é correto escrever a solução do sistema dinâmico nesse formato e quais as hipóteses que devemos colocar para que isso aconteça.

Para a primeira parte da pergunta a resposta é que a representação acima não é correta e não ser para alguns tipos muito especiais de ruído. Em grande parte dos problemas de Engenharia gostaríamos de colocar  $w(t, \cdot)$  acima como um ruído branco e, para esse tipo de ruído, a representação acima é certamente incorreta já que qualquer integral envolvendo ruído branco seria infinito.

No entanto, poderíamos usar formalmente a solução proposta se entendermos o último termo à direita como sendo uma representação de  $\int \underline{\Phi}(t, \tau) G(\tau) d\beta(\tau)$  onde  $d\beta(\tau)$  é uma diferença dividida envolvendo processo ou movimento Browniano. Essa afirmação significa que se colocarmos a expressão (\*) na forma:

$$dx_t = F(t) x(t) dt + B(t) u(t) dt + G(t) d\beta(t) \quad (***)$$

onde  $d\beta(t)$  é um processo Browniano, então a expressão acima é correta e fornecerá uma solução:

$$x(t) = \underline{\Phi}(t, t_0) x(t_0) + \int_{t_0}^t \underline{\Phi}(t, \tau) B(\tau) u(\tau) d\tau + \int_{t_0}^t \underline{\Phi}(t, \tau) G(\tau) d\beta(\tau)$$

É possível notar, então, que a expressão (\*) estaria correta se o processo branco pudesse ser definido como a derivada em relação ao tempo do processo Browniano. Como se é conhecido mais adiante, o problema é que movimento Browniano é não-diferenciável e isso invalida a representação (\*) que, a partir de agora, terá perda丧失 o sentido de uma forma de colocação simplificada da equação (\*\*).

→ Para conceituarmos movimento Browniano, precisamos antes conceituar o que é um processo a incrementos independentes.

Def.: Seja  $t_0 < t_1 < t_2 < \dots < t_N$  uma partição do intervalo ou conjunto de parâmetros  $\mathbb{T}$ . Se os incrementos do processo  $\mathbb{X}(\cdot, \cdot)$ , ou seja, o conjunto de  $N$  variáveis aleatórias

$$\delta_1(\cdot) = [\mathbb{X}(t_1, \cdot) - \mathbb{X}(t_0, \cdot)]$$

$$\delta_2(\cdot) = [\mathbb{X}(t_2, \cdot) - \mathbb{X}(t_1, \cdot)]$$

:

$$\delta_N(\cdot) = [\mathbb{X}(t_N, \cdot) - \mathbb{X}(t_{N-1}, \cdot)]$$

é tal que os  ~~$\delta_i(\cdot)$~~   $\delta_i(\cdot)$ ,  $i_1 = 1, 2, \dots, N$  são mutuamente independentes para qualquer partição de  $\mathbb{T}$ , então,  $\mathbb{X}(\cdot, \cdot)$  é um processo de incrementos independentes.

Def.:  $\beta(\cdot, \cdot)$  é um processo (ou movimento) Browniano no escalar de difusão constante se:

a) é um processo de incrementos independentes;

b) os incrementos são v.z. Gaussianas de forma

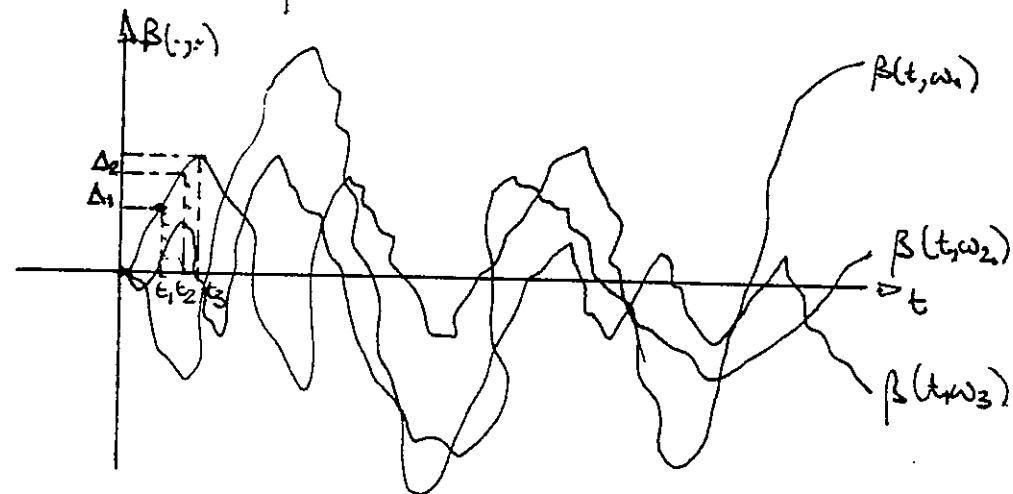
que para  $t_1$  e  $t_2$  instantes quaisquer em  $\mathbb{T}$ :

$$E[\beta(t_2) - \beta(t_1)] = 0$$

$$E[(\beta(t_2) - \beta(t_1))^2] = q |t_2 - t_1|$$

$$c) \beta(t_0, \omega_i) = 0 \text{ para todo } \omega_i \in \Omega$$

Este tipo de processo serve para descrever, por exemplo, o movimento das moléculas de um gás (Teoria Quântica dos Gases) e tem uma representação abaixo.



Da figura:

$$\Delta_1 = \delta_1(\omega_1) = [\beta(t_2, \omega_1) - \beta(t_1, \omega_1)]$$

$$\Delta_2 = \delta_2(\omega_1) = [\beta(t_3, \omega_1) - \beta(t_2, \omega_1)]$$

e, sendo o processo um movimento Browniano,  $\Delta_1$  e  $\Delta_2$  são independentes.

A constante  $q$ , usada na definição, é dita difusão do processo Browniano e serve como um indicador de quão rápido o valor médio quadrático de  $\beta(\cdot, \cdot)$  diverge de seu valor inicial, zero, no instante  $t_0$ . Este fato decorre da afirmação abaixo que é a chave para o entendimento do processo Browniano.

Desde que  $\beta(t_i, \cdot)$  para um dado  $t_i \in \mathbb{T}$  é uma v.a. composta pelos somos de incrementos Gaussianos independentes,  $\beta(t_i, \cdot)$  é também Gaussiano com estatística:

$$\mu_{\beta}(t_i) = E[\beta(t_i)] = 0$$

$$\mathbb{C}_{\beta\beta} = P_{\beta\beta}(t_i) = E[\beta^2(t_i)] = q(t_i - t_0)$$

Esse resultado permite, então, gerar, para o movimento Browniano, as funções de densidade conjunta para qualquer conjunto finito de v.z.s  $\beta(t_0), \dots, \beta(t_N)$ . Essas funções densidade conjunta serão também Gaussianas, de médias nulas.

Para mostrar a continuidade de um processo Browniano é antes necessário definir o conceito de convergência. A convergência de um processo estocástico pode ser colocada em termos de 3 ebórdagens: convergência em média quadrática (l.i.m.), convergência com probabilidade 1 (w.p.1) ou quase certa e convergência em probabilidade.

É possível demonstrar que um processo Browniano é contínuo em qualquer das três conceitos de convergência. No entanto, demonstra-se que movimento Browniano é não diferenciável em qualquer dos três conceitos.

Isso nos leva a associar, incorretamente, o ruído branco Gaussiano escalar como a derivada de um movimento Browniano escalar, i.e., se  $\beta(\cdot, \cdot)$  fosse diferenciável, definirímos um processo branco  $w(\cdot, \cdot)$ , gaussiano tal que:

$$\beta(t, \cdot) = \int_{t_0}^t w(\tau, \cdot) d\tau$$

$$\text{ou } w(t, \cdot) = \frac{d\beta(t, \cdot)}{dt}$$

e esse processo branco teria como propriedades estatísticas

$$E[w(t)] = 0 \quad \text{para todo } t \in \mathbb{T}$$

$$E[w(t)w(\tau)] = q \delta(t-\tau)$$

A discussão feita até aqui pode ser generalizada para movimentos Brownianos escalares com difusão variante no tempo, bastando, para isso, uma redefinição ao item b

para:

$$E[(\beta(t_2) - \beta(t_1))^2] = \int_{t_1}^{t_2} q(t) dt \quad , \quad t_2 > t_1$$

com  $q(t) \geq 0$  para todo  $t \in \mathbb{I}$ .

O correspondente ruído branco Gaussiano não estacionário (incorreto), derivado do movimento Browniano clássico é dado por:

$$E[w(t)] = 0$$

$$E[w(t) w(t')] = q(t) \delta(t-t') \quad \text{para todo } t, t' \in \mathbb{I}.$$

A extensão para o caso vetorial é também direta. Um processo vetorial de movimento Browniano, de  $n$  componentes,  $\beta(\cdot, \cdot)$  é um processo de incrementos Gaussianos indeterminados com

$$E[\beta(t)] = 0$$

$$E[(\beta(t_2) - \beta(t_1))(\beta(t_2) - \beta(t_1))^T] = \int_{t_1}^{t_2} Q(t) dt$$

para  $t_2 \geq t_1$ , com  $Q$ , matriz  $n \times n$ , simétrica, semidefinita positiva para todo  $t \in \mathbb{I}$  e, ao mínimo, contínua por partes.

Por analogia, define-se o ruído branco vetorial Gaussiano formalmente através de:

$$E[w(t)] = 0$$

$$E[w(t) w^T(t')] = Q(t) \delta(t-t') \quad , \quad t, t' \in \mathbb{I}$$

com a matriz  $Q$  nas mesmas hipóteses da definição anterior.

Considerando todas as considerações sobre movimento Browniano e processos brancos, resta dar sentido à integral  $\int_t \Phi(t,\tau) G(\tau) d\beta(\tau)$  que aparece em (\*\*\*) para a diferenciação estocásticas para, com isso, gerar soluções para as equações diferenciais estocásticas.

Se  $a(\cdot)$  é uma função conhecida do tempo, contínua por partes, e  $\beta(\cdot, \cdot)$  é um processo Browniano escalar de difusão  $q(\cdot)$  para todo  $t \in \mathbb{T} = [0, \infty)$ , então uma integral estocástica escalar tem a forma:

$$I(t, \cdot) = \int_{t_0}^t a(\tau) d\beta(\tau, \cdot)$$

A própria notação sugere que  $I(t, \cdot)$  é, para um dado instante  $t$ , uma v.a. e, logicamente,  $I(\cdot, \cdot)$  é um processo estocástico. Provaremos, então, que a integral estocástica definida acima é um processo estocástico com propriedades estatísticas possíveis de serem determinadas desde que  $\int_{t_0}^t a^2(\tau) q(\tau) d\tau$  seja finita.

A prova acima é feita através do conceito de convergência em média quadrática. Como movimento Browniano é Gaussiano e só operações lineares foram executadas,  $I(t, \cdot)$  é também gaussiano com média e covariância dados por:

$$E[I(t)] = 0$$

$$E[I^2(t)] = \int_{t_0}^t a^2(\tau) q(\tau) d\tau$$

As integrais estocásticas exibem as mesmas propriedades lineares das integrais ordinárias, saber:

$$1) \int_{t_0}^{t_2} a(\tau) d\beta(\tau) = \int_{t_0}^{t_1} a(\tau) d\beta(\tau) + \int_{t_1}^{t_2} a(\tau) d\beta(\tau)$$

$$2) \int_{t_1}^t [a(\tau) + b(\tau)] d\beta(\tau) = \int_{t_1}^t a(\tau) d\beta(\tau) + \int_{t_1}^t b(\tau) d\beta(\tau)$$

$$3) \int_{t_0}^t a(\tau) d[\beta(\tau) + \beta'(\tau)] = \int_{t_0}^t a(\tau) d\beta(\tau) + \int_{t_0}^t a(\tau) d\beta'(\tau)$$

$$4) \int_{t_0}^t c \alpha(\tau) d\beta(\tau) = c \int_{t_0}^t \alpha(\tau) d\beta(\tau) = \int_{t_0}^t \alpha(\tau) d[c\beta(\tau)]$$

$$5) \int_{t_0}^t \alpha(\tau) d\beta(\tau) = \alpha(\tau) \beta(\tau) \Big|_{t_0}^t - \int_{t_0}^t \beta(\tau) d\alpha(\tau)$$

A propriedade 5) demonstra que a integração por partes é também válida para integrais estocásticas.

A integral estocástica na forma definida é também contínua, o que completa a relação das propriedades necessárias para caracterizar  $I(\cdot, \cdot)$  também como um processo Browniano com novo coeficiente de difusão:

$$q_I(t) = \int_{t_0}^t \sigma^2(\tau) q(\tau) d\sigma$$

Para fazer a extensão das integrais estocásticas para o caso vetorial basta considerar um processo Browniano vetorial de dimensão  $s$ , com estatística

$$E[\beta(t_i)] = 0$$

$$E[(\beta(t_2) - \beta(t_1)) (\beta(t_2) - \beta(t_1))^T] = \int_{t_1}^{t_2} Q(\tau) d\tau$$

onde  $Q(t)$ , matriz de difusão é simétrica, definida positiva e, pelo menos, contínua por partes.

Sendo  $A(t)$  uma matriz  $(s \times s)$  de funções conhecidas do tempo, define-se integral estocástica vetorial por:

$$I(t, \cdot) = \int_{t_0}^t A(\tau) d\beta(\tau)$$

O vetor aleatório  $I(t, \cdot)$  é Gaussiano, com estatística

$$E[I(t)] = 0$$

$$E[I(t) I(t)^T] = \int_{t_0}^t A(\tau) Q(\tau) A^T(\tau) d\tau$$

De mesma forma, demonstra-se que  $I(\cdot, \cdot)$  é um processo Browniano vetorial com difusão redefinida por:

$$E\left[\{I(t_2) - I(t_1)\}\{I(t_2) - I(t_1)\}^T\right] = \int_{t_1}^{t_2} A(\tau) Q(\tau) A(\tau)^T d\tau$$

Visto que processo Browniano é não diferenciável, forma diferenciais que, no caso do cálculo diferencial e integral, são definidas através das derivadas, no caso de processos estocásticos devem ser definidas através de integrais estocásticas. Nesse caso, dada a integral estocástica

$$I(t) = I(t_0) + \int_{t_0}^t A(\tau) d\beta(\tau)$$

a diferencial estocástica de  $I(t)$  é definida por:

$$dI(t) = A(t) d\beta(t)$$

Para completar o longo resto dos requisitos que permitem a interpretação das equações diferenciais estocásticas lineares, resta concetrar o que é a diferencial do produto de uma função do tempo por uma integral estocástica.

Admita  $\mathcal{S}(\cdot, \cdot)$  como integral estocástica (processo Browniano) dada por:

$$\mathcal{S}(t) = \mathcal{S}(t_0) + \int_{t_0}^t A(\tau) d\beta(\tau)$$

Admita também, que  $D(\cdot)$  é uma matriz conhecida de funções diferenciáveis e que se define um novo processo  $y(\cdot, \cdot)$  através de:

$$y(t) = D(t) \mathcal{S}(t)$$

Nesse caso, é possível demonstrar que a diferencial

estocástico de  $y_{11,..}$  é dado por:

$$dy(t) = \tilde{J}(t) S(t) dt + J(t) dS(t)$$

O que comprova que a diferençaial estocástica de uma forma linear (como  $y = DS$ ) obedece exatamente as mesmas regras de uma diferenciação linear do cálculo diferencial - integral.

