Universidad de Santiago de Chile
Facultad de Ingeniería
Depto. de Ingeniería Informática

Análisis de Datos
Capítulo VIII
"Análisis de Series Temporales"

continuación

Profesor: Dr. Max Chacón.

8.4. Análisis de sistemas lineales

Modelo: abstracción de la realidad.
Funciones

Reducir la complejidad: permitié ver las características importantes de un proceso, eliminando detalles.

Predecir: Dado un conjunto de casos conocidos, determinar casos nuevos de los cuales, no se conoce su salida.

Tanto para el análisis y la predicción de un problema real se utiliza la esta abstracción de la realidad.

Modelos y Sistemas

Modelos no Paramétricos

Modelos que necesitan una gran cantidad de parámetros (*puede llegar a ser infinita*) para representar el fenómeno. (Ej: función de transferencia, respuesta al impulso).

Modelos Paramétricos

Modelos matemático que están completamente caracterizado por un conjunto finito de parámetros. (descripciones matemáticas de las relaciónes entre las variables del sistema (Ej: FIR, ARX, ARMAX, OE).

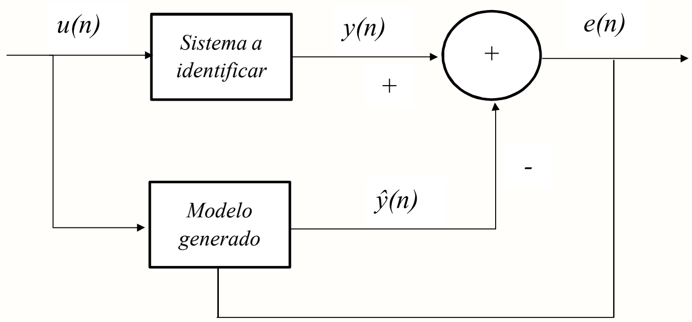
Identificación de Sistemas (IS)

Metodología que consiste en construir modelos matemáticos del sistema estudiado.

Reglas:

- Obtención de datos de entrada y salida del sistema que se desea identificar.
- Pre-procesamiento de los datos de entrada y salida. (Examinar, limpiar y filtrar los datos).
- Elección del conjunto de modelos.
- Obtención de los parámetros del modelo y validación.
-Si modelo tiene buenos resultados (alcanzo un error aceptable), terminar sino realizar cambios en punto anterior e iterar.

Estructura general de la Identificación de sistemas



$e(n) = y(n) - \hat{y}(n)$

- ***ŷ(n)*** la salida del modelo generado y ***e*** el error. La identificación de los parámetros de los modelos para Sistemas Lineales se basa en el error de predicción ***e(n)= y(n)- ŷ(n)***

8.5. Modelos no paramétrico (Función de Transferencia, TF)

La base de la construcción de estos modelos esta en la transformada de Fourier.

$$X(f) = \mathscr{F}[x(t)] = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt$$

En su forma discreta:

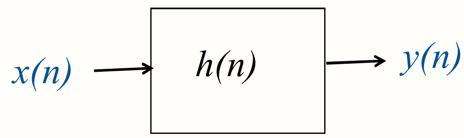
$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} \qquad k = 0,1,2,\ldots,N-1$$

k: frecuencia discreta

Transformada Inversa

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X(k)e^{j2\pi kn/N} \qquad n = 0,1,2\ldots,N-1$$

Para Sistema Lineal Invariante en el Tiempo (SLI)



En un SDLI, la salida es la convolución de la entrada con la función de transferencia del sistema

$$y(n) = \sum_{k=-\infty}^{\infty} h(n)x(n-k) = \sum_{k=-\infty}^{\infty} x(n)h(n-k)$$

En la frecuencia es la multiplicación

$$Y(f) = X(f)H(f)$$

Multiplicando por la entrada

$$Y(f)X(f) = X(f)X(f)H(f) = X^2(f)H(f)$$

$$H(f) = X(f)Y(f)/X^2(f)$$

Donde:

Espectro cruzado entrada salida: $G_{xy}(f) = X(f)Y(f)$

Auto-espectro de la entrada: $G_{xx}(f) = X^2(f)$

$$H(f) = \frac{G_{xy}(f)}{G_{xx}(f)}$$

En forma practica los espectros cruzados y la auto-correlación se calcula de las estimaciones.

