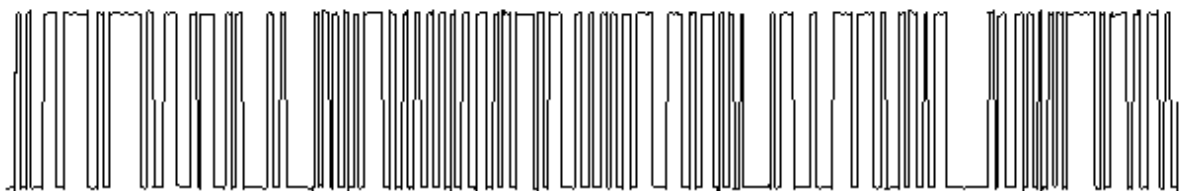
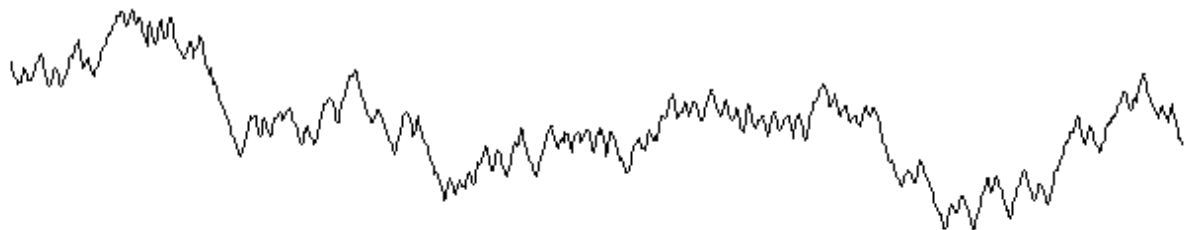




Universidad Nacional de la Plata
Facultad de Ingeniería
Dpto. de Electrotecnia

Cátedra de Control y Servomecanismos

IDENTIFICACIÓN DE SISTEMAS DINÁMICOS



Cristian Kunusch
Año 2003

INDICE

Introducción	2
1. Sistemas Lineales Invariantes en el Tiempo	3
1.1 Respuesta Impulsiva, Funciones de Transferencia y Densidad Espectral de Potencia:.....	3
Respuesta Impulsiva.....	3
Muestreo.....	3
Perturbaciones.....	4
Caracterización de Perturbaciones.....	4
Función Covarianza.....	5
Función de Autocorrelación.....	5
Funciones de Transferencia.....	5
1.2 Espectro de señal.....	6
Definición de Densidad Espectral de Potencia.....	7
Espectro de un Proceso Estocástico Estacionario.....	7
Transformación del espectro introducida por un Sistema Lineal.....	8
2. Métodos de Identificación No Paramétrica	9
2.1 Análisis espectral.....	9
<i>Aplicación en Matlab</i>	10
2.3 Análisis por Correlación (estimación de la respuesta impulsiva).....	10
<i>Aplicación en Matlab</i>	12
3. Predicción	13
Invertibilidad del Modelo de Ruido.....	13
Predicción de v	13
Predicción de y	14
4. Modelos Paramétricos de Sistemas Lineales Invariantes en el Tiempo	15
4.1 Modelos lineales y sets de modelos lineales.....	15
4.2 Familias de modelos de funciones de transferencia.....	16
Estructura ARX.....	16
Estructura ARMAX.....	17
Estructura ARARX.....	18
Estructura ARARMAX.....	18
Estructura de Error de Salida.....	18
Estructura Box-Jenkins.....	19
Familia General de Estructuras.....	20
<i>Aplicación en Matlab</i>	20
5. Selección de Estructura y Validación del Modelo	21
5.1 Aspectos generales en la elección de una estructura modelo.....	21
Calidad del Modelo.....	21
Precio del Modelo.....	22
Consideraciones generales.....	22
5.2 Consideraciones a priori.....	22
<i>Tipo del Modelo</i>	22
<i>Orden del Modelo</i>	23
5.3 Selección de Estructura basada en análisis preliminares de los datos.....	23
Estimación del Tipo del modelo.....	23
Estimación del Orden del modelo.....	23
5.4 Comparación de estructuras modelo.....	25
Comparación de modelos utilizando sets de datos frescos (Validación Cruzada).....	25
Comparación de modelos con sets de datos de "Segunda mano".....	25
5.5 Validación del modelo.....	26
Validación del modelo con respecto al propósito del modelado.....	26

Viabilidad de los Parámetros Físicos	26
Consistencia del Modelo Entrada-Salida	26
Reducción del Modelo	27
Simulación	27
5.6 Conclusiones finales	27
5.7 Aplicación en Matlab	28
Comparación de Diferentes Estructuras	28
Chequeo de cancelaciones Polo-Cero.....	29
Análisis Residual	29
Simulaciones libres de ruido	30
Evaluación de la incertidumbre del modelo	30
Comparación de Diferentes Modelos	31
6. Acondicionamiento de los Datos.....	32
6.1 Perturbaciones de alta frecuencia.....	32
6.2 Bursts y Outliers	32
6.3 Perturbaciones Lentas: Comentarios Generales	32
6.4 Evaluación de las Aproximaciones.....	33
6.5 Drift, Trends y Variaciones de Baja Frecuencia.....	34
6.6 Aplicación en Matlab	34
Niveles de Offset	34
Outliers	34
Filtrado de los datos	34
Realimentación de los Datos.....	35
Retardos	35
7. Nociones Básicas para el Diseño de la Señal de Excitación.....	36
7.1 Secuencia Binaria de Alternancia Psudoaleatoria	36
8. Conclusiones Generales	38
Que nos puede ofrecer la identificación de sistemas ?.....	38
Posibilidades	38
Limitaciones: Calidad de los Datos	38
Limitaciones: Estructuras Modelo	38
Bibliografía.....	39

INTRODUCCIÓN

En este apunte se presentan los conceptos de la teoría de Identificación de Sistemas Dinámicos, siendo esta la disciplina encargada de la realización de modelos matemáticos de sistemas, partiendo del análisis de datos experimentales, mediciones y observaciones del sistema. Su base teórica se sustenta en parte por la teoría de sistemas dinámicos y señales de naturaleza estocástica, como así también por métodos y algoritmos matemáticos para la correcta estimación de los parámetros involucrados. Cada rutina de identificación presentada se encuentra complementada con su correspondiente aplicación en Matlab ©.

Debido a que los sistemas dinámicos abundan en nuestro medio ambiente, las técnicas de identificación de sistemas han cobrado gran relevancia en diversas áreas del conocimiento (ingenierías, economía, biotecnología, etc.), donde se requiere de un modelo preciso para fines de análisis, predicción, simulación, diseño y control.

En particular, las técnicas de control actuales requieren de modelos matemáticos cada vez más exactos para el análisis y el diseño. Siendo que en muchos casos, tales modelos no pueden ser obtenidos en forma sencilla y económica a partir de las leyes que rigen cada proceso. Aquí es entonces donde juega un papel decisivo la Identificación de Sistemas Dinámicos, que es una herramienta capaz de proporcionar los métodos necesarios para obtener de manera relativamente sencilla los modelos matemáticos buscados con un alto grado de exactitud.