→ Temos, então, em suas as ferramentas matemáticas que nos permitem tratar uma equação diferencial estocástica linear na forma

$$\rightarrow dx_t = f(t) x(t) dt + B(t) u(t) dt + G(t) d\beta(t) \quad (***)$$

No entanto, é sempre mais simples escrever e compreender a equação assim em termos de processos de ruído branco. A forma assim será então resumida a:

$$\dot{x}(t) = F(t) x(t) + B(t) u(t) + G(t) w(t) \quad (*)$$

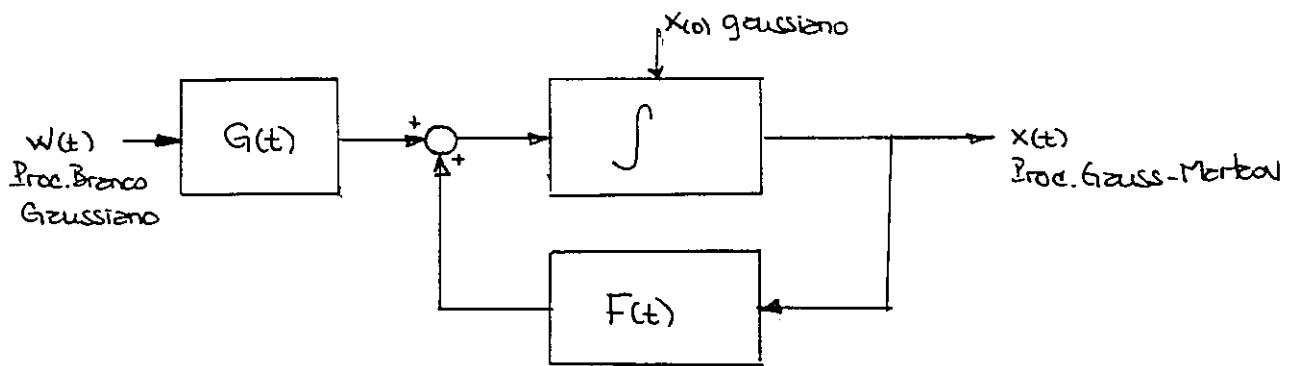
para maior facilidade de entendimento, mas estará sempre implícito que a representação correta é dada por (\*\*\*)

Compreendida essa observação, podemos enunciar o seguinte teorema, fundamental para os processos de Gauss-Markov.

→ Teorema:

Um processo aleatório de Gauss-Markov pode sempre ser representado pelo vetor de estado de um sistema dinâmico linear contínuo forçado por um processo puramente aleatório (branco) onde o vetor de estado inicial é gaussiano.

$$\dot{x}(t) = F(t) x(t) + G(t) w(t) + (B(t) u(t))$$



Se admitirmos que, sendo  $x(t)$  um vetor ( $n \times 1$ ) e  $w(t)$  um vetor ( $m \times 1$ ), a distribuição inicial, correspondente ao instante  $t_0$ , é gaussiana com

$$a) E[x(t_0)] = \mu_x(t_0) \quad \text{conhecida}$$

$$b) E[(x(t_0) - \mu_x(t_0))(x(t_0) - \mu_x(t_0))^T] = P(t_0) \quad \text{conhecida}$$

que o ruído é branco, gaussiano, com estatísticas:

$$c) E[w(t)] = \bar{w}(t)$$

$$d) E[(w(t) - \bar{w}(t))(w(t) - \bar{w}(t))^T] = Q(t) \delta(t-t)$$

e, finalmente, que o ruído é não correlacionado com o estado, isto é:

$$e) E[(w(t) - \bar{w}(t))(x(t_0) - \mu_x(t_0))^T] = 0$$

então, as equações diferenciais que governam a evolução da média e da covariância do processo de Gauss-Markov são:

1) Para a média:

$$\dot{\mu}_x(t) = F(t) \mu_x(t) + G(t) \bar{w}(t) + B u(t)$$

2) Para a matriz de Covariância:

$$\dot{P}_{xx}(t) = F(t) P_{xx}(t) + P_{xx}(t) F^T(t) + G(t) Q(t) G^T(t)$$

A equação 2) é uma equação de Riccati, de uso bastante comum em Engenharia, porém de solução bastante complicada.

As equações 1) e 2) podem ser reescritas em termos

da matriz de transição do sistema dinâmico linear, ficando assim formular equivalente:

3) Para a média:

$$\mu_x(t) = \Phi(t, t_0) \mu_x(t_0) + \int_{t_0}^t \Phi(t, s) G(s) \bar{w}(s) ds$$

4) Para a covariância:

$$P_{xx}(t) = \Phi(t, t_0) P_{xx}(t_0) \Phi^T(t, t_0) + \int_{t_0}^t \Phi(t, s) G(s) Q(s) G^T(s) \Phi^T(s) ds$$

Exemplos:

a) Considere um oscilador harmônico de frequência 1 rad/s forçado por ruídos brancos Gaussianos. A equação de estado é dada por:

$$\ddot{x}(t) = F(t) - x(t) + G_1(t) w_1(t)$$

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}$$

Admita condições iniciais em  $t_0 = 0$  tais que o oscilador parte precisamente de  $x_1(0) = 1$  e  $x_2(0) = 3$ . Admita  $w_1$  e  $w_2$ , independentes, de média nula e intensidades (difusões) 1 e 2, respectivamente, isto é:

$$E[w_1(t) w_1(t+\tau)] = 1 \delta(t)$$

$$\bar{w} = 0$$

$$E[w_2(t) w_2(t+\tau)] = 2 \delta(t)$$

$$E[w_1(t) w_2(t+\tau)] = 0$$

Desde que as condições iniciais são conhecidas com

total certeza,  $x(t_0)$  pode ser modelado como uma v.z. gaussiana com matriz de covariância nula, i.e.:

$$\mu_x(0) = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad P_{xx}(0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

A matriz  $Q$  de covariâncias do ruído, em função dos dados do problema, é dada por

$$Q(t) = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

A determinação da matriz de transição para um sistema como o proposto já foi feita diversas vezes no decorrer do curso.

$$\underline{\Phi}(t, t_0) = \begin{bmatrix} \cos(t-t_0) & \sin(t-t_0) \\ -\sin(t-t_0) & \cos(t-t_0) \end{bmatrix}$$

Teremos, então, para a média:

$$\mu_x(t) = \underline{\Phi}(t, t_0) \mu_x(t_0) + \int_{t_0}^t \underline{\Phi}(t, s) G(s) \tilde{w}(s) ds = \underline{\Phi}(t, t_0) \mu_x(t_0)$$

$$\mu_x(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} \cos t + 3 \sin t \\ 3 \cos t - \sin t \end{bmatrix}$$

Para a matriz de covariâncias:

$$P_{xx}(t) = \underline{\Phi}(t, t_0) P_{xx}(t_0) \underline{\Phi}^T(t, t_0) + \int_{t_0}^t \underline{\Phi}(t, s) G(s) Q(s) G^T(s) \underline{\Phi}^T(t, s) ds$$

$$P_{xx}(t) = \int_0^t \begin{bmatrix} \cos(t-s) & \sin(t-s) \\ -\sin(t-s) & \cos(t-s) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \cos(t-s) & -\sin(t-s) \\ \sin(t-s) & \cos(t-s) \end{bmatrix} ds$$

$$P_{xx}(t) = \begin{bmatrix} \left(\frac{\epsilon^2 + 2b^2}{2}\right)t + \left(\frac{\epsilon^2 - 2b^2}{4}\right)\sin 2t & \left(\frac{2b^2 - \epsilon^2}{2}\right)\sin 2t \\ \left(\frac{2b^2 - \epsilon^2}{2}\right)\sin 2t & \left(\frac{\epsilon^2 + 2b^2}{2}\right)t + \left(\frac{2b^2 - \epsilon^2}{4}\right)\sin 2t \end{bmatrix}$$

Note que a covariância está divergindo, com as entradas de diagonal crescendo linearmente com o tempo e com suas saídas sobreposta.

b) Considere o mesmo exemplo a) agora com um coeficiente de amortecimento  $\zeta \neq 0$ , de tal forma que a matriz  $F$  se torna:

$$F = \begin{bmatrix} 0 & 1 \\ -1 & -2\zeta \end{bmatrix}$$

Como a matriz de transição nesse caso já é um pouco complicada, vamos apenas mudar as equações diferenciais para média e matriz de covariâncias. Para a média:

$$\dot{\mu}_{xx}(t) = F(t) \mu_{xx}(t) + G(t) \bar{w}(t) = F(t) \mu_{xx}(t)$$

$$\dot{\mu}_{xx}(t) = \begin{bmatrix} \dot{\mu}_{x_1}(t) \\ \dot{\mu}_{x_2}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -2\zeta \end{bmatrix} \begin{bmatrix} \mu_{x_1}(t) \\ \mu_{x_2}(t) \end{bmatrix} = \begin{bmatrix} \mu_{x_2}(t) \\ -\mu_{x_1}(t) - 2\zeta \mu_{x_2}(t) \end{bmatrix}$$

$$\dot{P}_{xx}(t) = P_{xx}(t) F^T(t) + F(t) P_{xx}(t) + G(t) Q(t) G^T(t)$$

$$\begin{bmatrix} \dot{P}_{11} & \dot{P}_{12} \\ \dot{P}_{12} & \dot{P}_{22} \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & -2\zeta \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -1 & -2\zeta \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & b \end{bmatrix}$$

$$\begin{bmatrix} \dot{P}_{11} & \dot{P}_{12} \\ \dot{P}_{12} & \dot{P}_{22} \end{bmatrix} = \begin{bmatrix} 2P_{12} + \zeta^2 & -P_{11} - 2\zeta P_{12} + P_{22} \\ -P_{11} - 2\zeta P_{12} + P_{22} & -2P_{12} - 4\zeta P_{22} + 2\zeta^2 \end{bmatrix}$$

Esta matriz de covariâncias não cresce sem limite com

valor de regime pode ser obtido fazendo-se  $\dot{P}_{xx}(t) = 0$ , o que leva a:

$$P_{xx}(t \rightarrow \infty) = \begin{bmatrix} \frac{z^2(1-4\beta^2)+2b^2}{4\beta} & z/2 \\ z/2 & \frac{z^2+b^2}{4\beta} \end{bmatrix}$$

### 6.11. EQUAÇÕES ESTOCÁSTICAS LINEARES DE DIFERENÇAS

Equações estocásticas lineares de diferença são de extrema importância por dois motivos. O primeiro é que muitas vezes o modelo do sistema dinâmico envolvido no nosso estudo é fornecido diretamente como um sistema de múltiplos estágios, isto é, a equação que rege a evolução do sistema é uma equação de diferenças. O segundo é que, para simulação no computador, um modelo discretizado pode ser mais facilmente implementado. A transformação para o modelo discretizado é feita a partir de:

$$\dot{x}(t) = F(t)x(t) + G(t)w(t)$$

que, na forma de matriz de transição pode ser escrita como:

$$x(t_{k+1}) = \Phi(t_{k+1}, t_k)x(t_k) + \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, s)G(s)w(s)ds$$

ou, na forma correta, envolvendo movimento Browniano e ruído branco:

$$x(t_{k+1}) = \Phi(t_{k+1}, t_k)x(t_k) + \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, s)G(s)d\beta(s)$$

Essa última relação é equivalente a se escrever:

$$x(t_{k+1}) = \underline{\Phi}(t_{k+1}, t_k) x(t_k) + w_d(t_k)$$

onde  $w_d(t_k)$  (ou  $w_d(\cdot, \cdot)$ ) é uma sequência estocástica Gaussiana branca, vetorial e discreta, com estatísticas

$$E\{w_d(t_k)\} = \bar{w}(t_k)$$

$$\begin{aligned} E[(w_d(t_k) - \bar{w}(t_k))(w_d(t_k) - \bar{w}(t_k))^T] &= Q_d(t_k) = \\ &= \int_{t_k}^{t_{k+1}} \underline{\Phi}(t_{k+1}, \tau) G(\tau) Q(\tau) G^T(\tau) \underline{\Phi}^T(t_{k+1}, \tau) d\tau \end{aligned}$$

$$E[w_d(t_i) \cdot w_d(t_j)] = 0 \quad , \quad t_i \neq t_j$$

Estas propriedades decorrem do fato de que a integral estocástica de um processo Browniano é também um processo Browniano, cujo ruído branco "associado" tem as estatísticas acima.

As sequências estocásticas têm um teorema análogo ao sugerido para processos estocásticos no que se refere aos processos de Gauss-Markov. Esse teorema é de fundamental importância para os desenvolvimentos seguintes.

Teorema:

Uma sequência de Gauss-Markov pode sempre ser representada pelo estado de um sistema linear de múltiplos estágios, gerado por uma sequência branca (puramente aleatória) desde que o vetor de estado inicial seja Gaussiano.

Tomando o sistema discreto:

$$x(k+1) = \underline{\Phi}(k+1, k) x(k) + I(k) w(k)$$

Se admitirmos  $X(0)$  gaussiano, com as seguintes hipóteses:

- estatística do ruído }   
 a)  $E[X(0)] = \mu_0$  conhecida  
 b)  $E[(x(0) - \mu_0)(x(0) - \mu_0)^T] = P(0)$  conhecida  
 c)  $E[W(k)] = \bar{W}(k)$  conhecida para todo  $k$   
 d)  $E[(W(k) - \bar{W}(k))(W(k) - \bar{W}(k))^T] = \begin{cases} 0 & k \neq l \\ Q(k) & k = l \end{cases}$

Admitindo, também que o ruído em qualquer instante e os estados iniciais sejam independentes, o que vai permitir escrever:

$$e) E[(x(0) - \mu_0)(W(k) - \bar{W}(k)^T] = 0$$

Com essas hipóteses, lembrando que transformações lineares de um vetor gaussiano preservam o caráter gaussiano, temos que a função densidade de transição  $p[x(k+1) | x(k)]$  é também gaussiana, com isso, é possível determinar a evolução de  $\mu(k)$  e  $P(k)$ .

Tomando a esperança em ambos os lados de:

$$x(k+1) = \Phi(k+1, k)x(k) + \Gamma(k)W(k)$$

chegamos a:

$$\mu(k+1) = \Phi(k+1, k)\mu_k + \Gamma(k)\bar{W}(k)$$

que define a evolução da média  $\mu(k)$ .

Subtraindo a última relação da anterior:

$$x(k+1) - \mu(k+1) = \Phi(k+1, k)[x(k) - \mu(k)] + \Gamma(k)[W(k) - \bar{W}(k)]$$

Multiplicando a relação acima pelo transposto e tomado a esperança em ambos os lados:

$$\begin{aligned} E[(x(k+1) - \mu(k+1))(x(k+1) - \mu(k+1))^T] &= E[\Phi(k+1, k)[x(k) - \mu(k)][x(k) - \mu(k)]^T\Phi^T(k+1, k) \\ &\quad + \Phi(k)[x(k) - \mu(k)][W(k) - \bar{W}(k)]^T\Gamma^T(k) + \Gamma(k)[W(k) - \bar{W}(k)][x(k) - \mu(k)]^T\Phi^T(k+1, k)] \\ &\quad \times \Phi^T(k) + \Gamma^T(k)[W(k) - \bar{W}(k)][W(k) - \bar{W}(k)]^T\Gamma^T(k)] \end{aligned}$$

Eq. Riccati discreta

$$\underline{P}(k+1) = \underline{\Phi}(k+1, k) \underline{P}(k) \underline{\Phi}^T(k+1, k) + \underline{\Gamma}(k) \underline{Q}(k) \underline{\Gamma}^T(k)$$

↳ MATRIZ COVARIÂNCIA

## 6.12. SHAPING FILTERS

Em muitas ocasiões, o uso de modelos de ruído branco Gaussiano para descrever todos os ruídos envolvidos no sistema real pode resultar não adequado. Seria desejável, nesses casos, gerar autocorrelações empíricas ou dados da densidade espectral de potência e ai desenvolver um modelo matemático que produza como saída o ruído realmente observado.