Los espectros son funciones de la frecuencia complejas. Se requiere mostrar el espectro en modulo y fase.

$$|H(f)| = \sqrt{|H_R(f)|^2 + |H_I(f)|^2}$$

$$\phi(H(f)) = \tan^{-1}\frac{|H_I(f)|}{|H_R(f)|}$$

La estimación de la función de transferencia tiene dos maneras para ser implementada:

- Para evitar los errores aleatorios en el calculo de la estimación de la FT, es necesario cortar la señal en ventas y promediarlas en forma traslapada.
- Esto también genera distorsiones (fuga espectral), para evitar estas distorsiones las señales se multiplican por ventanas, que suavizan este espectro como: triangulares, coseno o Hanning.
- Los errores se reducen al promediar los espectros estimados sobre múltiples segmentos de datos (método de Welch)

Otra forma no paramétrica de modelar es obtener la respuesta al impulso del sistema

Sabemos que en un SDLI

$$y(n) = \sum_{k=0}^{\infty} h(n)x(n-k) = \sum_{k=0}^{n} x(n)h(n-k)$$

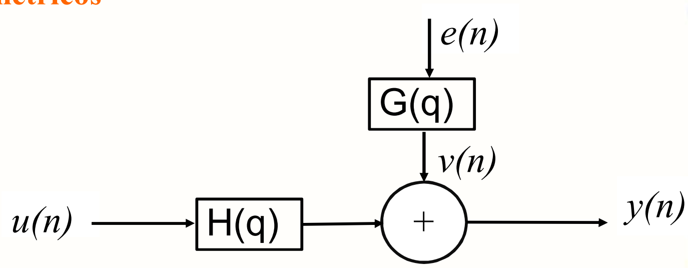
Entonces si se aplica una función impulso a la entrada del sistema, se tendrá:

$$y(n) = \sum_{k=0}^{n} \delta(n)h(n-k) = h(n)$$

Por lo tanto al ingresar una función impulso unitario en cero se obtendrá la función de transferencia del sistema

8.6 Modelos Paramétricos

8.6.1. Estructura general de los modelos lineales paramétricos



Donde:

- ***u(t):*** *Entrada del sistema*
- ***e(t):*** *Entrada de ruido blanco (señal aleatoria no correlacionada a lo largo del tiempo)*
- ***v(t):*** *salida de ruido*
- ***y(t):*** *salida debida a u(t) y v(t)*
- ***H(t),G(t):*** *funciones de transferencia del sistema y ruido.*

La salida del sistema dependerá de la entrada al sistema y del ruido

$$y(n)=\sum_{k=1}^{\infty}h(k)u(n-k)+v(n)$$

$$v(n)=\sum_{k=0}^{\infty}g(k)e(n-k)$$

Operador de retraso y adelanto en el tiempo ***q***

Definición: operador adelanto ***q*** en el tiempo

q u(t)=u(n+1)

Definición: operador retardo q^{-1} en el tiempo

$q^{-1}u(n)=u(n-1)$

Este operador permite representar algebraicamente las operaciones de las funciones de transferencia

$$H(q)=\sum_{k=1}^{\infty}h(k)q^{-k}$$

$$H(q)u(n)=\sum_{k=1}^{\infty}h(k)q^{-k}u(n)$$

$$y(n)=H(q)u(n)=\sum_{k=1}^{\infty}h(k)u(n-k)$$

$$v(n)=G(q)e(n)=\sum_{k=0}^{\infty}g(k)e(n-k)$$

$$con:G(q)=\sum_{k=0}^{\infty}g(k)q^{-k}$$

Entonces un SDLI con ruido será representado por:

$$y(n) = H(q)u(n) + G(q)e(n)$$

Donde las funciones *H(q)* y *G(q)* son funciones racionales, de un numero de parámetros a determinar.