Objetivo general

Obtener una visión general de los métodos de identificación paramétrica y no paramétrica, dar un fundamento matemático de los mismos e introducir los conceptos esenciales para el diseño de un ensayo de identificación típico con la ayuda del software adecuado, de manera que se puedan comparar las ventajas y desventajas de cada uno de ellos.

1. SISTEMAS LINEALES INVARIANTES EN EL TIEMPO

1.1 Respuesta Impulsiva, Funciones de Transferencia y Densidad Espectral de Potencia:

Respuesta Impulsiva

Consideremos un sistema con una única entrada $u(t)$ y una única salida $y(t)$ (SISO). El sistema será *invariante en el tiempo* si su respuesta a una cierta señal de entrada no depende en absoluto del tiempo, y será *lineal* si su respuesta a una combinación lineal de entradas, es la misma combinación lineal de las respuestas a las entradas individuales. De la misma forma podemos decir que un sistema es *causal* si la salida en un tiempo determinado solo depende de la entrada actual y de las entradas anteriores a dicho instante de tiempo.

Entonces un sistema lineal, invariante en el tiempo y causal puede ser perfectamente descrito por su respuesta al impulso $g(\tau)$ de la siguiente manera:

$$y(t) = \int_{\tau=0}^{\infty} g(\tau) \cdot u(t - \tau) d\tau \quad (1.1)$$

Muestreo

Dado que de ahora en más vamos a trabajar casi exclusivamente con señales de tiempo discreto, entonces asumiremos que $y(t)$ será observada solo en los instantes de muestreo $t_k = kT$ con $k=1,2,3,4,\dots$:

$$y(kT) = \int_{\tau=0}^{\infty} g(\tau) \cdot u(kT - \tau) d\tau \quad (1.2)$$

siendo T el periodo de muestreo.

Llegado este punto es importante destacar que en aplicaciones de identificación la entrada $u(t)$ es comúnmente mantenida constante entre los intervalos de muestreo, de modo que:

$$u(t) = u_k \quad \text{cuando} \quad kT \leq t < (k+1)T \quad (1.3)$$

Esta situación es por lo general implementada por razones de practicidad en la implementación, y también involucra una gran simplicidad en el análisis del sistema. Reemplazando (1.3) en (1.2):

$$\begin{aligned} y(kT) &= \int_{\tau=0}^{\infty} g(\tau) \cdot u(kT - \tau) d\tau = \sum_{n=1}^{\infty} \int_{\tau=(n-1)T}^{nT} g(\tau) \cdot u(kT - \tau) d\tau = \\ &= \sum_{n=1}^{\infty} \left[\int_{\tau=(n-1)T}^{nT} g(\tau) \cdot d\tau \right] \cdot u_{k-n} = \sum_{n=1}^{\infty} g_T(n) \cdot u_{k-n} \end{aligned} \quad (1.4)$$

donde definimos:

$$g_T(n) = \int_{\tau=(n-1)T}^{nT} g(\tau) \cdot d\tau$$

Pero la mayor parte del tiempo asumiremos un intervalo de muestreo unitario y usaremos la variable t para enumerar los instantes de muestreo. De esta manera podemos escribir la ecuación (1.4) de la siguiente forma:

$$y(t) = \sum_{k=1}^{\infty} g_{T=1}(k) \cdot u(t-k) = \sum_{k=1}^{\infty} g(k) \cdot u(t-k) \quad \text{con} \quad t = 0,1,2,3,\dots \quad (\text{convolución en tiempo discreto}) \quad (1.5)$$

Nota: esta notación será la que utilizaremos de aquí en adelante, sin perder de vista que es ligeramente distinta a la que comúnmente se utiliza para enumerar las muestras de una señal de tiempo discreto. De hecho, dicha modificación surge como intención ser consistentes con la mayor parte de la bibliografía existente.

Perturbaciones

De acuerdo con la ecuación anterior la salida puede ser calculada en forma exacta una vez conocida la entrada al sistema. Pero en la mayoría de los casos esto es imposible debido a que siempre existen señales espurias que afectan al sistema y se escapan de nuestro control. Ahora, bajo un marco teórico lineal, podemos asumir que dichos efectos pueden ser representados por un término aditivo $v(t)$ a la salida (ver figura 1.1):

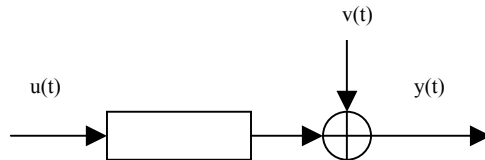


Figura 1.1 Sistema con perturbación

Luego:

$$y(t) = \sum_{k=1}^{\infty} g(\tau).u(t-k) + v(t) \quad (1.6)$$

Hay muchas fuentes y causas de perturbaciones, entre las cuales podemos destacar las diferentes clases de ruido como así también las entradas exógenas al sistema, que se comportan como variables de control pero que no pueden ser controladas por el usuario.

Caracterización de Perturbaciones

La característica más importante de una perturbación es que su valor no se puede conocer de antemano. Sin embargo, información acerca de perturbaciones pasadas podría resultar de mucha utilidad para hacer adecuadas suposiciones acerca de valores futuros. Por lo tanto, como siguiente paso será necesario emplear un marco probabilístico para describir futuras perturbaciones.

Una simple aproximación de $v(t)$ podría ser la siguiente:

$$v(t) = \sum_{k=0}^{\infty} h(k).e(t-k) \quad (1.7)$$

donde $\{e(t)\}$ es una secuencia idénticamente distribuida de variables aleatorias independientes, con una cierta función de densidad que comúnmente se la asume como Normal. Mientras que la secuencia $\{h(k)\}$, como veremos posteriormente, corresponde a la respuesta impulsiva del modelo de ruido. Esta descripción es muy versátil y suficiente para la mayoría de nuestros propósitos prácticos.

Nota: vale aclarar que $\{e(t)\}$ y $\{v(t)\}$, como fueron definidas anteriormente, son *procesos estocásticos*. Pero las perturbaciones que observamos y que son sumadas a la salida del sistema son *realizaciones* del *proceso estocástico* $\{v(t)\}$. O sea que estrictamente hablando deberíamos distinguir en la notación, la diferencia entre el proceso y su realización. No obstante, vamos a hacer un pequeño abuso de la notación a modo de simplificar la lectura, y por ahora no haremos explícita dicha distinción.

Función Covarianza

Con la descripción (1.7) de $v(t)$, podemos calcular su valor esperado como,

$$E\{v(t)\} = \sum_{k=0}^{\infty} h(k) \cdot E\{e(t-k)\} = 0$$

y su función de covarianza como,

$$\begin{aligned} C_v(\tau) &= E\{v(t) \cdot v(t-\tau)\} = \sum_{k=0}^{\infty} \sum_{s=0}^{\infty} h(k) \cdot h(s) \cdot E\{e(t-k) \cdot e(t-\tau-s)\} \\ &= \sum_{k=0}^{\infty} \sum_{s=0}^{\infty} h(k) \cdot h(s) \cdot \delta(k-\tau-s) \cdot \sigma^2 \\ &= \sigma^2 \cdot \sum_{k=0}^{\infty} h(k) \cdot h(k-\tau) \quad (\text{aquí } h(r) = 0 \text{ siempre que } r < 0, \text{ ya que el sistema es causal}) \end{aligned} \quad (1.8)$$

Función de Autocorrelación

Si definimos la función de autocorrelación de una señal determinística como:

$$R_{xx}(\tau) = \int_{-\infty}^{\infty} x(t+\tau) \cdot x(t) dt \quad (1.9a)$$

y si dicha señal es periódica, entonces esta expresión puede ser simplificada de la siguiente forma:

$$R_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t+\tau) \cdot x(t) dt \quad (1.9b)$$

Llegado este punto podemos enunciar el siguiente teorema:

Teorema 1.1 :

sea $x(t)$ un proceso estocástico ergódico, entonces dado un τ :

$$R_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t+\tau) \cdot x(t) dt = E\{x(t+\tau) \cdot x(t)\} = C_x(\tau) \quad (\text{Covarianza } \{x(t)\}) \quad (1.10)$$

Vale aclarar que para procesos ergódicos la autocorrelación de una realización coincide con la covarianza del proceso.