Se os dados observados forem, de fato, amostras de um movimento Browniano ou de um processo estacionário Gaussiano cuja densidade espectral conhecida, então um sistema linear invariante no tempo, o "shaping filter", fornece o modelo correto do ruído. Mesmo no caso de densidades espectrais complicadas, pode-se fazer aproximações por densidades espetrais simples e usar o mesmo procedimento. Embora não sejam desenvolvidos neste item, é possível também projetar "shaping filters" variáveis no tempo.

Consideramos então o nosso sistema de interesse ser descrito por:

$$\dot{x}(t) = F(t)x(t) + G(t)\eta(t)$$

onde  $\eta(\cdot, \cdot)$  é um ruído Gaussiano correlacionado no tempo, i.e., não branco. Também, admitimos que  $\eta(\cdot, \cdot)$  possa ser gerado por um shaping filter linear:

$$\dot{x}_f(t) = \underbrace{F_f(t)}_{x_f(t)} + G_f(t) w(t)$$

$$\eta(t) = H_f(t) x_f(t)$$

onde o subscrito f denota filtro e onde  $w(\cdot, \cdot)$  é um ruído branco Gaussiano.

Então, a saída do "shaping filter" pode ser usada para forçar o sistema original. Definindo um vetor de estado aumentado  $x_a(\cdot, \cdot)$  através de

$$x_a(\cdot, \cdot) = \begin{bmatrix} x(\cdot, \cdot) \\ x_f(\cdot, \cdot) \end{bmatrix}$$

reescrivemos a equação diferencial do estado aumentado:

$$\begin{bmatrix} \dot{x}(t) \\ \dot{x}_f(t) \end{bmatrix} = \begin{bmatrix} F(t) & \begin{bmatrix} G(t)H_f(t) \\ 0 \end{bmatrix} \\ 0 & \begin{bmatrix} F_f(t) & 0 \end{bmatrix} \end{bmatrix} \begin{bmatrix} x(t) \\ x_f(t) \end{bmatrix} + \begin{bmatrix} 0 \\ G_f(t) \end{bmatrix} w(t)$$

$$\dot{x}_a(t) = F_a(t)x_a(t) + G_a(t)w(t)$$

Se a equação de observações do sistema original é dada por:

$$z(t) = H(t)x(t) + v(t)$$

então, para o estado aumentado o modelo de observações será dado por:

$$z(t) = [H(t); 0] \begin{bmatrix} x(t) \\ x_f(t) \end{bmatrix} + v(t)$$

$$z(t) = H_a(t)x_a(t) + v(t)$$

O modelo final obtido é então um sistema linear forçado apenas por ruído branco Gaussiano.

Um desenvolvimento análogo é possível para o caso de ruído correlacionado no tempo nas medidas. Considerando o sistema:

$$\dot{x}(t) = F(t)x(t) + G(t)w(t)$$

$$z(t) = H(t)x(t) + n(t) + v(t)$$

onde  $w(\cdot, \cdot) \in v(\cdot, \cdot)$  são ruídos brancos e  $n(\cdot, \cdot)$  é não branco.

$n(\cdot, \cdot)$  pode ser gerado pela saída de um shaping filter forçado por ruído branco Gaussiano  $w_f(\cdot, \cdot)$ :

$$\dot{x}_f(t) = F_f(t) x_f(t) + G_f(t) w_f(t)$$

$$n(t) = H_f(t) x_f(t)$$

O estado aumentado pode ser definido da mesma maneira que a já feita, levando a um sistema aumentado descrito por:

$$\begin{bmatrix} \dot{x}(t) \\ \dot{x}_f(t) \end{bmatrix} = \begin{bmatrix} F(t) & 0 \\ 0 & F_f(t) \end{bmatrix} \begin{bmatrix} x(t) \\ x_f(t) \end{bmatrix} + \begin{bmatrix} G(t) & 0 \\ 0 & G_f(t) \end{bmatrix} \begin{bmatrix} w(t) \\ w_f(t) \end{bmatrix}$$

$$z(t) = \begin{bmatrix} H(t) & H_f(t) \end{bmatrix} \begin{bmatrix} x(t) \\ x_f(t) \end{bmatrix} + v(t)$$

que, de novo, está na forma de um sistema linear forçado apenas por ruído branco Gaussiano.

Algumas configurações de shaping filters são bastante comuns para modelamento de certos processos e vamos discuti-las individualmente:

a) O caso trivial é o do próprio ruído branco Gaussiano, de média  $m_0$  e autocorrelação

$$E[w(t) w(t+\tau)] = [P_0 + m_0^2] \delta(\tau) = X_0 \delta(\tau)$$

b) O modelo de bias ou constante eletrônica é gerado como a saída de um integrador sem entradas, mas com condições iniciais modeladas como v.z. Gaussiana  $x(0)$  com média  $\mu_0$  e variância  $P_0$  especificadas. Então

$$\dot{x}(t) = 0$$

Este é uma forma degenerada do shaping filter em que nenhum ruído força a equação de estado do filtro. Desde que as amostras são constantes no tempo, a autocorrelação é constante

para qualquer  $t$ , resultando numa densidade espectral impulsiva com toda potência concentrada em  $\omega = 0$ :

$$\Phi_{xx}(t_0) = E[x(t)x(t+\tau)] = [B_0 + m_0^2]$$

$$\bar{\Phi}_{xx}(\omega) = 2\pi[B_0 + m_0^2]\delta(\omega)$$

Esse modelo é muito usado para "biç" de giroscópios e outros sensores: de um período de operação para outro, o nível de "biç" pode mudar mas se manter constante enquanto o instrumento é ~~reajustado~~. Muito cuidado deve ser tomado quando se usar esse modelo num estimador ótimo: o modelo supõe que embora não saibamos o nível de "biç" a priori, sabemos que esse "biç" não muda seu valor no tempo. Com isso, o estimador tende a estimar o valor do "biç" usando as primeiras informações e a desprezar as informações que chegam depois. Se ~~se~~ pretendemos manter o "biç" em estimativas razoáveis é preferível usar valores que variam lentamente através do modelo de movimento browniano descrito a seguir.

c) Movimento Browniano (ou passeio elétórico) é dado pelo secc de um integrador forçado por ruído Branco Gaussiano

$$\ddot{x}(t) = w(t) \quad x(t_0) = 0$$

onde  $w(t)$  tem média nula e  $E[w(t)w(t+\tau)] = Q\delta(\tau)$ . A equação de média será a mesma que faz a constante elétórica ( $\mu_x(t) = 0$ ) mas a estatística de segunda ordem é diferente:

$$\dot{\bar{x}}_{xx}(t) = Q \quad \text{ao invés de } \dot{\bar{x}}_{xx}(t) = 0$$

de tal forma que o valor médio quadrático cresce linearmente com o tempo:

$$E[\bar{x}^2(t)] = Q(t - t_0)$$

d) Processos Exponenciais Fruente Correlacionados no Tempo ou processos de Markov de primeira ordem são casos de 1º ordem forçados por ruído branco Gaussiano de média nula e intensidade de  $Q$ . Então, para produzir uma saída com autocorrelação

$$\Psi_{xx}(\tau) = E[x(t)x(t+\tau)] = Q^2 e^{-|t|\tau/\sigma^2}$$

isto é, com tempo de correlação  $T$  e valor médio quadrático  $Q^2$ , com média nula, o modelo deve ser descrito por:

$$\dot{x}(t) = -\left(\frac{1}{T}\right)x(t) + w(t)$$

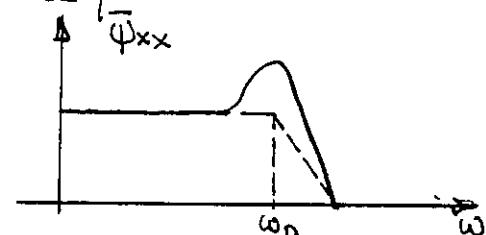
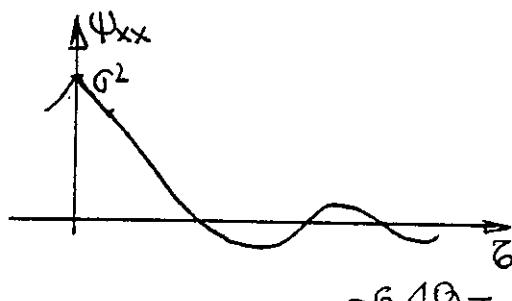
Nesse caso,  $Q = 2\sigma^2/T$ , i.e.,  $E[x^2(t)] = QT/2$  e a densidade espectral associada é:

$$\bar{\Psi}_{xx}(\omega) = \frac{2\sigma^2/T}{\omega^2 + \left(\frac{1}{T}\right)^2}$$

Este particular tipo de shaping filter fornece boa aproximação para uma grande variedade de ruídos observados que sejam tipo de banda limitada, seja a banda estreita ou larga. Para uma representação gráfica desse ruído veja a seção 6.5.

e) Um processo de Markov de segunda ordem fornece um modelo muito bom para fenômenos aleatórios oscilantes como vibrações e flexões em veículos aeroespaciais. A forma geral de sua autocorrelação é:

$$\Psi_{xx}(\tau) = E[x(t)x(t+\tau)] = \frac{\sigma^2}{\cos\eta} e^{-\xi\omega_n|\tau|} \cos(\sqrt{1-\xi^2}\omega_n|\tau| - \eta)$$



Esse ruído pode ser gerado passando um ruído branco Gaussiano estacionário  $w(\cdot, \cdot)$  de intensidade  $Q=1$  através de um sistema de segunda ordem:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\zeta\omega_n \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} z \\ c \end{bmatrix} w(t)$$

onde  $x_1(t)$  é a saída do sistema e

$$z = [(20^2/\cos\gamma)\omega_n \sin(\alpha - \gamma)]^{1/2}$$

$$c = b - 2\zeta\omega_n$$

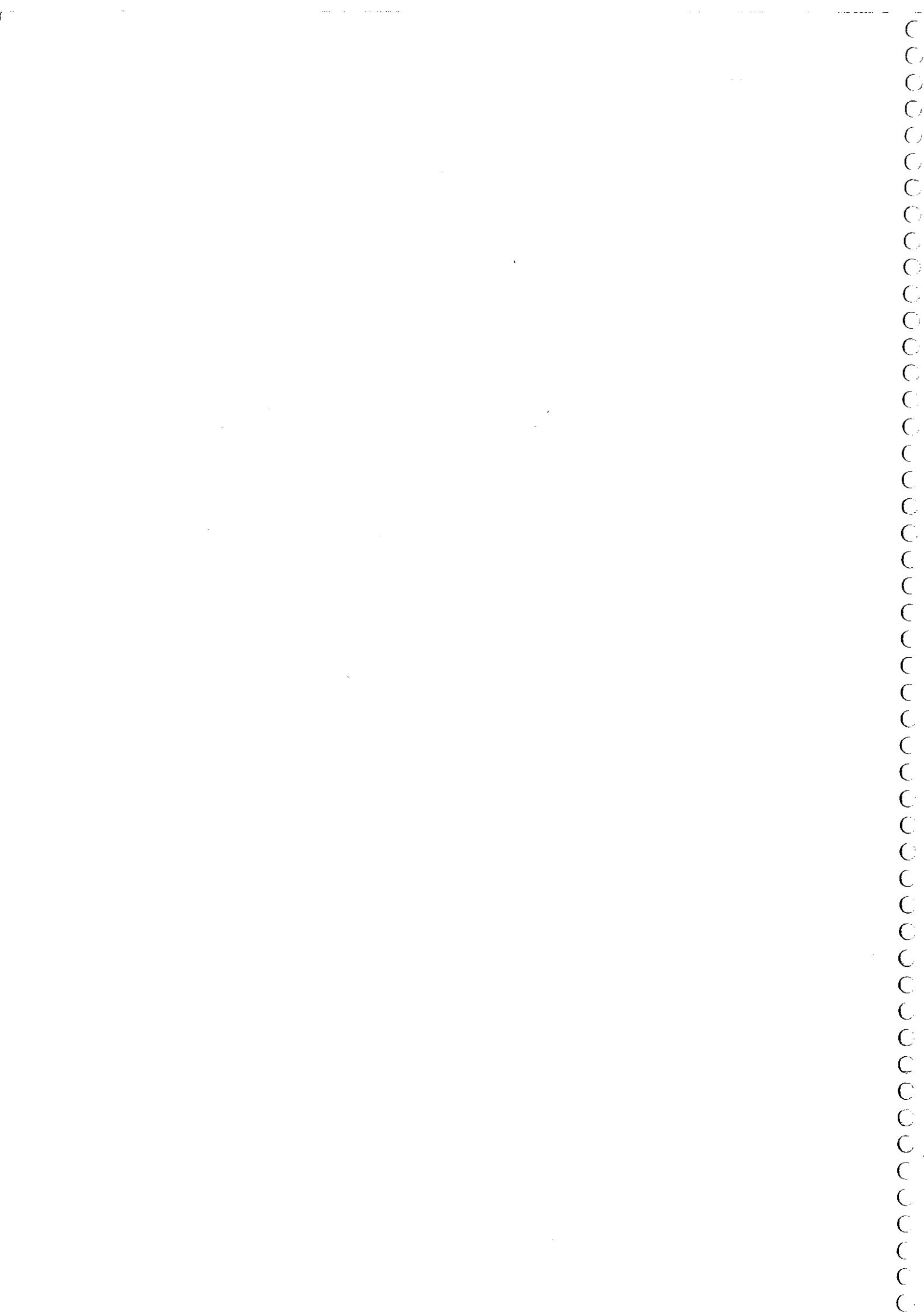
$$\alpha = \operatorname{arctg} [\bar{z}/\sqrt{1-\bar{z}^2}]$$

## CAPÍTULO 7 - FILTRO DE KALMAN PARA SISTEMAS LINEARES

Ao longo dos capítulos anteriores foram descritas e montadas as diversas ferramentas para se obter a forma do filtro de Kalman. Este capítulo é, então, uma síntese do que foi visto visto para montar os pacotes computacionais que compõe o chamado filtro de Kalman.

Serão obtidas três versões do filtro: a versão de menor utilização em sistemas de Engenharia, onde a dinâmica é contínua no tempo e as observações são feitas de modo discreto; a versão de dinâmica discreta com observações discretas no tempo e, finalmente, o caso de dinâmica contínua com medidas contínuas no tempo.

A maneira de se obter o filtro de Kalman não segue aqui o desenvolvimento de Kalman que usou resultados da análise funcional para chegar à forma final. O desenvolvimento que foi feito nestas notas permite obter o filtro através de considerações sobre funções densidade condicional, processos estocásticos e estimadores para sistemas estôsticos. Este conteúdo nos levará à forma de Kalman para filtragem e predição em sistemas dinâmicos, isto é, dado uma observação num instante  $t_k$ , sermos capazes de estimar o estado do sistema no instante  $t_k$ , dito processo de filtragem, e de estimar o estado do sistema num instante de futuro,  $t_{k+1}$ , dito processo de predição. A mesma teoria se aplica facilmente ao problema de alinhamento, isto é, dado a observação no instante  $t_k$ , como determinar o estado do sistema no instante  $t_i$ ,  $t_i < t_k$ . O processo de alinhamento, de extrema importância em tratamento de dados aleatórios, não será desenvolvido e recomenda-se aos interessados nesse assunto a leitura de Izzwinski (1970) ou Bryson e Ho (1975).