$$f(q) = \frac{P(q)}{Q(q)}$$

Donde *P(q)* y *Q(q)* son polinomios cuyos grados determinaran los parámetros que estiman los modelos lineales.

$$y(n) = H(q,\theta)u(n) + G(q,\theta)e(n)$$

Donde:

$$H(q,\theta) = \frac{B(q)}{F(q)} = \frac{b_1 q^{n_k} + b_2 q^{n_k - 1} + \ldots + b_{n_b} q^{n_b}}{1 + f_1 q + \ldots + f_{n_f} q^{n_f}}$$

La función del ruido será:

$$G(q,\theta) = \frac{C(q)}{D(q)} = \frac{1 + c_1 q + \ldots + c_{n_c} q^{n_c}}{1 + d_1 q + \ldots + d_{n_d} q^{n_d}}$$

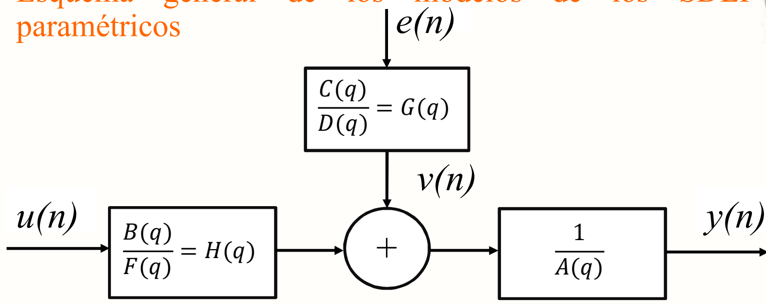
Donde:

- n_k: numero de retardos de la entrada del modelo
- n_b, n_f, n_c, n_d: orden de los polinomios
- b_n, c_n: numero de coeficientes de los numeradores
- c_n, d_n: numero de coeficientes de los denominadores

También es posible ingresar retardos en la salida del sistema, para eso se generará el polinomio ***A(q)***

$$A(q)=1+a_1 q+a_2 q^2+\ldots+a_{na} q^{na}$$

Esquema general de los modelos de los SDLI paramétricos



$$A(q)y(n) = \frac{B(q)}{F(q)}u(n) + \frac{C(q)}{D(q)}e(n)$$

Entonces el problema de identificación de sistemas consistirá: en encontrar un modelo, que se comporte como un sistema lineal, dependiendo de la forma que adopten ***H(q)*** y ***G(q)***.

Según la forma general de estas fracciones racionales, se pueden definir diferentes tipos de sistemas.

Según las formas de las fracciones: ***H(q)*** y ***G(q)*** se pueden definir algunos de los modelos que se describen enseguida:

FIR

Finite Impulse Response

ARX

Auto-regressive with exogenous input model

ARMAX

Auto-regressive Moving Average with eXogenous inputs

OE

Output Error

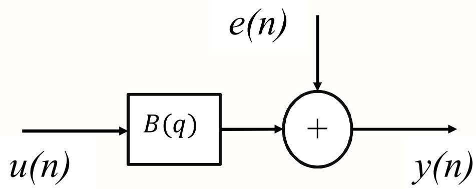
➢ **FIR:** Finite Impulse Response

Es un sistema cuya salida solo depende de las entradas y los tiempos anteriores de ésta.