Funciones de Transferencia

A partir de ahora será conveniente introducir una notación de la cual vamos a hacer mucho uso en adelante, y consiste en definir el *operador de adelanto* q de esta forma

$$q \cdot u(t) = u(t+1)$$

y el *operador de atraso* q^{-1}

$$q^{-1} \cdot u(t) = u(t-1)$$

En estas condiciones, (1.5) se puede escribir así:

$$y(t) = \sum_{k=1}^{\infty} g(k) \cdot u(t-k) = \sum_{k=1}^{\infty} g(k) \cdot (q^{-k}u(t)) = \left[\sum_{k=1}^{\infty} g(k) \cdot q^{-k} \right] \cdot u(t) = G(q) \cdot u(t) \tag{1.11}$$

siendo:

$$G(q) = \sum_{k=1}^{\infty} g(k) \cdot q^{-k} \tag{1.12}$$

Y en esta expresión hemos introducido el término $G(q)$ que llamaremos *operador de transferencia* o *función de transferencia* de un sistema lineal como el descripto por (1.5). Ya que (1.11) describe la relación entre las secuencias $y(t)$ y $u(t)$.

Nota: hemos elegido q como argumento de G en lugar de q^{-1} (el cual podría ser más natural) y esto es para ser consistentes con la notación de *transformada z* y *transformada de Fourier*. Además, estrictamente hablando, la denominación de *función de transferencia* deberíamos reservarla para la transformación al plano z de $\{g(k)\}_1^{\infty}$ (secuencia de datos $g(k)$), que es,

$$G(z) = \sum_{k=1}^{\infty} g(k) \cdot z^{-k} \quad \text{con } z = e^{Tj\omega} \tag{1.13}$$

Análogamente podemos trabajar con (1.7) y obtener,

$$H(q) = \sum_{k=0}^{\infty} h(k) \cdot q^{-k} \quad \text{con } q = e^{Tj\omega} \tag{1.14a}$$

o sea que si $T=1$:

$$H(e^{j\omega}) = \sum_{k=0}^{\infty} h(k) \cdot e^{-kj\omega} = \sum_{k=0}^{\infty} h(k) \cdot q^{-k} = H(q) \tag{1.14b}$$

(lo mismo se extiende para $G(q)$ y demás expresiones dependientes de q)

luego,

$$v(t) = H(q) \cdot e(t) \tag{1.15}$$

De modo que nuestra descripción para un sistema lineal con perturbación aditiva será la siguiente:

$$\boxed{y(t) = G(q) \cdot u(t) + H(q) \cdot e(t)} \tag{1.16}$$

siendo $\{e(t)\}$, por lo general, una secuencia de variables aleatorias independientes de media nula y varianza σ^2 .

1.2 Espectro de señal

Ahora desarrollaremos un marco teórico que nos permitirá describir las características espectrales de las señales con las cuales trabajaremos, que van a ser tanto de naturaleza determinística como estocástica. Y en particular será muy común de aquí en adelante que utilicemos señales descriptas como procesos estocásticos con componentes determinísticas. La razón es básicamente que preferimos considerar la secuencia de entrada como determinística o al menos parcialmente, mientras que las perturbaciones sobre el sistema es conveniente que sean descriptas como variables aleatorias. Entonces mirando (1.16) encontramos que:

$$E\{y(t)\} = E\{G(q) \cdot u(t) + H(q) \cdot e(t)\} = G(q) \cdot E\{u(t)\} + H(q) \cdot E\{e(t)\} = G(q) \cdot u(t)$$

O sea que $\{y(t)\}$ no es un proceso estacionario, ya que su esperanza varía con el tiempo.

Para resolver este problema se introduce el concepto de proceso cuasi-estacionario, que no es más que un proceso $\{s(t)\}$ que debe cumplir con las siguientes reglas:

$$\begin{aligned}
 \text{(i)} \quad E\{s(t)\} &= m_s(t) & \text{si} & \quad |m_s(t)| \leq C \quad \forall t & \quad \text{(media acotada)} \\
 \text{(ii)} \quad E\{s(t).s(r)\} &= R_{ss}(t,r) & \text{si} & \quad |R_{ss}(t,r)| \leq C & \quad \text{(covarianza acotada)} \\
 \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N R_{ss}(t, t-\tau) &= R_{ss}(\tau) & & \quad \forall \tau & \quad (1.17)
 \end{aligned}$$

Nota (1): si $\{s(t)\}$ es una secuencia determinística, la esperanza ya no tiene sentido y la idea de función cuasi-estacionaria se refiere a que $\{s(t)\}$ sea una secuencia acotada siempre que el límite de $R_{ss}(\tau)$ exista.

Nota (2): si $\{s(t)\}$ es un proceso estocástico estacionario entonces se satisfacen (i) y (ii) en forma trivial.

Definición de Densidad Espectral de Potencia

Se define densidad espectral de potencia o espectro (en inglés spectra) de $\{s(t)\}$ como,

$$\phi_{ss}(\omega) = \sum_{\tau=-\infty}^{\infty} R_{ss}(\tau).e^{-j\tau\omega}$$

y espectro cruzado entre $\{s(t)\}$ y $\{w(t)\}$

$$\phi_{sw}(\omega) = \sum_{\tau=-\infty}^{\infty} R_{sw}(\tau).e^{-j\tau\omega}$$

siempre y cuando las sumas infinitas existan.

En tanto que $\phi_{ss}(\omega)$ siempre es una función real, y $\phi_{sw}(\omega)$ es por lo general una función compleja de ω .

Espectro de un Proceso Estocástico Estacionario

Si ahora tomamos un proceso estacionario $\{v(t)\}$ de covarianza (1.10), y tenemos en cuenta que según nuestro último desarrollo (1.17) coincide con (1.10), veremos que la definición de espectro adoptada coincide con la convencional. Por lo tanto, si el proceso $v(t)$ viene dado por (1.7) su función de covarianza será (1.8), y su espectro:

$$\begin{aligned}
 \phi_{vv}(\omega) &= \sum_{\tau=-\infty}^{\infty} R_{vv}(\tau).e^{-j\tau\omega} = \sum_{\tau=-\infty}^{\infty} \sigma^2 . e^{-j\tau\omega} \sum_{k=\max(0,\tau)}^{\infty} h(k).h(k-\tau) \\
 &= \sigma^2 . \sum_{\tau=-\infty}^{\infty} \sum_{k=\max(0,\tau)}^{\infty} h(k).e^{-jk\omega} . h(k-\tau).e^{j(k-\tau)\omega} \\
 &= \sigma^2 . \sum_{s=0}^{\infty} h(s).e^{js\omega} . \sum_{k=0}^{\infty} h(k).e^{-jk\omega} = \sigma^2 . |H(e^{j\omega})|^2 \quad \text{(usando (1.14))} \quad (1.18) \\
 &\quad \text{(siendo } s=k-\tau)
 \end{aligned}$$

Este va a ser un resultado muy importante para un uso futuro.