#### 4.1. MODELO LINEAR CONTÍNUO / DISCRETO

Tendo em vista sua maior utilização em problemas de Engenharia, o filtro de Kalman será obtido primeiramente para um sistema dinâmico contínuo com observações discretas no tempo. Nesse caso, nosso modelo linear é descrito por uma equação diferencial forcingada por ruído branco, o que, se rechocado corretamente, é uma equação de incrementos forcingada por um processo de movimento Browniano:

$$\dot{\hat{x}}(t) = F(t) \hat{x}(t) + G(t) w(t)$$

ou ainda:

$$dx_t = F(t) x(t) dt + G(t) d\beta(t)$$

onde

$x(t)$  é o vetor ( $n \times 1$ ) do estado do sistema;

$F(t)$  é uma matriz ( $n \times n$ ) de funções conhecidas do tempo, pelas quais continuo por tratos;

$G(t)$  é uma matriz ( $n \times s$ ) de funções conhecidas

$w(t)$  é um vetor ( $s \times 1$ ) de ruído branco gaussiano com estatística

$$E[w(t)] = 0$$

$$E[w(t) w(t')] = Q(t) \delta(t-t')$$

No caso de movimento Browniano  $d\beta$  seria um vetor ( $s \times 1$ ) com estatística dada por:

$$E[\beta(t)] = 0$$

$$E[(\beta(t) - \beta(t'))(\beta(t) - \beta(t'))^T] = \int_{t'}^t Q(\tau) d\tau$$

Supõe-se, como dado indistintivamente é montado o filtro que o estado inicial do sistema dinâmico (processo de Gauss-Markov) seja Gaussiano e conhecido com estatística dada por:

$$E[x(t_0)] = \hat{x}_0$$

$$E[(x(t_0) - \hat{x}_0)(x(t_0) - \hat{x}_0)^T] = P_0$$

O modelo de observações, sendo feitas em instantes discretos no tempo, é descrito por uma sequência de Gauss-Markov, forçada por uma sequência branca:

$$z(t_i) = H(t_i)x(t_i) + v(t_i)$$

onde

$z$  é um vetor ( $m \times 1$ ) de medidas (nos) no instante,

$H$  é uma matriz ( $m \times n$ ) conhecida

$v$  é um vetor ( $m \times 1$ ) branco com estatística:

$$E[v(t_i)] = 0$$

$$E[v(t_i)v^T(t_j)] = \begin{cases} R(t_i) & t_i = t_j \\ 0 & t_i \neq t_j \end{cases}$$

Para completar as hipóteses sobre o modelo colocado, as matrizes de covariância de ruídos, no estorso,  $Q(t)$  e, no ruído,  $R(t)$ , são conhecidas para qualquer instante no intervalo de tempo considerado no nosso problema, sendo que  $Q(t)$  é pelo menos semidefinita positiva e  $R(t)$  é sempre definida positiva, ambas simétricas. Finalmente, a última e importíssima hipótese, admite-se que  $x(t_0)$ ,  $w(\cdot, \cdot)$  (ou  $\beta(\cdot, \cdot)$ ) e  $v(\cdot, \cdot)$  são processos mutuamente independentes. Como são Gaussianos, todos serão também não correlacionados entre si.

O critério de desempenho do estimador é também um dado fundamental. O filtro de Kalman é otimizado segundo o ponto de vista Bayesiano, isto é, procurar-se maneiras de prover a função densidade de probabilidade condicionada a todos



história das observações feitas.

Com esse critério, a função densidade condicionada, usando as hipóteses já colocações, permanecerá sempre Gaussiana e, com isso, a média, a moda, a mediana e qualquer escolha lógica de estimativa baseada na densidade condicionada convergem para o mesmo valor estimado.

Para coletar o histórico das observações feitas, introduz-se dois vetores a saber:

$$\underline{Z}(t_i) = \begin{bmatrix} z(t_1) \\ \vdots \\ z(t_i) \end{bmatrix}$$

vetor de dimensão crescente que contém toda história das medidas realizadas, acrescentadas uma a uma;

$$\underline{z}_i = \begin{bmatrix} z_1 \\ \vdots \\ z_i \end{bmatrix}$$

vetor também de dimensão crescente que contém todos os vetores individuais de números medidas no decorrer do tempo.

## 7.2. FILTRO DE KALMAN PARA SISTEMAS CONTÍNUOS / DISCRETOS.

Descreto o modelo completo no item anterior, vamos nos preocupar com a montagem do filtro no intervalo  $(t_{i-1}; t_i)$  contido no intervalo  $(t_0, t_f)$  de interesse.





Admitimos processada a medida  $z(t_{i-1}, w_j) = z_{i-1}$ , onde  $z(t_{i-1}, w_j)$  é componente do vetor  $\bar{z}$  e  $z_{i-1}$  é uma componente do vetor de números  $\bar{z}$ . Admitimos também, como resultado desse processamento, que  $f_{x(t_{i-1}) | \bar{z}(t_{i-1})}(\bar{z} | \bar{z}_{i-1})$ , isto é, a função densidade do estado em  $t_{i-1}$ , condicionada a toda história das medidas até  $t_{i-1}$ , que está contida em  $\bar{z}(t_{i-1})$ , é Gaussiana com estatística dada por:

$$\text{Média} - \hat{x}(t_{i-1}^+)$$

$$\text{Matriz de Covariância} - P(t_{i-1}^+)$$

Isto pode ser mostrado a partir do fato de que o estado inicial do sistema  $x(t_0)$  é Gaussiano com função densidade  $f_{x(t_0)}(\bar{z})$  conhecido e não condicionada, já que não há informações de medidas no instante  $t_0$ . Usando regras de Bayes e as propriedades dos processos envolvidos, Maybeck (1979) e Jezwinski (1970) demonstram a afirmação acima.

Sendo  $f_{x(t_{i-1}) | \bar{z}(t_{i-1})}(\bar{z} | \bar{z}_{i-1})$  Gaussiana conhecida, então o primeiro problema é conseguir produzir uma estimativa do estado no instante  $t_i$  onde temos disponível uma nova observação  $z_i$ . Se pudermos gerar tal estimativa, esta servirá como uma estimativa a priori quando da incorporação da nova observação.

Em termos de função densidade condicionada, como não há nenhuma observação sendo feita no intervalo  $(t_{i-1}^+, t_i)$ , o problema se resume em gerar  $f_{x(t_i) | \bar{z}(t_{i-1})}(\bar{z} | \bar{z}_{i-1})$ .

É possível provar que a função acima, i.e., a densidade no instante  $t_i$  condicionada a todo histórico das informações etc o instante  $t_{i-1}$ , é também Gaussiana (Maybeck, 1979) com estatística dada por:

$$\text{Média: } \hat{x}(t_i^-) = \underline{\Phi}(t_i, t_{i-1}) \hat{x}(t_{i-1}^+)$$

$$\text{Covariância: } \underline{P}(t_i^+) = \underline{\Phi}(t_i, t_{i-1}) \underline{P}(t_{i-1}^+) \underline{\Phi}^T(t_i, t_{i-1}) + \\ + \int_{t_{i-1}}^{t_i} \underline{\Phi}(t_i, \tau) G(\tau) Q(\tau) G^T(\tau) \underline{\Phi}^T(t_i, \tau) d\tau$$

Esse desenvolvimento é chamado ciclo de projeção do filtro de Kalman, levando estado e covariância desde o instante  $t_{i-1}^+$ , imediatamente posterior à incorporação da medida  $z_{i-1}$ , até o instante  $t_i$ , imediatamente anterior à incorporação da nova medida  $z_i$ .

A solução colocada acima corresponde à solução do problema as formas discretizadas através do uso da matriz de transição do estado. As formas correspondentes para a projeção são que as diferenciais são dadas por:

$$\text{Média: } \hat{x}(t) = f(t) \hat{x}(t) \quad t_{i-1}^+ \leq t \leq t_i^-$$

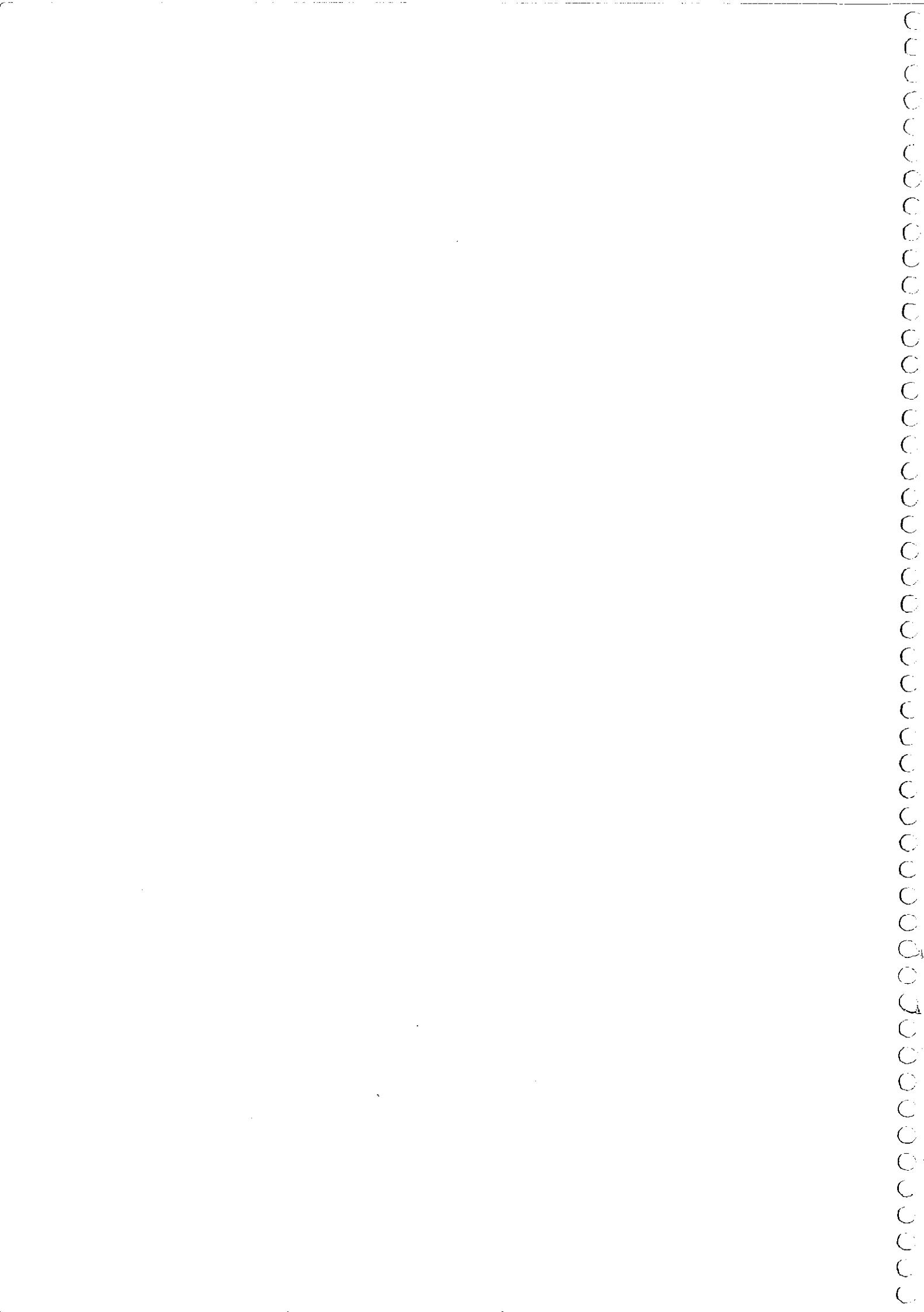
$$\text{Covariância: } \dot{\underline{P}}(t) = f(t) \underline{P}(t) + \underline{P}(t) F^T(t) + G(t) Q(t) G^T(t) \quad t_{i-1}^+ < t < t_i^-$$

Resolvendo o problema de se gerar uma estimativa do estado no instante em que será incorporado o novo vetor de medidas, deseja-se agora buscar a função densidade do estado condicionada a todas as observações inclusive a medida recente chegada  $z_i$ .

Dase-se, então, procurar gerar a densidade condicionada  $f_{x(t_i)|z(t_i)}(z|z_i)$ . Demonstra-se (Maybeck, 1979) que essa densidade é também Gaussiana com:

$$\text{Média: } \hat{x}(t_i^+) = \hat{x}(t_i^-) + K(t_i) [z_i - H(t_i) \hat{x}(t_i^-)]$$

$$\text{Covariância: } \underline{P}(t_i^+) = \underline{P}(t_i^-) - K(t_i) H(t_i) \underline{P}(t_i^-)$$



onde  $t_i^+$  se refere ao instante imediatamente posterior à incorreção da nova medida e  $K(t_i)$  é o fator de kalman dado por:

$$K(t_i) = P(t_i^-) H^T(t_i) [H(t_i) P(t_i^-) H^T(t_i) + R(t_i)]^{-1}$$

As três últimas relações correspondem ao chamado ciclo de atualização do filtro contínuo / discreto onde o novo vetor de medidas  $z_i$  é processado para melhorar nossas estimativas. Note que essas equações são extensões das mesmas desenvolvidas na seção 5.6. do capítulo 5 sobre a forma de kalman para sistemas invariantes no tempo, agora usando  $\hat{x}(t_i^-)$  como estimativa a priori do estado com matriz de covariância  $P(t_i^-)$ .

\* Para resumir o algoritmo, temos que dadas as condições iniciais

$$\hat{x}(t_0) = E[x(t_0)] = \hat{x}_0$$

$$P(t_0) = E[(x(t_0) - \hat{x}_0)(x(t_0) - \hat{x}_0)^T] = P_0$$

O filtro de kalman é um algoritmo recursivo que para sistemas contínuos com observações discretas consiste de dois passos:

a) Passo de propagação entre  $t_{i-1}^+$  e  $t_i^-$

$$\hat{x}(t_i^-) = \underline{\Phi}(t_i, t_{i-1}) \hat{x}(t_{i-1}^+) \quad (\text{ou } \hat{x} = F(t) \hat{x})$$

$$P(t_i^-) = \underline{\Phi}(t_i, t_{i-1}) P(t_{i-1}^+) \underline{\Phi}^T(t_i, t_{i-1}) + \int_{t_{i-1}}^{t_i} \underline{\Phi}(t_i, \tau) G(\tau) Q(\tau) G^T(\tau) \underline{\Phi}^T(t_i, \tau) d\tau$$

$$(\text{ou } \hat{P}(t) = F(t) \hat{P}(t) + P(t) F^T(t) + G(t) Q(t) G^T(t))$$

b) Passo de atualização entre  $t_i^-$  e  $t_i^+$

$$K(t_i) = P(t_i^-) H^T(t_i) [H(t_i) P(t_i^-) H^T(t_i) + R(t_i)]^{-1}$$

$$\hat{x}(t_i^+) = \hat{x}(t_i^-) + K(t_i) [z_i - H(t_i) \hat{x}(t_i^-)]$$

$$P(t_i^+) = P(t_i^-) - K(t_i) H(t_i) P(t_i^-)$$

Como uma última observação, note que o ciclo de propagação gera  $\hat{x}(t_i)$ , a melhor estimativa do estado antes da medida em  $t_i$ ,  $z(t_i, w_j) = z_j$ . O ciclo de atualização gera  $H(t_i)\hat{x}(t_i)$  que é a melhor previsão do que será a medida no instante  $t_i$  antes que a medida seja realmente feita.

O resíduo de medida  $r(t_i)$  é então gerado como a diferença entre o valor observado  $z_i$  e a melhor previsão dele antes que seja realizada, i.e.:

$$r(t_i) = z_i - H(t_i)\hat{x}(t_i)$$

O resíduo de medida  $r(t_i)$  é uma quantidade importante para considerações de convergência do filtro e muitas vezes é usado em processos adaptativos, como veremos mais adiante.

Exemplo:

Como exemplo consideraremos uma representação bastante simples para o teste de um giro. O exemplo será baseado num sistema dinâmico escalar. Sabemos que os giros estão sujeitos a deriva ("drifts") de longas durações e gostaríamos de estimar a taxa de deriva a partir de dados de laboratório.