Aquí: *A(q)=F(q)=C(q)=D(q)=1*

$$1y(n)=\frac{B(q)}{1}u(n)+\frac{1}{1}e(n)$$

$$y(n)=B(q)u(n)+e(n)$$



Donde:

- ***y(t):*** la salida estimada
- ***u(t):*** la entrada
- ***e(t):*** el ruido blanco o error

y con: *A(q)=F(q)=C(q)=D(q)=1*

El problema aquí será determinar, cual es el grado del polinomio (n_b)

$$B(q)=b_0+b_1q^{-1}+b_2q^{-2}+\ldots+b_{n_b}q^{-n_b}$$

y los valores de n_b +1 y los valores de los parámetros b_i

$$y(n)=b_0u(n)+b_1u(n-1)+\ldots+b_{n_b}u(n-n_b)+e(n)$$

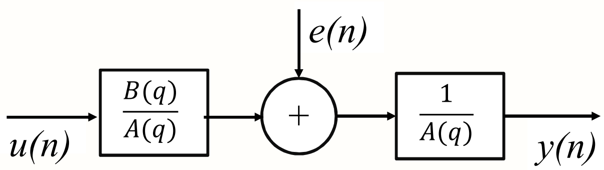
➢ **ARX:** modelo auto regresivo con entrada exógenas.

La salida depende de las entradas (actuales y anteriores) pero también de la salidas anteriores.

Aquí: ***C(q)= D(q)= F(q)= 1***

$$A(q)y(n) = \frac{B(q)}{1}u(n) + \frac{1}{1}e(n)$$

$$A(q)y(n) = B(q)u(n) + e(n)$$



Ahora se tienen dos polinomios a determinar:

$$A(q) = 1 + a_1q^{-1} + b_2q^{-2} + \ldots + b_{n_a}q^{-n_a}$$

$$B(q) = b_0 + b_1q^{n-1} + b_2q^{-2} + \ldots + b_{n_b}q^{-n_b}$$

Con $\boldsymbol{n_a}$ y $\boldsymbol{n_b}$ número de retardos en entrada y salida respectivamente.

$$y(n) = \frac{B(q)}{A(q)}u(n) + \frac{1}{A(q)}e(n)$$

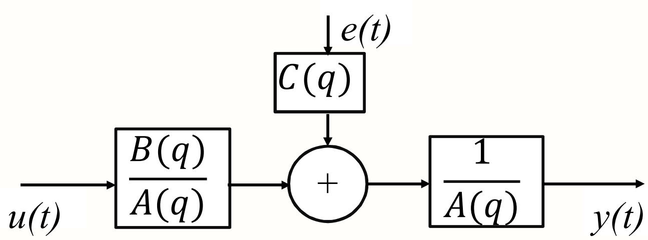
$$y(n) = -a_1y(n-1) - a_2y(n-2) - \ldots - a_{n_a}y(n-n_a)$$
$$+ b_0u(n) + b_1u(n-1) + \ldots + b_{n_b}u(n-n_b) + e(n)$$

➢ **ARMAX:** modelo auto regresivo con entrada exógenas y media móvil. La salida depende de las entradas (actuales y anteriores), de las salidas anteriores, en conjunto con un promedio el ruido.

Aquí: ***F(q)=D(q)= 1***

$$A(q)y(n) = \frac{B(q)}{1}u(n) + \frac{C(q)}{1}e(n)$$

$$A(q)y(n) = B(q)u(n) + C(q)e(n)$$



Se tienen que determinar tres polinomios:

$$A(q) = 1 + a_1 q^{-1} + b_2 q^{-2} + \ldots + b_{n_a} q^{-n_a}$$

$$B(q) = b_0 + b_1 q^{n-1} + b_2 q^{-2} + \ldots + b_{n_b} q^{-n_b}$$

$$C(q) = 1 + c_1 q^{-1} + c_2 q^{-2} + \ldots + c_{n_c} q^{-n_c}$$

Con n_a, n_b y n_c número de retardos entrada, salida y ruido.