El proceso estocástico descrito por $v(t) = H(q).e(t)$, donde $\{e(t)\}$ es una secuencia de variables aleatorias independientes de media nula y varianza σ^2 , tiene la siguiente densidad espectral de potencia:

$$\phi_{vv} = \sigma^2 . |H(e^{j\omega})|^2 \quad (1.19)$$

Transformación del espectro introducida por un Sistema Lineal**Teorema 1.2 :**

tomemos la secuencia $\{y(t)\}$ dada por:

$$y(t) = G(q).u(t) + H(q).e(t)$$

donde $\{u(t)\}$ es una señal determinística, cuasi-estacionaria y de espectro $\phi_{uu}(\omega)$, mientras que $\{e(t)\}$ es ruido blanco con varianza σ^2 . Entonces, si G y H son filtros estables, $\{y(t)\}$ será cuasi-estacionaria y estará dada por:

$$\begin{aligned} \phi_{yy}(\omega) &= |G(e^{j\omega})|^2 \phi_{uu}(\omega) + \phi_{vv}(\omega) = |G(e^{j\omega})|^2 \phi_{uu}(\omega) + \sigma^2 \cdot |H(e^{j\omega})|^2 \\ \phi_{yu}(\omega) &= G(e^{j\omega}) \cdot \phi_{uu}(\omega) \end{aligned} \quad (1.20)$$

2. MÉTODOS DE IDENTIFICACIÓN NO PARAMÉTRICA

Debido a que un sistema lineal invariante en el tiempo puede ser descrito por su función de transferencia, o bien por su correspondiente respuesta impulsiva, lo que iniciaremos ahora será un camino en busca de algunos métodos que nos permitan determinar estas funciones por técnicas directas y sin hacer una preselección de los posibles modelos. Estos métodos son frecuentemente denominados métodos no paramétricos debido a que no emplean (en forma explícita) un vector finito de parámetros en la búsqueda de una mejor descripción del sistema.

A continuación presentaremos una técnica que puede contener varios niveles de sofisticación llamada "análisis espectral", la cual tiene lugar en el dominio de la frecuencia. Y por otro lado un método muy sencillo denominado "análisis por correlación", que trabaja en el dominio del tiempo.

2.1 Análisis espectral

Este método para determinar funciones de transferencia de sistemas lineales fue desarrollado a partir de métodos de estimación espectral. Y lo que nosotros haremos aquí, será un ligero acercamiento al tema y tomar algunos conceptos básicos para realizar estimadores matemáticos que permitan arribar a un resultado acorde a la situación.

De acuerdo a (1.16), si queremos caracterizar un sistema bastaría con hallar su función de transferencia $G(e^{j\omega})$, como así también su comportamiento ante perturbaciones exógenas $H(e^{j\omega})$. Como resultado inmediato del análisis de (1.20), es claro que estimando las tres densidades espectrales de potencia involucradas (hablamos de ϕ_{uu} , ϕ_{yy} y ϕ_{yu}) es posible estimar $G(e^{j\omega})$, como así también $\phi_{vv}(\omega)$, que a través de (1.19) se vincula con el módulo de $H(e^{j\omega})$.

Es evidente que en este análisis habremos perdido la fase de $H(e^{j\omega})$, sin embargo esta función no es de vital importancia a la hora de analizar el comportamiento del ruido $v(t)$.

Principales pasos a seguir en el análisis:

1) Como primer paso se realizan los estimadores de las funciones de intercorrelación $\hat{R}_{yy}(\tau)$, $\hat{R}_{yu}(\tau)$, $\hat{R}_{uu}(\tau)$, usando:

$$\begin{aligned}\hat{R}_{yy}(\tau) &= \frac{1}{N} \sum_{t=1}^N y(t+\tau) \cdot y(t) \\ \hat{R}_{yu}(\tau) &= \frac{1}{N} \sum_{t=1}^N y(t+\tau) \cdot u(t) \\ \hat{R}_{uu}(\tau) &= \frac{1}{N} \sum_{t=1}^N u(t+\tau) \cdot u(t)\end{aligned}\tag{2.1}$$

2) A partir de esto se forman los estimadores de los correspondientes espectros,

$$\begin{aligned}\hat{\phi}_{yy}(\omega) &= \sum_{\tau=-M}^M \hat{R}_{yy}(\tau) \cdot W_M(\tau) \cdot e^{-j\omega\tau} \\ \hat{\phi}_{yu}(\omega) &= \sum_{\tau=-M}^M \hat{R}_{yu}(\tau) \cdot W_M(\tau) \cdot e^{-j\omega\tau} \\ \hat{\phi}_{uu}(\omega) &= \sum_{\tau=-M}^M \hat{R}_{uu}(\tau) \cdot W_M(\tau) \cdot e^{-j\omega\tau}\end{aligned}\tag{2.2}$$

donde $W_M(\tau)$ es una ventana de tiempo, utilizada para disminuir el ripple que presenta el espectro de la señal al considerar un número finito de muestras. En la actualidad existen varios tipos de ventana, y por lo general, en este tipo de aplicaciones se suele utilizar la ventana de Hamming o la de Bartlett, las cuales poseen una buena relación entre el ancho del lóbulo principal, y la atenuación en bandas laterales. Para nuestro caso, dicha ventana es conocida como *ventana de retardo* y el índice M corresponde al ancho temporal de la misma.

3) Una vez que poseemos la información correspondiente a los tres espectros incógnita, se continúa el proceso estimando $G(e^{j\omega})$ y $\phi_{vv}(e^{j\omega})$:

$$\boxed{\hat{G}_N(e^{j\omega}) = \frac{\hat{\phi}_{yu}(\omega)}{\hat{\phi}_{uu}(\omega)}} \quad ; \quad \hat{\phi}_{vv}(\omega) = \hat{\phi}_{yy}(\omega) - \frac{|\hat{\phi}_{yu}(\omega)|^2}{\hat{\phi}_{uu}(\omega)} \Rightarrow \boxed{|\hat{H}_N(e^{j\omega})| = \frac{\sqrt{\hat{\phi}_{vv}(\omega)}}{\sigma^2}}$$

Este procedimiento es conocido como *análisis espectral*.

Aplicación en Matlab

Para un sistema de una salida y de una o más entradas, la función 'spa' lleva a cabo el análisis espectral de acuerdo al procedimiento anteriormente mencionado.

```
>>[G,FIV] = spa(z)
```

Aquí z es una matriz compuesta de dos columnas que contiene los datos de salida y entrada como primera y segunda columna respectivamente. G y FIV son matrices que contendrán la función estimada \hat{G}_N y la estimación del espectro de perturbación $\hat{\phi}_{vv}$. Las mismas son codificadas en un formato especial llamado *freqfunc format* que contiene información acerca de los valores de frecuencia, amplitud, y fases, como así también sus desviaciones estándar. Esto le va a permitir a uno graficar las estimaciones en forma inmediata usando la función 'bodeplot' o 'ffplot'.