Admitimos que os giros e suas derivas podem ser adequadamente modelados como processos Gaussiano estacionários correlacionados exponencialmente no tempo, o que fornece um bom modelo para os efeitos de deriva predominantes.

Por outro lado temos que podemos medir instantaneamente a taxa de deriva. Essa hipótese não é muito realista, mas serve para ilustrar a discussão sobre o filtro de Kalman.

A taxa de deriva do giro é o processo de estado  $x(\cdot, \cdot)$  e, em se tratando de um processo exponencialmente correlacionado Gaussiano, corresponde à saída de um "shaping filter" de 1ª ordem

forçado por ruído branco Gaussiano  $w(\cdot, \cdot)$

A frequência de corte do "shaping filter" é colocada como  $1\text{rad}/\text{h}$  o que implica num intervalo de correlação  $T = 1\text{h}$ . A estatística de  $w(\cdot, \cdot)$  é dada por

$$E[w(t)] = 0$$

$$E[w(t)w(t+\tau)] = Q\delta(\tau), \quad Q = 2 \text{ } \sigma^2/\text{h}$$

Com isso, o shaping filter tem suas dinâmicas dadas por:

$$\dot{\hat{x}}(t) = -\left(\frac{1}{T}\right)x(t) + w(t)$$

$$\dot{\hat{x}}(t) = -x(t) + w(t)$$

de onde identificamos

$$F(t) = -1 \quad \text{e} \quad G(t) = +1$$

Admitimos que as medidas sejam tomadas a cada  $0,25\text{ h}$  e modeladas por:

$$z(t_i) = x(t_i) + v(t_i)$$

onde  $v(t_i, \cdot)$  é uma sequência branca Gaussiana com

$$E[v(t)] = 0$$

$$E[v(t_i)v(t_j)] = R\delta_{ij}, \quad R = 0.5 \text{ } \sigma^2/\text{h}^2$$

Desejamos, agora, processar as medidas  $z$  para obter uma estimativa ótima da taxa de deriva do giro  $x(t)$ .

Antes de mais nada verifiquemos que a matriz de transição de estado nesse caso é dada por

$$\Phi(t, t_{i-1}) = \exp[-(t_i - t_{i-1})] = e^{-0.25} \stackrel{\downarrow}{=} 0.78$$

Como admitimos que  $x(\cdot, \cdot)$  é um processo estacionário

podemos determinar a variancia em regime permanente como valor inicial da variancia pois este é o valor que podemos calcular.

A equação geral para a covariância é dada por:

$$\dot{P}(t) = F(t) P(t) + P(t) F^T(t) + G(t) Q(t) G^T(t)$$

que se reduz a:

$$\dot{P}(t) = -\frac{1}{2} P(t) + P(t) (-I) + (I) Q (I)$$

$$\dot{P}(t) = -2P(t) + Q(t) = -2P + Q$$

Em regime permanente  $\dot{P}(t) = 0$  e

$$P = \frac{Q}{2} \Rightarrow P = 1 \text{ } \sigma^2/\text{h}^2$$

Então, antes de realizar qualquer medida temos como condições iniciais

$$\hat{x}(t_0) = \hat{x}_0 = 0 \quad (\text{hipótese})$$

$$P(t_0) = P = 1 \text{ } \sigma^2/\text{h}^2$$

Nesse ponto temos as condições necessárias para montagem do filtro totalmente especificadas. As equações de propagação na forma da matriz de transição serão dadas por:

$$\hat{x}(t_i^-) = \bar{\Phi}(t_i, t_{i-1}^-) \hat{x}(t_{i-1}^+) \Rightarrow \hat{x}(t_i^-) = 0.78 \hat{x}(t_{i-1}^+)$$

$$P(t_i^-) = \bar{\Phi}^2(t_i, t_{i-1}^-) P(t_{i-1}^+) + \int_{t_{i-1}^-}^{t_i^-} \bar{\Phi}^2(t_i, \tau) G^2 Q d\tau$$

$$= 0.78^2 P(t_{i-1}^+) + \int_{t_{i-1}^-}^{t_i} [e^{-(4-\tau)^2}]^2 \cdot 1^2 \cdot 2 d\tau$$

$$= 0.6024 P(t_{i-1}^+) + 2 \int_{t_{i-1}^-}^{t_i} e^{-2(t_i-\tau)} d\tau$$

$$P(t_i^-) = 0.6024 P(t_{i-1}^+) + 0.39$$

Nas representações por equações diferenciais escreverímos quando para o ciclo de propagação:

$$\dot{\hat{x}}(t) = F(t) \hat{x}(t) \Rightarrow \dot{\hat{x}}(t) = -\hat{x}(t)$$

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)Q(t)G(t) \Rightarrow \dot{P}(t) = -2P(t) + 2$$

O ciclo de atualização para este exemplo pode ser escrito:

$$K(t_i^-) = P(t_i^-) H^T(t_i) [H(t_i)P(t_i)H^T(t_i) + R(t_i)]^{-1} = \frac{P(t_i^-) H(t_i)}{H(t_i)P(t_i)H^T(t_i) + R(t_i)}$$

$$K(t_i^-) = \frac{1 \cdot P(t_i^-)}{1 \cdot P(t_i^-) + 0.5} \Rightarrow K(t_i^-) = \frac{P(t_i^-)}{P(t_i^-) + 0.5}$$

$$\hat{x}(t_i^+) = \hat{x}(t_i^-) + K(t_i^-) [z_i - H(t_i) \hat{x}(t_i^-)]$$

$$\Rightarrow \hat{x}(t_i^+) = \hat{x}(t_i^-) + \frac{P(t_i^-)}{P(t_i^-) + 0.5} [z_i - \hat{x}(t_i^-)] \Rightarrow$$

$$P(t_i^+) = P(t_i^-) - K(t_i^-)H(t_i)P(t_i^-) = P(t_i^-) - \frac{P(t_i^-)^2}{P(t_i^-) + 0.5}$$

$$\Rightarrow P(t_i^+) = \frac{0.5 P(t_i^-)}{P(t_i^-) + 0.5}$$

### 7.3. FILTRO DE KALMAN PARA SISTEMAS CONTÍNUOS / CONTINUOS

Vista a representação para sistemas com dinâmica

continuo e observações discretas do filtro de Kalman, os resultados obtidos podem ser facilmente estendidos para o caso de sistemas contínuos com observações contínuas no tempo.

O modelo da dinâmica do sistema é dado por:

$$\dot{x}(t) = F(t)x(t) + G(t)w(t)$$

ou seja, é o mesmo modelo proposto para o caso do filtro contínuo/discreto que acabamos de desenvolver.

O modelo de observações proposto é agora contínuo, modelado por um processo  $z_c(\cdot, \cdot)$  definido por

$$z_c(t) = H(t)x(t) + v_c(t)$$

onde o índice  $c$  se refere a medidas contínuas no tempo e  $v_c(\cdot, \cdot)$  é um ruído branco Gaussiano de média nula assim

$$E[v_c(t)v_c^T(t+\tau)] = R_c(t)\delta(\tau)$$

Admitindo para esse caso as mesmas hipóteses de não correlação entre os processos  $x(\cdot, \cdot)$ ,  $v_c(\cdot, \cdot)$  e  $w(\cdot, \cdot)$ , isto é:

$$E[w(t)v_c^T(\tau)] = 0$$

$$E[x(t)w^T(\tau)] = 0$$

$$E[x(t)v_c^T(\tau)] = 0$$

é possível provar que a média e a matriz de covariâncias, parâmetros suficientes para caracterizar totalmente a função densidade de probabilidade condicionada a todo histórico das medidas, obedecem as seguintes equações diferenciais para sua propagação:

Média:

$$\hat{x}(t) = F(t)\hat{x}(t) + K(t)[z(t) - H(t)\hat{x}(t)]$$

Covariância:

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)Q(t)G^T(t) - P(t)H^T(t)R^{-1}(t)H(t)P(t)$$

Azulho de Kalman

$$K(t) = P(t)H^T(t)R^{-1}(t)$$

Note que a equação da covariância é uma equação completa de Riccati (matricial) e que, no caso contínuo, não há separações entre ciclos de propagação e atualização visto que o filtro funciona com elementos contínuos de informações.

É importante observar, também, que a equação de Riccati para a covariância dos erros independe do processo observado, f.e., para sua implementação precisamos apenas da matriz de covariância dos ruidos  $R_e(t)$ , das observações. Com isso, a equação de Riccati pode ser resolvida e os valores da matriz  $P(t)$  armazenados para cada instante do intervalo de interesse do problema. Da mesma forma, o azulho de Kalman  $K(t)$  pode ser armazenado antes da correspondente informação chegar. Dessa forma, o filtro de Kalman contínuo/contínuo pode ser implementado com as equações de propagação/atualização acima a partir da condição inicial do estado, Gaussian, com estatísticas

$$E[x(0)] = \hat{x}_0$$

$$E[(x(0) - \hat{x}_0)(x(0) - \hat{x}_0)^T] = P_0$$

A grande dificuldade no tratamento do caso contínuo contínuo é a solução da equação de Riccati de covariância que verificamos ser não linear. Apesar de exibir todas as propriedades desejadas de existência, unicidade e continuidade das soluções, equações de Riccati são difíceis de serem integradas em computador por serem muito sensíveis aos passos de integração utilizados. Para tentar evitar a integração desse tipo de equa-

ção, podemos propor a substituição por um sistema equivalente de dimensão  $2n \times n$ , expresso por

$$\tilde{P}(t) = \tilde{U}(t) \tilde{V}^{-1}(t)$$

onde  $\tilde{U}(t)$  e  $\tilde{V}(t)$  são matrizes  $n \times n$  que satisfazem o sistema linear homogêneo:

$$\begin{bmatrix} \dot{\tilde{U}}(t) \\ \dot{\tilde{V}}(t) \end{bmatrix} = \begin{bmatrix} -F(t) & | G(t) Q(t) G^T(t) \\ H^T(t) R_c(t) H(t) & -F^T(t) \end{bmatrix} \begin{bmatrix} \tilde{U}(t) \\ \tilde{V}(t) \end{bmatrix}$$

com a condição inicial

$$\begin{bmatrix} \tilde{U}(t_0) \\ \tilde{V}(t_0) \end{bmatrix} = \begin{bmatrix} P_0 \\ I \end{bmatrix}$$

Demonstra-se facilmente que a condição inicial é satisfeita já que

$$P(t_0) = \tilde{U}(t_0) \tilde{V}^{-1}(t_0) = P(t_0) I^{-1} = P_0$$

Diferenciando,  $\dot{P}(t) = \dot{\tilde{U}}(t) \tilde{V}^{-1}(t)$

$$\dot{P}(t) = \ddot{\tilde{U}}(t) \tilde{V}^{-1}(t) + \dot{\tilde{U}}(t) \frac{d}{dt} \tilde{V}^{-1}(t)$$

Lembrando que  $\frac{d}{dt} \tilde{V}^{-1}(t) = -\tilde{V}^{-1} \dot{\tilde{V}} \tilde{V}^{-1}$  ( $\tilde{V}^{-1} \tilde{V} \tilde{V}^{-1}$ )

$$\dot{P} = \ddot{\tilde{U}} \tilde{V}^{-1} - \dot{\tilde{U}} \tilde{V}^{-1} \dot{\tilde{V}} \tilde{V}^{-1}$$

e substituindo  $\dot{\tilde{U}}$  e  $\dot{\tilde{V}}$  das expressões acima:

$$\dot{P} = (F \tilde{U} + G Q G^T \tilde{V}) \tilde{V}^{-1} - \tilde{U} \tilde{V}^{-1} (H^T R_c^{-1} H \tilde{U} - F^T \tilde{V}) \tilde{V}^{-1}$$

$$\dot{P} = (F U V^{-1} + G Q G^T V V^{-1} - U V^{-1} H^T R_c^{-1} H U V^{-1}, U V^{-1} F V V^{-1})$$

$$\dot{P} = F P + G Q G^T - P H^T R_c^{-1} H P + P F^T$$

que é a equação de Riccati para a covariância.

Exemplo:

Voltamos ao exemplo do giro em teste que desenvolvemos num sistema contínuo/discreto e admitimos agora que as medidas estejam disponíveis continuamente, o que pode ser traduzido por:

$$z_c(t) = x(t) + v_c(t)$$

onde  $v_c(\cdot, \cdot)$  é um ruído branco Gaussiano com média nula e variação

$$E[v_c(t)v_c(t+\tau)] = R_c(t)\delta(\tau)$$

Para aproximar razoavelmente o sistema contínuo do sistema discreto desenvolvido, colocarmos para a intensidade  $R_c(t)$  do ruído branco contínuo

$$R_c(t) = R_c = [R_c(t_i)\Delta t] = (0.5 \text{ } \Omega^2/\text{Hz}) \cdot (0.25 \text{ s})$$

$$R_c = 0.125 \text{ } \Omega^2/\text{s}$$

Montamos agora o estimador lembrando que  $F(t) = -I$ ,  $G(t) = I$ ,  $H(t) = I$

\* Para a média:

$$\hat{x} = F \cdot x + K(t)[z(t) - H \hat{x}]$$

onde agora o Gráfico de Kfazendo é:

$$K(t) = P(t) H^T(t) P_c^{-1}(t) \Rightarrow K(t) = \frac{P(t)}{R_c(t)} \Rightarrow K(t) = \frac{P(t)}{0.125}$$

$$K(t) = 8P(t)$$

resultando

$$\Rightarrow \dot{\hat{x}}(t) = -\hat{x}(t) + 8P(t)[z(t) - \hat{x}(t)]$$

que deve ser integrada com a condição inicial  $\hat{x}(0) = 0$

\* Para a covariância:

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)Q(t)G^T(t) - P(t)H^T(t)P_c^{-1}(t)H(t)P(t)$$

$$\dot{\underline{P}} = -1 \cdot \underline{P} + 2 \cdot (-1) + 1 \cdot 2 \cdot 1 - \underline{P} \cdot 1 \cdot \underline{B} \cdot 1 \cdot \underline{B}$$

$$\dot{\underline{Q}} = -2\underline{P} + 8\underline{P}^2 + 2$$

A solução destas equações é dada por:

$$\underline{P}(t) = \frac{\sqrt{17} + \operatorname{tgh}(\sqrt{17}[t-t_0])}{\sqrt{17} + 9\operatorname{tgh}(\sqrt{17}[t-t_0])}$$

Integrada a partir da condição inicial  $\underline{P}(t_0) = \underline{P} = 1 \text{ m}^2/\text{Hz}$

#### 7.4. FILTRO DE KALMAN PARA SISTEMAS DISCRETOS / DISCRETOS

Apenas com o intuito de completar as idéias referentes às formas do filtro de Kalman para sistemas lineares, vamos descrever as equações do filtro para o caso de um sistema dinâmico discreto com modelo discreto de observações.