$$\begin{aligned} y(n) = &-a_1 y(n-1) - a_2 y(n-2) - \ldots - a_{n_a} y(n-n_a) \\ &+ b_0 u(n) + b_1 u(n-1) + \ldots + b_{n_b} u(n-n_b) + e(n) \\ &+ c_1 e(n-1) + c_2 e(n-2) + \ldots + c_{n_c} e(n-n_c) \end{aligned}$$

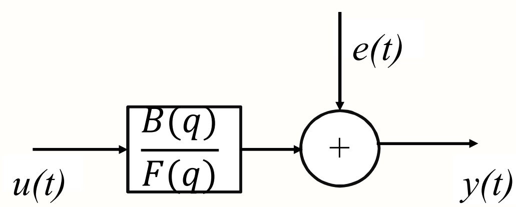
➢**OE: (Output Error)** modelo de error de salida

La salida en un momento determinado de tiempo, depende de las entradas y las salidas de los tiempos anteriores y de una entrada de ruido.

Aquí: ***A(q)= C(q)= D(q)= 1***

$$y(n) = \frac{B(q)}{F(q)}u(n) + \frac{1}{1}e(n)$$

$$F(q)w(n) = B(q)u(n) \quad y \quad y(n) = w(n) + e(n)$$



Se requiere determinar dos polinomios:

$$B(q) = b_0 + b_1 q^{n_{-1}} + b_2 q^{-2} + \ldots + b_{n_b} q^{-n_b}$$
$$F(q) = 1 + f_1 q^{-1} + f_2 q^{-2} + \ldots + f_{n_f} q^{-n_f}$$

Con $\boldsymbol{n_b}$ ***y*** $\boldsymbol{n_c}$ número de retardos en el denominador y numerador de la función de transferencia.

$$y(n) = -f_1 w(n-1) - f_2 w(n-2) - \ldots - f_{n_f} w(n-n_f)$$
$$+ b_0 u(n) + b_1 u(n-1) + \ldots + b_{n_b} u(n-n_b) + e(n)$$

$$w(n) = -f_1 w(n-1) - f_2 w(n-2) - \ldots - f_{n_f} w(n-n_f)$$
$$+ b_0 u(n) + b_1 u(n-1) + \ldots + b_{n_b} u(n-n_b)$$

$$y(n) = w(n) + e(n)$$

- **Box and Jenkins:** La salida depende de entradas y ruido, en tiempos actuales y anteriores.

Aquí: ***A(q)= 1***

$$y(n) = \frac{B(q)}{F(q)} u(n) + \frac{C(q)}{D(q)} e(n)$$

$$w(n) = -f_1 w(n-1) - f_2 w(n-2) - \ldots - f_{n_f} w(n-n_f)$$
$$+ b_0 u(n) + b_1 u(n-1) + \ldots + b_{n_b} u(n-n_b)$$

$$v(n) = -d_1 v(n-1) - d_2 v(n-2) - \ldots - d_{n_d} v(n-n_d)$$
$$+ e(n) + c_1 e(n-1) + \ldots + c_{n_c} e(n-n_c)$$

$$y(n) = w(n) + v(n)$$

Resumen

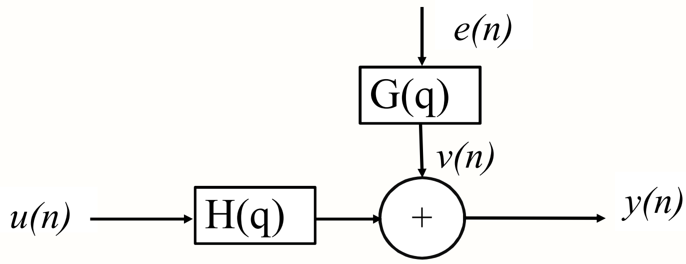
Modelo	A	B	C	D	F
FIR	1	Si	1	1	1
ARX	Si	Si	1	1	1
ARMAX	Si	Si	Si	1	1
OE	1	Si	1	1	Si
BJ	1	Si	Si	Si	Si

Las principales diferencias están en las estructuras

- BJ y ARMAX son estructuras de promedio de error
- FIR, ARX y OE son estructuras del tipo de error de salida.