En los cálculos (2.2), 'spa' usa una ventana de Hamming con un ancho por defecto, igual al mínimo entre 30 y un décimo del número de muestras. Pero el ancho M de la ventana puede ser modificado en forma arbitraria de esta forma,

```
>>[G,FIV] = spa(z,M)
```

La regla es que mientras M se incrementa, las funciones de frecuencia estimadas muestran detalles más agudos, pero son también más afectadas por perturbaciones aleatorias. En estos casos, para hacer la elección podría ser de utilidad implementar alguna secuencia típica de comandos que examine diferentes tamaños de ventana.

Nota: la función 'spa' permite la especificación del periodo de muestreo (T), mediante el uso del comando 'sett'.

```
>>G=spa(z)
>>G=sett(G,T)
```

2.2 Análisis por Correlación (estimación de la respuesta impulsiva)

Considerando la descripción del modelo (1.16) podemos describir la salida $\{y(t)\}$ en función de la respuesta al impulso unitario del sistema:

$$y(t) = \sum_{k=1}^{\infty} g(k).u(t-k) + v(t) \tag{2.3}$$

Si la entrada es una secuencia cuasi-estacionaria, de modo que

$$E\{u(t).u(t-\tau)\} = R_{uu}(\tau)$$

y

$$E\{u(t).v(t-\tau)\} = 0$$

entonces de acuerdo al teorema 1.2 (expresado en el dominio del tiempo),

$$E\{y(t).u(t-\tau)\}=R_{yu}(\tau)=\sum_{k=1}^{\infty}g(k).R_{uu}(k-\tau) \tag{2.4}$$

Si escogemos ruido blanco como entrada al sistema, su función de autocorrelación será la siguiente:

$$R_{uu}(\tau) = E\{u(t+\tau).u(t)\}=\sigma^2.\delta(\tau) = \begin{cases} \sigma^2 & \text{si } \tau = 0 \\ 0 & \text{si } \tau \neq 0 \end{cases}$$

Luego, la función de intercorrelación (correlación cruzada) entre la entrada y la salida será:

$$R_{yu}(\tau) = E\{y(t+\tau).u(t)\} = \sigma^2.g(\tau) \tag{2.5}$$

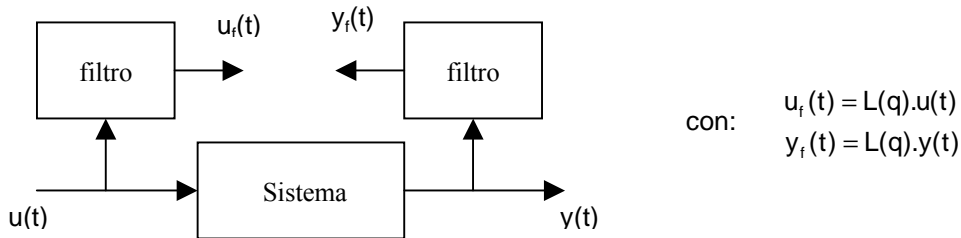
Donde $g(\tau)$ es la respuesta impulsiva del sistema. De este modo será fácil hallar, en función de la expresión anterior, una estimación de la respuesta impulsional:

$$\hat{g}(\tau) = \frac{1}{\sigma^2.N} \sum_{t=1}^N y(t+\tau).u(t) \tag{2.6}$$

Ahora, si en la entrada no disponemos de ruido blanco entonces lo que se utiliza por lo general es un filtro $L(q)$ denominado "blanqueante", cuya acción principal es tomar una secuencia de datos y obtener, a partir de ella, una secuencia no correlacionada de datos, que es una aproximación lo suficiente buena al ruido blanco como para seguir adelante con los cálculos. Por lo tanto la nueva secuencia tendrá esta forma:

$$u_f(t) = L(q).u(t)$$

filtrando la secuencia de salida con el mismo filtro y tomando esos nuevos datos filtrados, se puede obtener una buena estimación de los valores $g(k)$.



$$\begin{aligned} \Rightarrow R_{y_f, u_f}(\tau) &= E\{L(q).y(t+\tau).L(q).u(t)\} = L(q).E\{y(t+\tau).u_f(t)\} = \\ &= L(q).E\left\{\frac{1}{L(q)}y'(t+\tau).u_f(t)\right\} = \sigma_{u_f}^2.g(\tau) \end{aligned} \tag{2.7}$$

A este procedimiento se lo conoce como *análisis por correlación*.

Aplicación en Matlab

El procedimiento del *análisis por correlación* visto con anterioridad se puede implementar con la función 'cra' :

```
>>[ir,R] = cra(z,M,na)
```

Esta función retornará una estimación de la respuesta impulsiva del sistema en el vector *ir*, como así también una estimación de las funciones de autocorrelación de $y_f(t)$ y de $u_f(t)$ en la variable de salida *R*, para así poder evaluar el buen desempeño o no del filtro blanqueador.

Información de la matriz *R*:

- 1º columna: índices de las muestras.
- 2º columna: función de autocorrelación de la (posiblemente filtrada) salida.
- 3º columna: función de autocorrelación de la (posiblemente filtrada) entrada.
- 4º columna: función de intercorrelación $y_f(t)$ con $u_f(t)$.

El valor de *M* define el número de muestras para las cuales la autocorrelación será calculada. Esto es desde $-M$ hasta M , o sea que el largo de la correlación es $2M+1$ y la respuesta impulsiva es calculada desde 0 a M (el programa por defecto toma $M=20$).

Debido a que la entrada "preblanqueada" es ajustada a un modelo paramétrico "ar" de orden *na* (como los que se analizarán en los próximos capítulos), el tercer valor a definir es justamente dicho índice (el algoritmo toma por defecto $na = 20$).

Por último lo que se puede hacer, en base a los datos del vector *ir*, es utilizar el comando 'cumsum' que realiza una suma acumulativa de los elementos de un vector, para determinar en forma aproximada la respuesta al escalón del sistema.

```
>>sr = cumsum(ir)
```

3. PREDICCIÓN

Las descripciones del sistema dadas con anterioridad pueden ser utilizadas para resolver una gran variedad de problemas de diseño relacionados con los sistemas reales. Sin embargo, a los fines prácticos suele ser conveniente llevar a cabo el modelado en función de una cantidad fija de variables. En este sentido, la idea de predecir futuros valores de salida se va a tornar en un tema de esencial importancia en el desarrollo de métodos de identificación, es por eso que a continuación presentaremos las ideas básicas para hallar un predictor de la salida.

En una primera instancia deberemos desarrollar un método que nos permita predecir valores futuros de $v(t)$ en el caso de que su descripción sea:

$$v(t) = H(q).e(t) = \sum_{k=0}^{\infty} h(k).e(t-k) \quad (3.1)$$

donde asumimos que $H(q)$ es estable, o sea que el filtro no posee polos fuera del círculo unitario, debido a que:

$$\sum_{k=0}^{\infty} |h(k)| < \infty \quad (3.2)$$

Invertibilidad del Modelo de Ruido

Una propiedad muy importante de (3.1), la cual vamos a imponer, es que sea *invertible*. Esto quiere decir que si es conocida $v(t')$ para todo $t' \leq t$, entonces es posible hallar $e(t)$ como

$$e(t) = H^{-1}(q).v(t) = \sum_{k=0}^{\infty} \tilde{h}(k).v(t-k) \quad (3.3)$$

con

$$\sum_{k=0}^{\infty} |\tilde{h}(k)| < \infty \quad (3.4)$$

De esta forma se define el filtro H^{-1}

$$H^{-1}(q) = \frac{1}{H(q)} = \sum_{k=0}^{\infty} \tilde{h}(k).q^{-k} \quad (3.5)$$

Nota: para que la función $1/H(q)$ sea analítica en $|q| \geq 1$ es necesario que no tenga polos fuera del círculo unitario, esto quiere decir que $H(q)$ no debe tener ceros fuera del círculo unitario.