A dinâmica do sistema é então dada por uma equação de diferenças:

$$\underline{x}(k+1) = \underline{\Phi}(k+1, k) \underline{x}(k) + \underline{\Gamma}_E(k) \underline{w}(k)$$

onde  $\underline{\Phi}(k+1, k)$  é a matriz de transição de estado,  $\underline{\Gamma}(k)$  uma matriz ( $n \times s$ ) conhecida e  $\underline{w}(k)$  uma sequência branca Gaussiana de média nula e covariância

$$E[\underline{w}(k) \underline{w}^T(k)] = \underline{Q}(k) \delta_{kk}$$

O modelo de observações é dado por:

$$\underline{z}(k) = \underline{H}(k) \underline{x}(k) + \underline{v}(k)$$

onde  $H(k)$  é uma matriz conhecida e  $V(k)$  é uma sequência branca Gaussiana caracterizada por média nula e matriz de covariâncias:

$$E[V(k)V(k)^T] = R(k) \delta_{kk}$$

Supondo que o estado inicial seja Gaussiano com estatísticas

$$E[x(0)] = \bar{x}_0$$

$$E[(x(0) - \bar{x}_0)(x(0) - \bar{x}_0)^T] = P_0$$

O filtro de Kalman discreto/discreto corresponde a uma sucessão de propagações e atualizações onde reservando o subscrito - para os valores na propagação e ^ para os valores imediatamente após a atualização, temos:

- Ciclo de Propagação:

$$\text{Média: } \bar{x}(k+1) = \bar{\Phi}(k+1, k) \hat{x}(k) \quad (4)$$

Covariância:

$$\bar{P}(k+1) = \bar{\Phi}(k+1, k) \hat{P}(k) \bar{\Phi}^T(k+1, k) + \Gamma(k) Q(k) \Gamma^T(k) \quad (5)$$

- Ciclo de Atualização:

$$\text{Média: } \hat{x}(k+1) = \bar{x}(k+1) + K(k+1) [z(k+1) - H(k+1) \bar{x}(k+1)] \quad (5)$$

Covariância:

$$\hat{P}(k+1) = \bar{P}(k+1) - K(k+1) H(k+1) \bar{P}(k+1) \quad (3)$$

Gráfico de Kalman:

$$K(k+1) = \bar{P}(k+1) H(k+1) [H(k+1) \bar{P}(k+1) H^T(k+1) + R(k+1)]^{-1} \quad (2)$$

Note que as equações acima correspondem às equações obtidas no filtro contínuo/discreto quando se usa a representação de matriz de trazidas de estado, fato este que deverá ser esperado.

## 7.5. TÉCNICAS PARA IMPLANTAÇÃO EM COMPUTADOR

Desenvolvidas as diversas versões do filtro de Kaban para sistemas lineares poderia ser tentar a sua implantação em computador. Essa tarefa, aparentemente fácil, pode, dependendo do problema, representar um grande desafio mesmo para um competente programador. Em vista disso, a literatura recente traz propostas, para se tentar minorar os efeitos indesejáveis que surgem quando da implementação do filtro, de modo sistemático. Algumas dessas propostas se mostraram, ao longo de um intervalo de tempo relativamente curto, alternativas extremamente viáveis e hoje integram o rol dos procedimentos usuais para implantação. Este item tem a intenção de apresentar essas técnicas, com ênfase aos chamados métodos de fotografia.

## CAPÍTULO 8 - APLICAÇÕES DA TEORIA LINEAR A ESTIMAÇÃO DE SISTEMAS NÃO LINEARES

Neste capítulo iremos abordar o problema de filtragem e previsão para o caso de sistemas não lineares. Estimação de sistemas não lineares é ainda um problema cuja solução é mediada em que o desenvolvimento de uma teoria única é praticamente impossível. Apesar para citar as dificuldades inerentes à estimação de sistemas não lineares não seretos, deve ser dito que:

a) O problema é de dimensão infinito, visto que as equações de evolução da média e da matriz de covariância dos erros, condicionadas à história das observações, dependem, em geral, de todos os momentos da função densidade condicionadora;

b) Mesmo que a condição inicial do estado do sistema seja Gaussiana, como as transformações aplicadas são não lineares, não se pode garantir que qualquer estado subsequente seja Gaussiano. No caso linear, o que era Gaussiano permanece Gaussiano em vista da linearidade das transformações. Exemplo típico é o caso de integração estocástica proposta para os sistemas lineares que não vale mais para o caso não linear. Para o tratamento dos sistemas não lineares é necessária a colocação de um novo formalismo matemático, chamado cálculo de Itô;

c) Dificilmente existe uma solução analítica que otimiza um filtro mesmo nos casos não lineares mais simples, ao contrário do que acontece no caso linear onde o filtro de Kalman é a solução exata para o critério proposto e faz as hipóteses dadas.

Para se tentar contornar essas dificuldades existem duas linhas de abordagem. A primeira é caracterizada pelos

filtrros ditos sub-ótimos onde se usa como modelo no filtro uma aproximação linearizada do sistema real e se aplica sobre o modelo linearizado a teoria linear, principalmente as técnicas de Kalman. Essa abordagem gera uma série de técnicas simples, de grande utilidade em Engenharia, mas que podem apresentar problemas de convergência dependendo de como é feita a aproximação linearizada. A segunda linha, de desenvolvimento mais recente, é caracterizada pela aplicação das mais modernas técnicas numéricas de discretização para resolver as equações de propagação da função densidade condicionada, que, para o caso contínuo/contínuo, formam um sistema de equações integrais-diferenciais parciais. Apesar para cito, métodos de séries especiais e métodos de elementos finitos têm sido utilizados para tutor a solução por discretização nos sistemas não lineares.

No que se segue, desenvolvemos algumas técnicas sub-ótimas para aplicações a sistemas não lineares.

### 8.1. FILTRO ESTENDIDO DE KALMAN

Usaremos para obtenção das diversas formas do filtro de Kalman aplicado a sistemas não lineares a formulação modelo dinâmico do sistema contínuo/modelos de observações discrete, pois, como já afirmamos antes, este é a formulação mais comum em problemas de Engenharia. As versões dinâmico discreto/observações discrete e dinâmico contínuo/observações contínuas podem ser obtidas pela aplicação do mesmo desenvolvimento aqui feito para dinâmica contínua/observações discrete e encontram-se resumidas por exemplo em Gelb (1974).

Admitamos, então, que o nosso sistema dinâmico é des-

rito por uma equação diferencial estocástica vetorial

$$\dot{x}(t) = f(x(t), t) + G(t) w(t)$$

onde  $w(\cdot, \cdot)$  é um processo branco Gaussiano com média nula e

$$E[w(t) w(t)^T] = Q(t) \delta(t)$$

$f(x(t), t)$  é uma vetor de funções, de dimensão  $n$ , todas conhecidas, do vetor ( $n \times 1$ ) de estado  $x(t)$ .

$G(t)$ , por sua vez, é uma matriz ( $n \times s$ ) de funções conhecidas do tempo.

Supomos que o estado inicial do sistema é Gaussiano com estatística

$$E[x(t_0)] = \hat{x}_0$$

$$E[(x(t_0) - \hat{x}_0)(x(t_0) - \hat{x}_0)^T] = P_0$$

Agora, supomos ser capazes de gerar uma trajetória de referência ou trajetória nominal, determinística,  $\bar{x}(t)$  com suas deas condições iniciais  $\bar{x}(t_0)$ , satisfazendo:

$$\dot{\bar{x}}(t) = f(\bar{x}(t), t)$$

Definimos, então:

$$\delta x(t) = x(t) - \bar{x}(t)$$

como sendo a perturbação ou a variação da relação à trajetória em cada instante. Então  $\delta x(\cdot, \cdot)$  é um processo estocástico satisfazendo a equação diferencial:

$$\dot{\delta x} = \frac{d}{dt}(\delta x) = f(x(t), t) - f(\bar{x}(t), t) + G(t) w(t)$$

com condições iniciais Gaussiana com estatística dada por

$$E[\delta x(t_0)] = \bar{x}_0 - \bar{x}(t_0)$$

$$E[\delta x(t_0) \delta \dot{x}(t_0)] = ?_0$$

A trajetória de referência definida acima servirá tanto para a obtenção do filtro estendido como para a obtenção do filtro linearizado de Kalman, e, por isso, este definida de maneira genérica.

Fazemos agora a expansão de  $f(x,t)$  em torno da trajetória de referência

$$f(x(t), t) = f(\bar{x}(t), t) + F[t, \bar{x}(t_0)] \delta x(t) + \sigma(2)$$

A hipótese básica a ser feita é que  $\bar{x}(t)$  é suficientemente próximo de  $x(t)$  em  $[t_0, t_1]$ , por enquanto nesse intervalo de interesse, de maneira que os termos de 2.º ordem ou de ordem maior sejam desprezíveis. Com isso:

$$f(x(t), t) = f(\bar{x}(t), t) + F[t, \bar{x}(t_0)] \delta x(t)$$

onde

$$F[t, \bar{x}(t_0)] = \left[ \frac{\partial f(\bar{x}(t), t)}{\partial x_j} \right]_{\bar{x}(t)}$$

é a matriz ( $n \times n$ ) de derivadas parciais avaliadas ao longo da trajetória de referência.

Substituindo na expressão de  $\delta \dot{x}$ :

$$\delta \dot{x}(t) = f'(\bar{x}(t), t) + F[t, \bar{x}(t_0)] \delta x(t) - f(\bar{x}(t), t) + G(t) w(t)$$

$$\delta \dot{x}(t) = F[t, \bar{x}(t_0)] \delta x(t) + G(t) w(t)$$

agora uma expressão linear para a qual temos os modelos de propagação. Note nos argumentos de  $F$  a dependência da equação

com relação à trajetória nominal escolhida.

Abandonamos, por enquanto, a equação variacional e verificaremos o que fazer com o modelo não linear de observações, dado por:

$$z(t_i) = h(x(t_i), t_i) + v(t_i)$$

Linearizemos  $z(t_i)$  em torno do valor fornecido pela trajetória de referência,  $\bar{x}(t_i)$ . Para tanto, definimos uma observação perturbada

$$\bar{z}(t_i) = h(\bar{x}(t_i), t_i)$$

Com isso, definimos uma observação perturbada

$$\delta z(t_i) = z(t_i) - \bar{z}(t_i) = h(x(t_i), t_i) - h(\bar{x}(t_i), t_i) + v(t_i)$$

Desenvolvendo em série em torno de  $h(\bar{x}(t_i), t_i)$ :

$$\delta z(t_i) = h(\bar{x}(t_i), t_i) + \left[ \frac{\partial h}{\partial x} \right]_{\bar{x}(t_i)} (x(t_i) - \bar{x}(t_i)) - h(\bar{x}(t_i), t_i) + \sigma(2) + v(t_i)$$

Desprezando os termos de 2ª ordem e de maior ordem chegará a:

$$\delta z(t_i) = \underbrace{\left[ \frac{\partial h(\bar{x}(t_i), t_i)}{\partial x} \right]}_{\bar{x}(t_i)} \delta x(t_i) + v(t_i)$$

$$\delta z(t_i) = H[t_i, \bar{x}(t_i)] \delta x(t_i) + v(t_i)$$

um modelo linear de observações onde a notação da matriz  $H$  expõe explicitamente a dependência em relação à trajetória de referência escolhida.

Resumindo o que fizemos até aqui, partiu de um modelo não linear:

$$\dot{x}(t) = f(x(t), t) + G(t) w(t)$$

$$z(t_i) = h(x(t_i), t_i) + v(t_i)$$

com estado inicial  $x(t_0) \sim N(\bar{x}_0, P_0)$  e chegamos a um sistema linearizado ou perturbado:

$$\delta \dot{x}(t) = F(t_i; \bar{x}(t_0)) \delta x(t) + G(t) w(t)$$

$$\delta z(t_i) = H(t_i; \bar{x}(t_0)) \delta x(t_i) + v(t_i)$$

com estado inicial  $\delta x(t_0) \sim N(\bar{x}_0 - \bar{x}(t_0), P_0)$

O filtro estendido de Kalman corresponde à aplicação da teoria linear ao sistema linearizado acima para uma particular escolha da trajetória de referência. Nesse caso, a escolha óbvia para condição inicial na trajetória de referência é:

$$\bar{x}(t_0) = \bar{x}_0$$

Seu o ruído de medida não, a estimativa de média desse estado é propagada até  $t_i$ , instante em que vai ser incorporada a 1ª observação, através da aproximação de 1ª ordem:

$$\hat{x}(t) = f(\hat{x}(t), t) \quad \hat{x}(t_0) = \hat{x}_0$$

A covariância é propagada como no caso linear usando valores calculados na trajetória de referência, isto é:

$$\hat{P}(t) = F(\hat{x}(t), t) P(t) + P(t) F^T(\hat{x}(t), t) + G(t) Q(t) G^T(t), P(t)$$

Propagadas estas equações até  $t_i$ , teremos como estimativa a priori no ciclo de atualização  $\hat{x}(t_i)$  e  $\hat{P}(t_i)$  fornecidas pelas equações diferenciais acima. A incorporação dessas novas medidas se faz agora usando o mesmo esquema do caso linear com valores calculados na trajetória de referência. Teremos então:

$$\text{Média: } \hat{x}(t_i^+) = \hat{x}(t_i^-) + K(t_i) [z(t_i) - h(\hat{x}(t_i^-))]$$

$$\text{Covariância: } P(t_i^+) = P(t_i^-) - K(t_i) H(t_i, \hat{x}(t_i^-)) P(t_i^-)$$

$$\text{Grau de Kalman: } K(t_i) = P(t_i^-) H(t_i, \hat{x}(t_i^-)) [H(t_i, \hat{x}(t_i^-)) P(t_i^-) H^T(t_i, \hat{x}(t_i^-)) + R_{u_i}]^{-1}$$

$$\text{onde } R_{u_i} = E[\nu(t_i) \nu(t_i)^T]$$

Note que a saída do ciclo de atualização no instante  $t_1$  nos fornece as melhores estimativas  $\hat{x}(t_1^+)$  e  $P(t_1^+)$ , da média e da covariância. É então intuitivo que essas sirjam os valores largados como condições iniciais do ciclo de propagação entre  $t_1$  e  $t_2$ .

Teremos, então, para o filtro estendido de Kalman as seguintes relações de recorrência:

Ciclo de Propagação:

$$\text{Média: } \dot{\hat{x}}(t) = f(\hat{x}(t), t) \quad , \text{ cond. inicial } \hat{x}(t_k^+), t_k^+ < t < t_{k+1}^-$$

$$\text{Covariância: } \dot{P}(t) = F(\hat{x}(t), t) P(t) + P(t) F^T(\hat{x}(t), t) + G(t) Q(t) G^T(t) \\ \text{cond. inicial } P(t_k^+), t_k^+ < t < t_{k+1}^-$$

Ciclo de Atualização:

$$\text{Média: } \hat{x}(t_k^+) = \hat{x}(t_k^-) + K(t_k) [z(t_k) - h(\hat{x}(t_k^-))]$$

$$\text{Covariância: } P(t_k^+) = P(t_k^-) - K(t_k) H(t_k, \hat{x}(t_k^-)) P(t_k^-)$$

$$\text{Grau de Kalman: } K(t_k) = P(t_k^-) H(t_k, \hat{x}(t_k^-)) [H(t_k, \hat{x}(t_k^-)) P(t_k^-) H^T(t_k, \hat{x}(t_k^-)) + R(t_k)]^{-1}$$

## 7.2. FILTRO LINEARIZADO DE KALMAN

Algumas vezes nosso conhecimento sobre o problema que estamos tratando (ou outras considerações de ordem computacional) nos permite gerar uma boa trajetória de referência para a evolução do sistema, i.e., temos condições de garantir que o nosso sistema esteja sempre muito próximo da uma trajetória conhecida. Nesse caso, temos meios de simplificar o algoritmo do filtro estendido, trocando-o por um mais simples. Este último é chamado filtro linearizado de Kalman.

Admitindo válidas todos os desenvolvimentos feitos para a obtenção do filtro estendido, suponhamos que esteja à nossa disposição uma trajetória de referência  $\bar{x}(t)$ , válida para todo intervalo de interesse do problema,  $[t_0, t_f]$ .

Lógicamente, nesse caso, não haveria necessidade de se auxiliar as matrizes e parâmetros sobre as estimativas atualizadas, podendo isso ser feito sobre a trajetória de referência.