8.6.2. Predicción de la estructura general de los modelos lineales paramétricos

El modelo de predicción estándar será:



$$y(n) = H(q)u(n) + G(q)e(n) = H(q)u(n) + v(n)$$

La estimación de los parámetros se basa en la estimación de un solo paso adelante One-Step-a-Head

En la ecuación del sistema anterior se supone que se conocen todos los valores *y(n)* e *u(n)* para *n=k-1,k-2,…* de los valores anteriores a partir de un tiempo n_0. Suponiendo que por el momento conocemos *H(q)* y *G(q)* entonces también se conoce *v(n)=y(n)-H(q)u(n)* *(8.1).*

La predicción de un paso adelante será.

$$\hat{y}(n|n-1) = H(q)u(n) + \hat{v}(n|n-1) \qquad 8.2$$

Para resolver esta predicción primero se tiene que resolver la de *v(n),* para eso la función de ruido *G(q)* debe ser: Estable, invertible y ser representada como respuesta a función impulso.

Representándola como respuesta a función impulso.

$$\hat{v}(n|n-1) = \sum_{k=0}^{\infty} g(k)e(n-k) = [G(q)-1]e(n)$$

Como se supone que el ruido es invertible y estable.

$$e(n) = \frac{1}{G(q)} v(n)$$

Entonces $\hat{v}(n|n-1) = \frac{[G(q)-1]}{G(q)} e(n) = [1-G^{-1}(q)]e(n)$

Reemplazando en (8.2), y luego *v(n)* de (8.1)

$$\hat{y}(n|n-1) = H(q)u(n) + [1-G^{-1}(q)]v(n) \qquad 8.3$$

$$\hat{y}(n|n-1) = H(q)u(n) + [1-G^{-1}(q)](y(n) - H(q)u(n))$$

$$\hat{y}(n|n-1) = G^{-1}(q)H(q)u(n) + [1-G^{-1}(q)]y(n)$$

$$G(q)\hat{y}(n|n-1) = H(q)u(n) + [G(q)-1]y(n) \qquad (8.3)$$

De la expresión anterior se puede obtener el error

$$\varepsilon(n|n-1) = \hat{y}(n|n-1) - y(n)$$

$$\varepsilon(n|n-1) = G^{-1}(q)H(q)u(n) - G^{-1}(q)y(n)$$

La expresión anterior, representa la función de error para cualquier estructura general de los modelos que se quieran determinar los valores en el tiempo ***n*** con los datos del tiempo ***n-1***.

Esta formula fue obtenida suponiendo que ***H(q)*** y ***G(q)*** son funciones conocidas, pero no los son, pues son polinomios o fracciones de polinomios, de los cuales no se conocen sus términos constantes (los parámetros, los llamaremos ***θ***).

Entonces en función de los parámetros este error será:

$$\varepsilon(n|n-1,\theta) = G^{-1}(q,\theta)\left[H(q,\theta)u(n) - y(n)\right]$$

La solución entonces será obtener estos parámetros para que el error sea mínimo, en cada modelo

➢Predicción del modelo para la estructura FIR

Se requieren los parámetros b_0, b_1, b_2, …,b_{nb}

$$\vec{\theta} = \left[b_0\, b_1\, b_2\, \dots\, b_{n_b}\right]^T$$

Al comparar la estructura general con la particular del FIR se puede ser: ***G(q,θ)=B(q) y H(q,θ)=1***.

La función de transferencia del sistema depende directamente de los parámetros.