Predicción de v

Supongamos ahora que hemos observado $v(t')$ para $t' \leq t-1$ y que queremos predecir el valor de $v(t)$ basándonos en estas observaciones. Tenemos entonces, a partir de que el polinomio H es mónico,

$$v(t) = \sum_{k=0}^{\infty} h(k).e(t-k) = e(t) + \sum_{k=1}^{\infty} h(k).e(t-k) \quad (3.6)$$

Surge ahora trabajar con la *esperanza condicional* de $v(t)$ que se denota $\hat{v}(t|t-1)$, que se refiere al valor que se predice de $v(t)$ con la información obtenida de los instantes anteriores a t . Como el valor esperado de $e(t)$ es nulo, entonces:

$$\hat{v}(t|t-1) = \sum_{k=1}^{\infty} h(k).e(t-k) \quad (3.7)$$

por lo tanto

$$\hat{v}(t | t-1) = \left[\sum_{k=1}^{\infty} h(k) \cdot q^{-k} \right] \cdot e(t) = [H(q) - 1] \cdot e(t) = \frac{H(q) - 1}{H(q)} \cdot v(t) = [1 - H^{-1}(q)] \cdot v(t) \quad (3.8)$$

Predicción de y

Considerando una descripción del sistema como la (1.16), y asumiendo que tanto $y(t')$ como $u(t')$ son conocidas para $t' \leq t-1$ podemos decir que

$$v(t') = y(t') - G(q) \cdot u(t') \quad (3.9)$$

o sea que $v(t')$ también es conocida para $t' \leq t-1$. Por lo tanto la esperanza condicional de $y(t)$, dada la información anterior, es

$$\begin{aligned} \hat{y}(t | t-1) &= G(q) \cdot u(t) + \hat{v}(t | t-1) \\ &= G(q) \cdot u(t) + [1 - H^{-1}(q)] \cdot v(t) \\ &= G(q) \cdot u(t) + [1 - H^{-1}(q)] \cdot [y(t) - G(q) \cdot u(t)] \end{aligned}$$

Luego

$$\boxed{\hat{y}(t | t-1) = H^{-1} \cdot G(q) \cdot u(t) + [1 - H^{-1}(q)] \cdot y(t)} \quad (3.10)$$

Esta será, de aquí en adelante, la expresión que utilizaremos para predecir los valores futuros de la salida del sistema.

Nota: es importante destacar que hemos arribado a dicha ecuación sin hacer mayores consideraciones probabilísticas.

4. MODELOS PARAMÉTRICOS DE SISTEMAS LINEALES INVARIANTES EN EL TIEMPO

Un modelo de un sistema consiste en una descripción conveniente de algunas de sus propiedades, y de acuerdo a un propósito particular. El modelo no necesita ser una exacta descripción del sistema, y el usuario debe saber esto para poder llevar a cabo su propósito.

El primer paso en la identificación de sistemas es determinar una clase de modelos, dentro de la cual se hallará el modelo más conveniente.

4.1 Modelos lineales y sets de modelos lineales

Un modelo lineal invariante en el tiempo, como vimos con anterioridad, puede ser especificado por su respuesta impulsiva $\{g(k)\}_1^\infty$, el espectro $\phi_w(\omega) = \sigma^2 |H(e^{j\omega})|^2$ de la perturbación aditiva, y, posiblemente por la función de densidad de probabilidad (fdp) de la perturbación $e(t)$. Por lo tanto un modelo completo estaría dado por:

$$y(t) = G(q).u(t) + H(q).e(t) \quad (4.1)$$

$$f_e(x), \text{ fdp de } e(t)$$

con,

$$G(q) = \sum_{k=1}^{\infty} g(k).q^{-k}, \quad H(q) = 1 + \sum_{k=1}^{\infty} h(k).q^{-k} \quad (4.2)$$

En la mayoría de los casos es impracticable el hecho de hacer esta especificación tomando las secuencias infinitas $\{g(k)\}$ y $\{h(k)\}$ conjuntamente con la función $f_e(x)$. En lugar de esto uno elige trabajar con estructuras que permitan la especificación de G y H en términos de un número finito de valores. Tal es así que las funciones de transferencia racionales son típicos ejemplos de esto. Incluso es muy común que la *fdp* no sea especificada como función, pero descrita en función de unas pocas características numéricas como los momentos de primero y segundo orden:

$$E\{e(t)\} = \int x.f_e(x).dx = 0 \quad (4.3a)$$

$$E\{e^2(t)\} = \int x^2.f_e(x).dx = \sigma^2 \quad (4.3b)$$

Es también muy común asumir que $e(t)$ tiene una distribución Gaussiana, en cuyo caso la *fdp* queda enteramente especificada por (4.3).

La especificación (4.1) en términos de un número finito de valores, o coeficientes, se conoce como modelado paramétrico e involucra métodos de estimación y predicción conocidos *métodos de identificación paramétrica*.

Muy a menudo no es posible determinar a priori estos coeficientes partiendo del conocimiento de las leyes físicas que gobiernan el comportamiento del sistema. Es por eso que la determinación de todos o algunos de dichos valores debe ser dejada a los procedimientos de estimación. Esto quiere decir que los coeficientes en cuestión entran en el modelo como *parámetros a ser determinados*. Llamaremos entonces θ al vector que contenga todos los parámetros a estimar. O sea que ahora la descripción del modelo será la siguiente:

$$y(t, \theta) = G(q, \theta).u(t) + H(q, \theta).e(t) \quad (4.4a)$$

$$f_e(x, \theta), \text{ fdp de } e(t) ; \{e(t)\} \text{ ruido blanco} \quad (4.4b)$$

Vale destacar que este no es un modelo sino un *set de modelos*, aunque por lo general vamos a hablar despreocupadamente del "modelo (4.4)" aunque sea un abuso de notación desde un punto de vista formal.

Usando (3.10), vamos a poder predecir las muestras futuras para (4.4), y para enfatizar la dependencia que mantiene el estimador con el vector de parámetros θ , la notación se hace de la siguiente manera :

$$\hat{y}(t, \theta) = H^{-1}(q, \theta).G(q, \theta).u(t) + [1 - H^{-1}(q, \theta)].y(t) \tag{4.5}$$

Es notable que la forma de este predictor no depende de $f_e(x, \theta)$ ya que, como hemos expresado anteriormente, es posible llegar a (4.5) sin hacer consideraciones probabilísticas.

A continuación serán analizados diferentes modos de describir (4.4a) en términos de θ .

4.2 Familias de modelos de funciones de transferencia

Tal vez el camino más inmediato para parametrizar G y H sería representarlas como funciones racionales y dejar que los parámetros sean los coeficiente del numerador y del denominador. En lo que sigue presentaremos varias estructuras modelo conocidas como *modelos de caja negra*.