Desse maneira, definindo

$$F(\bar{x}(t), t) = \left[ \frac{\partial f(x(t), t)}{\partial x} \right]_{\bar{x}(t)}$$

$$H(\bar{x}(t_e), t_e) = \left[ \frac{\partial h(x(t_e), t_e)}{\partial x} \right]_{\bar{x}(t_e)}$$

onde  $\bar{x}(t_e)$  é a trajetória nominal disponível, o filtro linearizado de Kalman consiste em:

Ciclo de propagação:

Média:  $\hat{x}(t) = \overset{\circ}{f}(\bar{x}(t), t) + F(\bar{x}(t), t) [\hat{x}(t) - \bar{x}(t)]$

cond. inicial  $\hat{x}(t_0) = \bar{x}(t_0) = \hat{x}_0$  em  $t_0$

Covariância:  $\hat{P}(t) = F(\bar{x}(t), t) P(t) F(\bar{x}(t), t)^T + G(t) Q(t) G(t)^T$

cond. inicial  $P(t_0) = P_0$  em  $t_0$

Ciclo de atualização:

$$\text{Média: } \hat{x}(t_k^+) = \hat{x}(t_k^-) + K(t_k)[z(t_k) - h(\bar{x}(t_k), t_k) - H(\bar{x}(t_k), t_k)\hat{x}(t_k^-) - \bar{x}_k]$$

$$\text{Covariância: } P(t_k^+) = P(t_k^-) - K(t_k) H(\bar{x}(t_k), t_k) P(t_k^-)$$

$$\text{Ganhos de Kalman: } K(t_k) = P(t_k^-) H(\bar{x}(t_k), t_k) [H(\bar{x}(t_k), t_k) P(t_k^-) H^T(\bar{x}(t_k), t_k) + R(t_k)]^{-1}$$

É importante notar que as simplificações introduzidas pelo filtro linearizado são grandes, na medida em que sendo as matrizes  $f$  e  $H$  avalizadas sobre a trajetória de referência, as equações de projeção e atualização da matriz de covariância podem ser resolvidas "off-line" e os valores de  $P(t)$  armazenados. Com isso, o ganho de Kalman  $K(t)$  pode também ser pré-calculado e armazenado, diminuindo muito a carga computacional. O mesmo não é válido para o filtro estendido já que a trajetória de referência só pode ser gerada depois de cada ciclo de atualização. Em compensação, a precisão obtida com o filtro estendido é muito melhor do que a conseguida com o filtro linearizado.

Existem outras técnicas de aplicação da teoria linear à estimativa de sistemas não lineares, como o filtro estendido e iterado de Kalman ou filtro de Denham-Penes, mas a carga computacional é tão maior que dificilmente são aplicadas à solução de problemas de Engenharia.

### 7.3. ESTIMAÇÃO SIMULTÂNEA DE ESTADO E PARÂMETROS: FILTRO DE SCHMIDT-KALMAN

Em muitos casos de aplicações tanto o sistema dinâmico como o modelo de observações podem depender de certos parâmetros cujos valores não são conhecidos precisamente. Pode-

sc, então, usar a estrutura do filtro para estimar também esses parâmetros desconhecidos. O problema resultante é tipicamente não linear nos casos de aplicação mas comumente o desenvolvimento usando um sistema linear para depois fazer a passagem para o sistema não linear.

O parâmetros imprecisamente conhecidos sô considerados como variáveis aleatórias com estatística a priori conhecida. Consideremos primeiro um sistema linear contendo parâmetros incertos se for nez discreto/discreto, i.e., modelo dinâmico discreto / modelos de observações discrete. A alocação nas formas contínuo/discreto ou contínuo/contínuo é imediata.

Seja:

$$x(t_{k+1}) = \Phi(t_{k+1}, t_k) x(t_k) + \psi(t_{k+1}, t_k) m + \Gamma(k) w(k)$$

$$z(t_k) = H(t_k) x(t_k) + N(t_k) p + v(t_k)$$

onde os parâmetros a serem estimados são dados pelo vetor de v.z's  $m$ , ditos parâmetros dinâmicos, e pelo vetor de v.z's  $p$ , ditos parâmetros de medida.

É admitido o priori uma estatística para  $m$  e  $p$  dada por:

$$E[m] = 0$$

$$E[mm^T] = M_0$$

$$E[p] = 0$$

$$E[pp^T] = W_0$$

Supõe-se, também, que  $m$ ,  $p$ ,  $x(t_0)$ ,  $w(\cdot, \cdot)$  e  $v(\cdot, \cdot)$  sô não correlacionadas.

Os parâmetros  $m$  e  $p$  podem ser considerados como

séries de sistemas dinâmicos modelados por:

$$m(t_{k+1}) = m(t_k) \quad \text{cond. inicial } m(t_0) = m_0$$

$$\dot{p}(t_{k+1}) = \dot{p}(t_k) \quad \text{cond. inicial } p(t_0) = p_0$$

Com isso, definimos um vetor de estados aumentado através de:

$$\bar{x}(t_k) = \begin{bmatrix} x(t_k) \\ m(t_k) \\ \dot{p}(t_k) \end{bmatrix}$$

e podemos combinar todas as equações num sistema linear contínuo

$$\bar{x}(t_{k+1}) = \begin{bmatrix} \Phi(t_{k+1}, t_k) & \Psi(t_{k+1}, t_k) & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \bar{x}(t_k) + \begin{bmatrix} \Gamma(t_k) \\ 0 \\ 0 \end{bmatrix} w(t_k)$$

com as observações dadas por:

$$z(t_k) = \begin{bmatrix} h(t_k) & 0 & \bar{N}(t_k) \end{bmatrix} \bar{x}(t_k) + v(t_k)$$

Para este sistema podemos agora aplicar a forma linear do filtro de Kalman, gerando estimativas, por exemplo, para a média  $\bar{x}$  forma:

$$\hat{\bar{x}}(t_k^+) = \begin{bmatrix} \hat{x}(t_k^+) \\ \hat{m}(t_k^+) \\ \hat{\dot{p}}(t_k^+) \end{bmatrix}$$

O que permite uma combinação rápida e direta para estimativas dos parâmetros.

No entanto, a carga computacional adicional é uni-

to grande, o que torna o esquema instável, muitas vezes, inviável. Por outro lado, as incertezas nos parâmetros não podem ser desprezadas pois podem causar a divergência do filtro.

Uma alternativa, proposta por Schmidt, é levar em conta o efeito das incertezas nos parâmetros calculando de alguma forma o quanto essa incerteza degrada (piora) a estimativa do estado, sem estiver os parâmetros propriamente ditos. O filtro resultante é claramente sub-ótimo em relação ao esquema anterior já que os parâmetros não estão sendo estimados. Se estivessem sendo estimados, a incerteza sobre os parâmetros diminuiria, permitindo que a incerteza sobre o estado também descrese.

O filtro de Schmidt - Kalman é obtido primeiramente escrevendo-se o filtro linear padrão para o sistema aumentado e depois descartando as equações de estimativa dos parâmetros. A matriz de covariância dos erros  $\hat{P}$  para o sistema aumentado tem a forma:

$$\hat{P} = \begin{bmatrix} P & C_u & C_p \\ C_u^T & M_0 & 0 \\ C_p^T & 0 & N_0 \end{bmatrix}$$

já que estamos admitindo  $u$  e  $p$  não correcionadas. Na matriz acima  $P$  é a matriz de covariância usual do estado

$$P = E[(\bar{x}(t_k) - \hat{x}(t_k)) (\bar{x}(t_k) - \hat{x}(t_k))^T]$$

$$C_u = E[(x(t_k) - \hat{x}(t_k)) u^T]$$

$$C_p = E[(x(t_k) - \hat{x}(t_k)) p^T]$$

Então, demonstra-se que se mutuamente  $P$  e  $\hat{P}$  forem invariantes no tempo, podemos montar um filtro subótimo com

forma:

a) Ciclo de Propagação

\* Média:

$$\hat{x}(t_{k+1}^-) = \underline{\Phi}(t_{k+1}, t_k) \hat{x}(t_k^+)$$

\* Covariância do estado:

$$\begin{aligned} P(t_{k+1}^-) = & \underline{\Phi}(t_{k+1}, t_k) P(t_k^+) \underline{\Phi}^T(t_{k+1}, t_k) + \underline{\Phi}(t_{k+1}, t_k) C_m(t_k^+) \underline{\Phi}^T(t_{k+1}, t_k) + \\ & + \underline{\Psi}(t_{k+1}, t_k) C_m(t_k^+) \underline{\Psi}^T(t_{k+1}, t_k) + \underline{\Psi}(t_{k+1}, t_k) \underline{\Sigma}_0 \underline{\Psi}^T(t_{k+1}, t_k) + \\ & + \Gamma(k) Q(k) \Gamma^T(k) \end{aligned}$$

\* Covariância dos parâmetros dinâmicos

$$C_m(t_{k+1}^-) = \underline{\Phi}(t_{k+1}, k) C_m(t_k^-) + \underline{\Psi}(t_{k+1}, t_k) \underline{\Sigma}_0$$

\* Covariância dos parâmetros de medida

$$C_p(t_{k+1}^-) = \underline{\Phi}(t_{k+1}, t_k) C_p(t_k^+)$$

b) Ciclo de Atualização

\* Gráfico de Kalman:

$$K(t_{k+1}) = [P(t_{k+1}^-) H^T(t_{k+1}) + C_p(t_{k+1}^-) N^T(t_{k+1})] Y^{-1}(t_{k+1}^-)$$

onde

$$\begin{aligned} Y(t_{k+1}^-) = & U(t_{k+1}) P(t_{k+1}^-) H^T(t_{k+1}) + U(t_{k+1}) C_p(t_{k+1}^-) N^T(t_{k+1}) + \\ & + N(t_{k+1}) C_p^T(t_{k+1}^-) H^T(t_{k+1}) + N(t_{k+1}) \underline{\Sigma}_0 N^T(t_{k+1}) + R(t_{k+1}) \end{aligned}$$

\* Média:

$$\hat{x}(t_{k+1}^+) = \hat{x}(t_{k+1}^-) + K(t_{k+1}) [z(t_{k+1}) - H(t_{k+1}) \hat{x}(t_{k+1}^-)]$$

\* Covariância do estado:

$$\begin{aligned} P(t_{k+1}^+) = & P(t_{k+1}^-) - K(t_{k+1}) [U(t_{k+1}) P(t_{k+1}^-) + \\ & + N(t_{k+1}) C_p(t_{k+1}^-)] \end{aligned}$$

\* Covariância dos parâmetros dinâmicos

$$C_m(t_{k+1}^+) = C_m(t_{k+1}^-) - K(t_{k+1}) H(t_{k+1}) C_m(t_{k+1}^-)$$

\* Covariância dos parâmetros de medida:

$$C_p(t_{k+1}^+) = C_p(t_k^-) - \Delta t (t_{k+1}) [H(t_{k+1}) C_p(t_k^-) + N(t_k^-) W_0]$$

Note que  $M_0$  e  $\bar{W}_0$  são matrizes constantes durante os ciclos de estimação e propagação.

O filtro de Schmeidt-Kalman envolve, então, formas recursivas para as covariâncias  $C_M$  e  $C_p$ , além das equações usadas para  $\hat{x}$  e  $P$ , o que aumenta consideravelmente a carga computacional em relação ao filtro padrão. No entanto, em relação ao esquema com estimativas explícitas dos parâmetros, a carga é muito menor, o que torna esse tipo de filtro extremamente eficiente quando não há necessidade de se conhecer bem os parâmetros.

Obtida a forma do filtro para o sistema linear nos estados e nos parâmetros, a teoria linear pode ser aplicada como aproximação para os sistemas não-lineares. Este é o caso de interesse prático, como veremos a seguir.

Se escrevermos o sistema não-linear na forma:

$$\dot{\hat{x}}(t) = f(x(t), u, t) + G(t) w(t)$$

com estado inicial Gaussiano,  $N \sim (\hat{x}_0, P_0)$  e com modelo não linear de observações

$$\hat{z}(t_k) = h(x(t_k), p, t_k) + v_k$$

e admitirmos válidas as mesmas hipóteses feitas na obtenção do filtro estendido, verificamos que o importante caso em que:

$$\dot{\hat{x}}(t) = f(u, t)x(t) + G(t) w(t)$$

$$\hat{z}(t_k) = H(p, t_k)x(t_k) + v(t_k)$$

ié, onde os parâmetros das matrizes do sistema linear são desconhecidos, recai na formulação anterior.

A obtenção de um filtro estendido nesse caso é feita

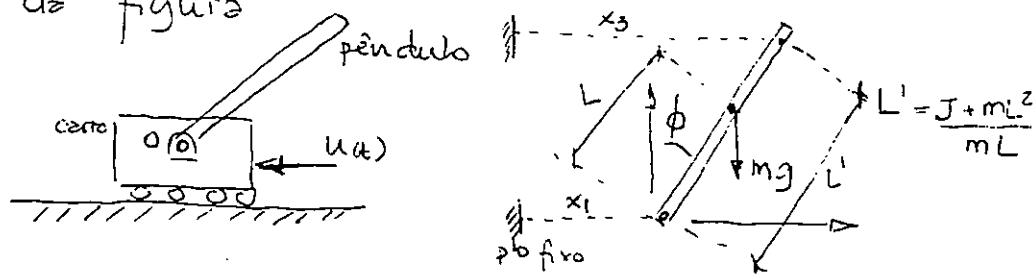
# PMCTSG - INTRODUÇÃO À TEORIA DE ESTIMAÇÃO

## SÉRIE DE EXERCÍCIOS

1. As equações de movimento linearizadas de um pêndulo invertido, colocadas na forma vetorial, são dadas por:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -c/m & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -g/L & 0 & g/L & 0 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ 1/m \\ 0 \\ 0 \end{bmatrix} u(t)$$

onde  $\mathbf{x}(t) = [x_1(t), x_2(t); x_3(t); x_4(t)]^T$  tem o significado da figura



Nesse caso,  $M$  é a massa do carro,  $c$  é o coeficiente de atrito carro-solo,  $u(t)$  é uma força de controle,  $m, L$  e  $J$  são a massa, o semi-comprimento e o momento de inércia do pêndulo.

Suponha que a variável observada seja o ângulo  $\phi$  entre o pêndulo e a vertical, i.e.,

$$y_1(t) = \left( -\frac{1}{L}, 0, \frac{1}{L}, 0 \right) \mathbf{x}(t)$$

e considere o problema de projetar um observador para esse sistema (identidade ou de ordem reduzida). Pede-se:

- 2) Mostre que é impossível achar um observador essencial

asintoticamente estável. Justifique.

b) Mostre que, se elém de  $\phi$ , mede-se o deslocamento do carro, i.e.  $y_2(t) = (1 \ 0 \ 0 \ 0) x(t)$

é adicionado ao vetor de variáveis observadas, então pode-se obter um observador assintoticamente estável.

c) Estabeleça um critério e projete o observador de b).

2. a) Determinar o valor  $b$  para que a função

$$f(x) = \begin{cases} bx(1-x) & 0 \leq x \leq 1 \\ 0 & \text{toda outra parte} \end{cases}$$

seja uma função densidade de probabilidade.

b) Nesse caso, qual a probabilidade da v.a. estar entre  $2 \leq x \leq 5$ ?

c) Idem para  $-\infty < x \leq 0.6$ ?

3. A função densidade exponencial é definida através de

$$f(x) = \begin{cases} q e^{-qx} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad \text{com } q > 0$$

a) Mostre que a função distribuição de probabilidade é

$$F(x) = \begin{cases} 1 - e^{-qx} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

b) Usando integração por partes, mostre que a média de uma v.a.  $X$  com essa distribuição é  $E[X] = \frac{1}{q}$

c) Usando integração por partes,ache  $E[X^2]$

d) Com os resultados de 3b e 3c, determine a variância de uma v.a. com essa distribuição.