El modelo predictor FIR será:

$$\hat{y}(n|\theta) = B(q)u(n) \qquad 8.4$$

Las mediciones de las entradas para *n-1*, son conocidas y llamadas el ***vector de regresión ϕ***.

$$\phi(n) = \left(u(n-1)u(n-2)\ \dots\ u(n-n_b)\right)^T$$

El predictor FIR se puede escribir en la ***forma de regresión***

$$\hat{y}(n|\theta) = \theta^T \phi(n) = \phi^T(n)\theta$$

Como la salida del modelo ***y(n)***, depende linealmente (es una combinación lineal) de parámetro $\boldsymbol{\theta}$, a esta forma se le llama un ***modelo de regresión lineal.***

Nota: Se debe considerara que la estructura de un modelo lineal, no implica que la dinámica (relación entrada salida) sea lineal

➢**Predicción del modelo para la estructura ARX**

Se requieren los parámetros corresponde a los polinomios *A(q)* y *B(q)*.

$$\vec{\theta} = \left[a_1\, a_2\, \dots\, a_{n_a}\, b_0\, b_1\, \dots\, b_{n_b}\right]^T$$

$$H(q|\theta) = \frac{B(q)}{A(q)} \qquad y \qquad G(q|\theta) = \frac{1}{A(q)}$$

Usando el predictor general de un paso adelante (8.3), se tiene:

$$\hat{y}(n|\theta) = B(q)u(n) + \left[1 - A(q)\right]y(n)$$

El vector de regresión será:

$$\phi(n) = \left(-y(n-1) \;\ldots\; -y(n-n_a)\; u(n-1) \;\ldots\; u(n-n_b)\right)^T$$

El modelo predictor, escrito como un modelo de regresión lineal será:

$$\hat{y}(n|\theta) = \theta^T \phi(n) = \phi^T(n)\theta$$

- **Predicción del modelo para la estructura ARMAX**

Se requieren los parámetros corresponde a los polinomios *A(q), B(q) y C(q)*.

$$\vec{\theta} = \left[a_1\, a_2\, \ldots\, a_{n_a}\, b_0\, b_1\, \ldots\, b_{n_b}\, c_1\, c_2\, \ldots\, c_{n_c}\right]^T$$

$$H(q|\theta) = \frac{B(q)}{A(q)} \qquad y \qquad G(q|\theta) = \frac{C(q)}{A(q)}$$

$$\hat{y}(n|\theta) = \frac{B(q)}{C(q)}u(n) + \left[1 - \frac{A(q)}{C(q)}\right]y(n)$$

Multiplicando esta expresión por *C(q)* y sumando a ambos lados la expresión $[1 - C(q)]\hat{y}(n|\theta)$

$$\hat{y}(n|\theta) = B(q)u(n) + [C(q) - A(q)]y(n) + [1 - C(q)]\hat{y}(n|\theta)$$

Usando: $\varepsilon(n|\theta) = y(n) - \hat{y}(n|\theta)$

$$\hat{y}(n|\theta) = B(q)u(n) + [1 - A(q)]y(n) + [C(q) - 1]\varepsilon(n|\theta)$$

Con este modelo podríamos obtener el *vector de regresión* para obtener el modelo predictor, pero este vector dependerá de del parámetro θ.

$$\phi(n|\theta) = (-y(n-1) \; \dots \; -y(n-n_a) \; u(n-1) \; \dots \; u(n-n_b)$$
$$\varepsilon(n-1|\theta) \; \varepsilon(n-2|\theta) \; \dots \; \varepsilon(n-n_c|\theta))^T$$

El modelo de regresión no es lineal, se denomina *modelo de regresión seudo-lineal*:

$$\hat{y}(n|\theta) = \theta^T \phi(n|\theta) = \phi^T(n|\theta)\theta$$

➢Predicción del modelo para la estructura OE

Se requieren los parámetros corresponde a los polinomios *B(q)* y *F(q)*.