Estructura ARX

Probablemente la relación entrada-salida más simple que se puede obtener sea la proveniente de una descripción del sistema como una ecuación lineal en diferencias.

$$y(t) + a_1y(t - 1) + a_2y(t - 2) + \dots + a_{n_a}y(t - n_a) = b_1u(t - 1) + b_2u(t - 2) + \dots + b_{n_b}u(t - n_b) + e(t) \tag{4.6}$$

Debido a que el término de ruido blanco $e(t)$ entra como un error directo en la ecuación en diferencias, el modelo (4.6) es también conocido como *modelo o estructura de ecuación de error*.

En este caso los parámetros a ajustar serán:

$$\theta = [a_1 \ a_2 \ \dots \ a_{n_a} \ b_1 \ b_2 \ \dots \ b_{n_b}]^T \tag{4.7}$$

si proponemos dos polinomios A(q) y B(q) de la forma,

$$A(q) = 1 + a_1 \cdot q^{-1} + a_2 \cdot q^{-2} + \dots + a_{n_a} \cdot q^{-n_a}$$

$$B(q) = b_1 \cdot q^{-1} + b_2 \cdot q^{-2} + \dots + b_{n_b} \cdot q^{-n_b}$$

vemos que (4.6) se corresponde con (4.4), siendo,

$$G(q, \theta) = \frac{B(q)}{A(q)} \ , \quad H(q, \theta) = \frac{1}{A(q)} \tag{4.8}$$

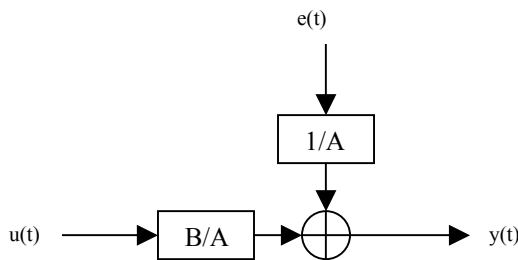


Figura 4.1 Estructura ARX

A este modelo se lo conoce como estructura “ARX”, donde “AR” hace referencia a la parte autorregresiva $A(q).y(t)$ y “X” al la entrada extra (extra input) $B(q).u(t)$ también conocida como variable exógena.

El flujo de señal de la Figura 4.1 nos indica que posiblemente este no sea el modelo más natural desde un punto de vista físico, ya que el ruido blanco es sumado a la salida luego de pasar a través del denominador del sistema dinámico. Sin embargo, el set de modelos de ecuación de error posee una propiedad importante que lo convierte en una buena primera elección en muchas aplicaciones. Y es que la forma del predictor define una regresión lineal.

Regresiones lineales:

En primer lugar hallemos un predictor para (4.6) reemplazando (4.8) en (4.5):

$$\hat{y}(t, \theta) = B(q, \theta).u(t) + [1 - A(q, \theta)].y(t) \quad (4.9)$$

Claramente, esta expresión podría haber sido hallada sin mayores problemas a partir de (4.6). O sea que, sin un marco probabilístico, el predictor (4.9) podría ser una elección natural si es que el término $e(t)$ en (4.6) es considerado "insignificante" o bien "difícil de estimar". Si somos un poco más estrictos podemos decir que a la hora de estimar el término $e(t)$, la mejor estimación del mismo es considerarlo nulo ya que la esperanza de un proceso con características de ruido blanco es nula. Es entonces perfectamente natural trabajar con (4.9) incluso para modelos determinísticos.

Introducimos ahora el vector

$$\varphi(t) = [-y(t-1) \quad -y(t-2) \quad \dots \quad -y(t-n_a) \quad -u(t-1) \quad -u(t-2) \quad \dots \quad -u(t-n_b)]^T \quad (4.10)$$

Así podemos rescribir (4.9) como

$$\hat{y}(t, \theta) = \theta^T \varphi(t) = \varphi^T(t) \theta \quad (4.11)$$

Esta es la importante propiedad de la que hacíamos referencia previamente. El predictor es un producto escalar entre un vector de datos conocidos $\varphi(t)$ y el vector de parámetros θ . A este modelo se lo conoce en estadística como *regresión lineal*, y al vector $\varphi(t)$ se lo llama *vector de regresión*. La importancia de esto radica en la simplicidad y eficacia de los métodos de estimación que pueden ser aplicados para la determinación de θ , como por ejemplo el criterio de cuadrados mínimos:

$$\epsilon(t, \theta) = y(t) - \hat{y}(t) = y(t) - \varphi^T(t) \theta \quad (4.12a)$$

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{t=1}^N \frac{1}{2} [y(t) - \varphi^T(t) \theta]^2 \quad (4.12b)$$

que se trata de ir modificando el vector θ con el objeto de minimizar la función V_N , o lo que es lo mismo reducir la diferencia entre $y(t)$ y su valor predicho. Es importante notar que el camino para hallar un modelo adecuado conduce necesariamente a cálculos iterativos.

Estructura ARMAX

La principal desventaja del modelo (4.6) reside en la escasez o falta de libertad en la descripción del término de perturbación. Sin embargo, es posible incorporar mayor flexibilidad al modelado si es que agregamos un término conocido como media en movimiento (moving average) del ruido blanco.

$$y(t) + a_1 y(t-1) + \dots + a_{n_a} y(t-n_a) = b_1 u(t-1) + \dots + b_{n_b} u(t-n_b) + e(t) + c_1 e(t-1) + \dots + c_{n_c} e(t-n_c) \quad (4.13)$$

con

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

Por lo tanto podemos escribir

$$A(q).y(t) = B(q).u(t) + C(q).e(t) \quad (4.14)$$

dicha ecuación se corresponde con (4.4) si

$$G(q, \theta) = \frac{B(q)}{A(q)}, \quad H(q, \theta) = \frac{C(q)}{A(q)} \quad (4.15)$$

siendo ahora el nuevo vector de parámetros:

$$\theta = [a_1 \ a_2 \ \dots \ a_{n_a} \ b_1 \ b_2 \ \dots \ b_{n_b} \ c_1 \ c_2 \ \dots \ c_{n_c}]^T \tag{4.16}$$

Debido al nuevo término de media en movimiento $C(q)e(t)$, el modelo (4.14) será llamado ARMAX. Este modelo se ha convertido en una herramienta estándar en control y econometría, tanto para la descripción de sistemas como para el diseño de control.

Nota: una versión de este modelo con una integración forzada se lo conoce como modelo ARIMAX, el cual es muy usado para describir sistemas con perturbaciones lentas. Y se obtiene reemplazando $\Delta y(t) = y(t) - y(t-1)$ en lugar de $y(t)$ en la ecuación (4.14).

Estructura ARARX

En lugar de modelar el término de ruido de la ecuación (4.6) con una media en movimiento, como hicimos en (4.13), se puede describir esta perturbación como una autorregresión, y de aquí el nombre del modelo.

$$A(q).y(t) = B(q).u(t) + \frac{1}{D(q)}.e(t) \tag{4.17}$$

con

$$D(q) = 1 + d_1q^{-1} + d_2q^{-2} + \dots + d_{n_d}q^{-n_d}$$

Estructura ARARMAX

Un forma más general aún es la que introduce la descripción ARMA en el modelo anterior, dando como resultado la estructura ARARMAX.