4. a) Determine o valor de  $b$  que torna

$$f(u,v) = \begin{cases} b e^{-2u-3v} & u>0, v>0 \\ 0 & \text{todas outras partes} \end{cases}$$

uma função densidade conjunta de duas v.a.  $X$  e  $Y$

b) Calcule as médias  $E[X]$  e  $E[Y]$  por integração por partes.

5. Seja  $(x_1, x_2, \dots, x_n)$  um conjunto de v.a's com a mesma média  $\mu$  e mesma variância  $\sigma^2$  e tais que  $E[(x_i - \mu)(x_j - \mu)] = 0$   $i \neq j$ . Sejam as estimativas  $\hat{\mu}$  e  $s^2$  de  $\mu$  e de  $\sigma^2$  dadas por

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu})^2$$

a)  $\hat{\mu}$  é um estimador não viésado de  $\mu$ ?

b) Mostre que  $n(\hat{\mu} - \mu) = \sum_{i=1}^n (x_i - \mu)$

c) Mostre que  $\left[ \sum_{i=1}^n (x_i - \mu) \right]^2 = \sum_{i=1}^n (x_i - \mu)^2 + \sum_{i \neq j}^n (x_i - \mu)(x_j - \mu)$

d) Usando o resultado c), mostre que:

$$E\left[\sum_{i=1}^n (x_i - \mu)\right]^2 = \sum_{i=1}^n E[(x_i - \mu)^2] = \sum_{i=1}^n \sigma^2 = n\sigma^2$$

e) Quadrando o resultado b) e usando c), mostre que:

$$E[(\hat{\mu} - \mu)^2] = \sigma^2/n \quad (\text{claramente lei do } 1/\sqrt{n})$$

f) Mostre que

$$\sum_{i=1}^n [(x_i - \mu) - (\hat{\mu} - \mu)]^2 = \sum_{i=1}^n (x_i - \mu)^2 - n(\hat{\mu} - \mu)^2$$

g) Usando f), demonstre que:

$$E\left[\sum_{i=1}^n (x_i - \hat{\mu})^2\right] = n\sigma^2 - n\frac{\sigma^2}{n} = (n-1)\sigma^2$$

h) Nesse caso,  $\sigma^2$  é um estimador não viésado de  $\sigma^2$

i)  $E[\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2]$  é estimador não viésado de  $\sigma^2$

j) Qual é valor da variança de  $\hat{\mu}$ ?

# ANS312 - Otimização em Sistemas Dinâmicos II

## Lista de Exercícios.

- 1) Dados um espaço amostral  $\Omega$  e uma  $\sigma$ -álgebra  $B$ , diz-se que  $P$  é uma função de probabilidade se:
- $P(\Omega) = 1$
  - $P(A) \geq 0$  para todo  $A \in B$
  - Se  $A_i \cap A_j = \emptyset$ , para todos  $A_i, A_j \in B$ ,  $i, j = 1, 2, \dots$ , então  $P(A_1 \cup A_2 \cup \dots) = \sum P(A_i)$

Com isso, demonstrar:

- $P(\emptyset) = 0$
- $P(A) = 1 - P(\bar{A}) \leq 1$ , onde  $\bar{A} \triangleq \Omega - A$
- $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$
- Se  $A_1 \subset A_2$ , então  $P(A_2) = P(A_1) + P(A_2 \cap \bar{A}_1)$

- 2) a) Determinar o valor  $b$  para que a função

$$f(x) = \begin{cases} bx(1-x) & 0 \leq x \leq 1 \\ 0 & \text{quero outro lugar} \end{cases}$$

seja uma função densidade de probabilidade

- Qual a probabilidade da v.z. estar entre  $2 \leq x \leq 5$ ?
- Qual a probabilidade da v.z. estar em  $-\infty < x < 0.6$ ?

- 3) A função densidade exponencial é definida por

$$f(x) = \begin{cases} qe^{-qx} & x > 0 \\ 0 & x \leq 0 \end{cases} \quad \text{onde } q \geq 0$$

- 3.a) Mostre que a função distribuição de probabilidade para este caso é dada por

$$F(x) = \begin{cases} 1 - e^{-qx} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

- 3.b) Usando integração por partes, mostre que a média de uma v.z.  $X$  com essa distribuição é  $E[X] = \frac{1}{q}$

- 3.c) Usando integração por partes, ache  $E[X^2]$

- 3.d) Com os resultados de 3.b) e 3.c) determine a variância de uma v.z. com essa distribuição.

4) a) determine o valor de  $b$  que torna

$$f(u,v) = \begin{cases} b e^{-2u-3v}, & u>0, v>0 \\ 0 & \text{toda outra parte} \end{cases}$$

uma função densidade conjunta de duas v.z.  $\underline{X}$  e  $\underline{Y}$

4b) calcule  $E[\underline{Z}]$  e  $E[\underline{Y}]$  usando integração por partes

5) Suponha que um vetor aleatório  $\underline{Z}$  esteja relacionado a outros dois vetores aleatórios  $\underline{X}$  e  $\underline{Y}$  através de:

$$\underline{Z} = A \underline{X} - \underline{Y}$$

$\underline{Z}, \underline{X}, \underline{Y}$  - vetores  $n \times 1$ ;  $A$  - matriz  $n \times n$  de constantes

Admita que sejam conhecidas, ainda, as matrizes

$$C_{\underline{X}} = \underline{\underline{E[X X^T]}}, C_{\underline{Y}} = \underline{\underline{E[Y Y^T]}}, C_{\underline{X}\underline{Y}} = \underline{\underline{E[X Y^T]}},$$

matrizes de  $2 \leq$  ordem.

5a) Mostre que  $C_{\underline{Z}\underline{X}} = A C_{\underline{X}} - C_{\underline{Y}\underline{X}}$

5b) Mostre que  $C_{\underline{Z}} = A C_{\underline{X}} A^T + C_{\underline{Y}} - A C_{\underline{X}\underline{Y}} - C_{\underline{X}\underline{Y}}^T A^T$

5c) Prove que  $C_{\underline{Y}\underline{Z}} = C_{\underline{X}\underline{Y}}^T A^T - C_{\underline{Y}}$

6) Sejam  $\underline{X}, \underline{Y}, \underline{Z}$  v.z. aleatórios conjuntamente distribuídos;  $c$  e  $d$  constantes;  $g(\cdot)$  função escalar fixa, não aleatória e admitimos que  $E[\underline{X}], E[\underline{Y}], E[\underline{Z}], E[(g(\underline{X})) \cdot \underline{X}]$  existem. Mostre que:

6a)  $E[\underline{X} | \underline{Y}] = E[\underline{X}]$  se  $\underline{X}$  e  $\underline{Y}$  são independentes

6b)  $E[\underline{X}] = E[E[\underline{X} | \underline{Y}]]$

6c)  $E[c | \underline{Y}] = c$

6d)  $E[g(\underline{Y}) | \underline{Y}] = g[\underline{Y}]$

6e)  $E[(c\underline{X} + d\underline{Z}) | \underline{Y}] = cE[\underline{X} | \underline{Y}] + dE[\underline{Z} | \underline{Y}]$

### 5-19 THE KALMAN FORM

The use of matrix identities will be used to transform our results into slightly different forms. These new formulas will be needed at a later time. Equation (5-71) multiplies out to obtain

$$\hat{x}^* = C_s^*(C_s^*)^{-1}\hat{x}^* + C_s^*B^{T*}(C_s^*)^{-1}\theta^*. \quad (5-73)$$

To save a little work, we now define  $K$  as follows:

$$K \equiv C_s^*B^{T*}(C_s^*)^{-1},$$

$$K = [(C_s^*)^{-1} + B^{T*}(C_s^*)^{-1}B^{T*}]^{-1}B^{T*}(C_s^*)^{-1}. \quad (5-74)$$

The identity (1-51) is used on the latter definition with

$$(C_s^*)^{-1} \rightarrow A^{-1}, \quad B^* \rightarrow B, \quad (C_s^*)^{-1} \rightarrow C^{-1},$$

to obtain

$$K = C_s^*B^{T*}(C_s^* + B^*C_s^*B^{T*})^{-1}. \quad (5-75)$$

The expression (5-72) for  $C_s^*$  will be transformed to a new form by means of the identity (1-45) with

$$\begin{aligned} (C_s^*)^{-1} &\rightarrow C^{-1}, \\ B^* &\rightarrow B, \\ (C_s^*)^{-1} &\rightarrow A^{-1}, \\ C_s^* - C_s^* = C_s^*B^{T*}(C_s^* + B^*C_s^*B^{T*})^{-1}B^*C_s^*. \end{aligned} \quad (5-76)$$

The expression (5-75) for  $K$  can be used to condense this last equation so that

$$\left\{ C_s^* - C_s^* = K B^* C_s^* \right\} \quad (5-77)$$

Finally, the definition for  $K$  in (5-74) and the value of  $C_s^*$  from (5-77) are placed into (5-73) to achieve the sequence:

$$\begin{aligned} \hat{x}^* &= (C_s^* - K B^* C_s^*)(C_s^*)^{-1}\hat{x}^* + K\theta^*, \\ \hat{x}^* - \hat{x}^* &= K B^* \hat{x}^* + K\theta^*. \end{aligned} \quad (5-78)$$

$$\left\{ \hat{x}^* = \hat{x}^* + K(\theta^* - B^*\hat{x}^*) \right\} \quad (5-79)$$

In this last equation, the matrix coefficient  $K$  may be viewed as a gain factor or constant to optimally weight together the new data  $\theta^*$  and the old estimate  $\hat{x}^*$  to obtain the new optimal estimate. Kalman [29] was the first to achieve this form by other arguments.

### 5-20 KALMAN VERSUS A PRIORI FORM

Let us collect our earlier formulas in order to provide a quick reference at a later time. Given an old estimate  $\hat{x}^*$ , old error matrix  $C_s^*$ , new data  $\theta^* = B^*\hat{x}^* + v^*$ ,

the hypotheses of  $H$ , and the assumption (5-67), the new estimate  $\hat{x}^n$  based on all the data is found by the sequence:

$$\begin{aligned} K &= C_e^* B^{T*} (C_e^* + B^* C_e^* B^{T*})^{-1}, \\ \hat{x}^n &= \hat{x}^e + K(\theta^* - B^* \hat{x}^e), \\ C_e^n &= C_e^* - K B^* C_e^*. \end{aligned} \quad (5-80)$$

The set of Eqs. (5-80) is, in fact, a generalization of the Kalman filter.

Let us also retrieve the original a priori results (5-71) and (5-72) for comparison to the Kalman form of (5-80):

$$\begin{aligned} \hat{x}^n &= C_e^n [(C_e^*)^{-1} \hat{x}^e + B^{T*} (C_e^*)^{-1} \theta^*], \\ C_e^n &= [(C_e^*)^{-1} + B^{T*} (C_e^*)^{-1} B^*]^{-1}. \end{aligned} \quad (5-81)$$

Note that the two sets of equations accomplish the same purpose—an old estimate and its error matrix are combined in an optimal way with a new set of data to arrive at a new optimal estimate and error matrix. The only advantage of one set over the other is numerical! If the dimension of the state vector  $x$  is greater than the number of data points in  $\theta^*$  then the Kalman form may be preferred, since the order of matrix inversion is less in the Kalman form than in the a priori form. If the dimension of the state vector  $x$  is less than the number of components of  $\theta^*$ , then the a priori form may be used to minimize the order of matrix inversion.

As a final comment, in any application of (5-80) or (5-81), the new data and the old estimate must be consistent! In other words, a new estimate and its variance based entirely upon the new set of data with no a priori information must lie within the range of uncertainty of the old estimate. If the two estimates are not consistent, then either the old estimate is incorrect or else the new set of data is erroneous, and in that event no combination of data should be attempted.

# DMC759 - Introdução à Teoria de Estimação

## 2ª Série de exercícios

1ª questão ]

Considere um rotor girando em torno de um eixo com velocidade constante, porém não conhecida. A posição angular é medida a cada T segundos de acordo com:

$$z_k = \theta_k + v_k \quad k = 1, 2, 3, \dots$$

onde  $\theta_k$  é a posição angular no instante  $t = kT$  e  $v_k$  é o erro de medida. As incertezas nas condições iniciais e nas medidas são descritas por:

$$E[\theta_0] = E[\dot{\theta}_0] = 0$$

$$E[\theta_0^2] = (20^\circ)^2, \quad E[\dot{\theta}_0^2] = (20 \text{ rad/s})^2; \quad E[\theta_0 \dot{\theta}_0] = 0$$

$$E[v_k^2] = (5^\circ)^2$$

Pede-se:

- 2) Escreva as equações de movimento do sistema e obtenha a matriz de transição de estados associada;
- 3) Monte as equações do filtro de Kalman para obter as estimativas ótimas do estado para as 2 primeiras observações e anexe os resultados.

2º questão]

Seja o sistema de primeira ordem escalar:

$$\dot{x} = w$$

onde  $w$  é um processo branco tal que:

$$E[w(t)] = 0$$

$$E[w(t)w(t+\tau)] = q \delta(t-\tau)$$

e o estado inicial  $x$  é gaussiano com:

$$E[x_0] = 0$$

$$E[x_0^2] = p_0$$

Admita um processo contínuo de medidas  $z(t)$

$$z(t) = x(t) + v(t)$$

onde  $v(t)$ , ruído de medida, é especificado por:

$$E[v(t)] = 0$$

$$E[v(t)v(t+\tau)] = r \delta(t-\tau)$$

Observe que  $q$ ,  $p_0$  e  $r$  são constantes escalares  
Mostre que o filtro ótimo para estimar  $x(t)$  é:

$$\hat{x}(t) = \left[ \frac{p(t)}{r} \right] (z(t) - \hat{x}(t))$$

onde  $p(t) = \sqrt{rq} \frac{1 + b e^{-2dt}}{1 - b e^{-2dt}}$  ;  $d = \sqrt{\frac{q}{r}}$  e

$$b = \frac{p_0 - \sqrt{rq}}{p_0 + \sqrt{rq}}$$

Sugestão: use a transformação  $p = r \frac{y}{y'}$  para integrar  
a Equação de Riccati

3ª questão]

Considere o sistema de primeiras ordens escalar:

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

onde  $u(t)$  é um controle e  $w(t)$  é um processo branco tal que:

$$E[w(t)] = 0; E[w(t)w(t+\delta)] = q \delta(t-\delta)$$

O estado inicial do sistema é admitido Gaussiano com estatística:

$$E[x(0)] = 0, E[x_0^2] = p_0$$

a) Deixe-se determinar a lei de controle ótimo  $u(t)$  que minimiza o funcional

$$J = E \left[ \int_0^{t_f} (x(t)^2 + p u(t)^2) dt + \frac{1}{2} x^2(t_f) \right]$$

onde  $t_f$  é conhecido e  $p$  é um parâmetro positivo.

Admita ~~que~~ que se tenha um processo contínuo de medidas  $z(t)$  dado por:

$$z(t) = x(t) + v(t)$$

onde  $v(t)$  é ruído de observações tal que:

$$E[v(t)] = 0; E[v(t)v(t+\delta)] = r \delta(t-\delta)$$

b) Considere que o instante final  $t_f$  se torne muito grande quando comparado às outras constantes de tempo do sistema, o que permite considerar a solução estacionária do problema. Admita, além disso, que você, como projetista, tenha que optar por fixar  $p$  em um das duas Níveis:

$$p_1 = 100 \text{ I/g} \quad \text{ou} \quad p_2 = 5 \text{ g}$$

Entre  $p_1$  e  $p_2$ , qual seria a sua escolha? Justifique sua escolha!