$$\vec{\theta} = \left[b_0 \, b_1 \dots b_{n_b} \; f_1 \, f_2 \dots f_{n_c}\right]^T$$

$$H(q|\theta) = \frac{B(q)}{f(q)} \qquad y \qquad G(q|\theta) = 1$$

$$\phi(n|\theta) = \left(u(n-1) \; \dots \; +u(n-n_b) - w(n-1|\theta) \; \dots \; -w(n-n_f|\theta)\right)^T$$

$$\hat{y}(n|\theta) = \theta^T \phi(n|\theta) = \phi^T(n|\theta)\theta$$

8.6.3 . Método de estimación de parámetros

Se usa la minimización de la función de costo, basada en la suma de cuadrados de la predicción de los errores.

La técnica matemática usado para implementar esta minimización, es el denominado ***método de los mínimos cuadrados****.*

El problema a resolver es:

Dada las mediciones de entrada y salida de N muestras: ***{u(n),y(n)}$_N$***. Determine los parametros del modelo con mínimo error.

$$J_N(\theta) = \frac{1}{N}\sum_{n=1}^{N} \varepsilon^2(n|\theta)$$

La estimación del parámetro θ para el conjunto de N muestras será:

$$\hat{\theta}_N = \arg\min_{\hat{\theta}\in D_M} J_M(\theta)$$

El caso de los modelos FIR y ARX el modelo de regresión es lineal, entonces el error no depende de θ

$$J_N(\theta) = \frac{1}{N}\sum_{n=1}^{N} \left[y(n) - \phi^T(n)\theta\right]^2$$

La solución se obtiene al igualar el gradiente a cero

$$\frac{\partial}{\partial\theta}J_N(\hat{\theta}_N)=0$$

Derivando cada cuadrado de la sumatoria:

$$\frac{\partial}{\partial\theta}\left[y(n)-\phi^T(n)\theta\right]^2=\frac{\partial}{\partial\theta}\left[\phi^T(n)\theta\right]\left[y(n)-\phi^T(n)\theta\right]$$

$$\frac{\partial}{\partial\theta}\left[y(n)-\phi^T(n)\theta\right]^2=-2\phi(n)\left[y(n)-\phi^T(n)\theta\right]=0$$

$$\frac{\partial}{\partial\theta}J_N(\hat{\theta}_N)=-\frac{2}{N}\sum_{n=1}^{N}\phi(n)\left[y(n)-\phi^T(n)\theta\right]=0$$

La solución esta dada por la solución de este sistema de ecuaciones: ***ecuaciones normales***

$$\left[\frac{1}{N}\sum_{n=1}^{N}\phi(n)\phi^T(n)\right]\hat{\theta}_N=\frac{1}{N}\sum_{n=1}^{N}\phi(n)y(n)$$

Usando representación vectorial

$$\vec{\Phi}=\begin{bmatrix}\Phi^T(1)\\ \Phi^T(2)\\ \vdots\\ \Phi^T(N)\end{bmatrix}\quad y\quad \vec{Y}=\begin{bmatrix}y(1)\\ y(2)\\ \vdots\\ y(N)\end{bmatrix}$$

$$J_N(\hat{\theta}_N)=\left\|\vec{Y}-\vec{\Phi}\vec{\theta}\right\|^2$$

$$\left[\vec{\Phi}^T\vec{\Phi}\right]\hat{\vec{\theta}}=\vec{\Phi}^T\vec{Y}$$

La solución será: $\hat{\vec{\theta}}=\left[\vec{\Phi}^T\vec{\Phi}\right]^{-1}\vec{\Phi}^T\vec{Y}$

Este método obtiene los parámetros de los modelos FIR y ARX.

Método de los mínimos cuadrados seudo-lineales

El problema es este caso es que el vector de datos $\phi(n,\theta)$, también depende de los parámetros θ que se requieren estimar. Esto es requerido para los modelos ARMAX y OE.

Para resolver este problema se requiere un método de optimización no lineal, que se resuelve el problema en forma iterativa.