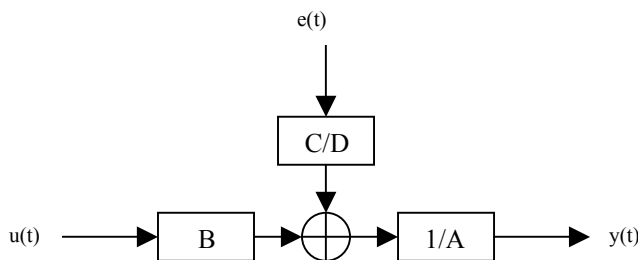


Figura 4.2 Estructura ARARMAX

$$A(q).y(t) = B(q).u(t) + \frac{C(q)}{D(q)}.e(t) \tag{4.18}$$

Estructura de Error de Salida

Hasta ahora hemos visto estructuras en las cuales la descripción de las funciones de transferencia tienen al polinomio $A(q)$ como factor común en sus denominadores. Pero desde un punto de vista físico sería más natural parametrizar estas transferencias en forma independiente.

Supongamos entonces por un momento que la relación entre la entrada y una salida no perturbada $w(t)$ puede ser representada como una ecuación en diferencias lineal, y que la perturbación consiste en ruido blanco.

$$w(t) + f_1w(t-1) + \dots + f_{n_f} w(t-n_f) = b_1u(t-1) + \dots + b_{n_b} u(t-n_b) \tag{4.19a}$$

$$y(t) = w(t) + e(t) \tag{4.19b}$$

Con

$$F(q) = 1 + f_1q^{-1} + f_2q^{-2} + \dots + f_{n_f} q^{-n_f}$$

podemos escribir la ecuación del modelo como:

$$y(t) = \frac{B(q)}{F(q)} \cdot u(t) + e(t) \tag{4.20}$$

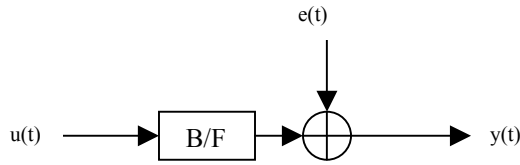


Figura 4.3 Estructura de Error de Salida

Estructura Box-Jenkins

El desarrollo del modelo de error de salida (4.20) es para adicionar al modelo las propiedades del ruido en la salida. Pero si describimos esto mismo en el marco de un modelo ARMA llegamos al siguiente resultado:

$$y(t) = \frac{B(q)}{F(q)} \cdot u(t) + \frac{C(q)}{D(q)} e(t) \tag{4.21}$$

Siendo esta la parametrización de dimensiones finitas más natural si es que partimos de (4.4). Ya que Las funciones de transferencia G y H son parametrizadas en forma independiente como funciones racionales. Entonces, de acuerdo a (4.5) el predictor para el set de modelos (4.21) será:

$$\hat{y}(t, \theta) = \frac{D(q) \cdot B(q)}{C(q) \cdot F(q)} \cdot u(t) + \frac{C(q) - D(q)}{C(q)} y(t) \tag{4.22}$$

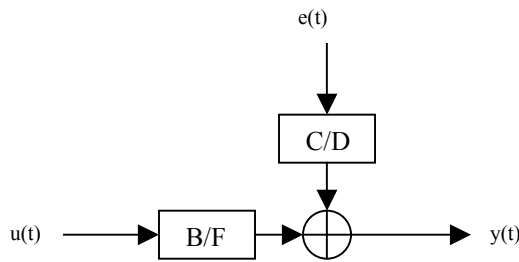


Figura 4.4 Estructura Box-Jenkins

Familia General de Estructuras

Las estructuras que hemos desarrollado anteriormente pueden dar lugar a 32 sets de modelos diferentes, dependiendo de cual de los cinco polinomios A, B, C, D o F son usados, aunque solo hemos mostrado seis de esas posibilidades. Entonces por conveniencia adoptaremos una estructura generalizada.

$$A(q) \cdot y(t) = \frac{B(q)}{F(q)} \cdot u(t) + \frac{C(q)}{D(q)} e(t) \tag{4.23}$$

Para propósitos prácticos, uno o más polinomios deberán ser llevados a la unidad debido a que este modelo es muy general para muchos casos. Sin embargo, el desarrollo de algoritmos y resultados para (4.23) también cubre todos los casos especiales correspondientes a sets de modelos más realísticos.

A partir de (4.5) sabemos que el predictor para (4.23) es:

$$\hat{y}(t, \theta) = \frac{D(q) \cdot B(q)}{C(q) \cdot F(q)} \cdot u(t) + \left[1 - \frac{D(q) \cdot A(q)}{C(q)} \right] y(t) \tag{4.24}$$

Comentario: muchas veces la dinámica desde $u(t)$ a $y(t)$ contiene un retardo, que para el caso discreto, puede ser contabilizado como un retardo de n_k muestras. Como resultado de esto, los primeros n_k coeficientes del polinomio $B(q)$ se hacen idénticamente nulos. Esto es:

$$B(q) = b_{n_k} \cdot q^{-n_k} + b_{n_k+1} \cdot q^{-n_k-1} + \dots + b_{n_k+n_b-1} \cdot q^{-n_k-n_b+1} = q^{-n_k} \bar{B}(q) \quad , \quad \text{siendo } b_{n_k} \neq 0$$

En consecuencia sería una buena idea expresar dicho retardo en forma explícita:

$$A(q).y(t) = q^{-n_k} \cdot \frac{\bar{B}(q)}{F(q)} \cdot u(t) + \frac{C(q)}{D(q)} e(t) \quad (4.25)$$

siendo n_k el número de muestras de retardo entre la entrada y la salida.

Aplicación en Matlab

El Toolbox de Identificación de Sistemas de Matlab contiene varias funciones para la estimación paramétrica de modelos. Todas ellas comparten la misma estructura de comando

```
>>mod= function ([y u],N);
```

Las variables de entrada "y" y "u" son vectores columna que contienen los datos de salida y entrada respectivamente, mientras que la matriz N especifica la estructura particular de los modelos a ser estimados. El modelo estimado resultante será contenido en mod, además de ser codificado en formato *theta*, que es el formato básico para la representación de modelos en este Toolbox. Dicho formato guarda la información acerca de la estructura del modelo y de sus órdenes, retardos, parámetros, y varianzas estimadas de los parámetros estimados en una matriz. El formato *theta* puede ser convertido a cualquier otro modelo que nos resulte de utilidad.

Comúnmente, para presentar la información del modelo se usa el comando 'present' :

```
>>mod= function ([y u],N);
>>present (mod)
```

Modelos ARX

Para estimar los parámetros a_i y b_i del modelo ARX (4.6) se usa la función 'arx' :

```
>>mod = arx (z,[n_a n_b n_k]);
```

Aquí n_a y n_b son los órdenes de $A(q)$ y $B(q)$ respectivamente. Mientras que n_k es el número de retardos entre la entrada y la salida definidos en la estructura. La función 'arx' implementa el método de estimación de mínimos cuadrados.

Familia General de Modelos de Caja Negra

Basándose en la estructura (4.25) es posible construir modelos de prácticamente cualquier tipo. Para esto existe la función 'pem' :

```
>>mod = pem (z,nn);
```

donde N determina todos los órdenes y retardos

```
>>N = [n_a n_b n_c n_d n_f n_k];
```

El comando 'pem' cubre todos los casos de modelos de sistemas lineales tipo caja negra. Sin embargo existen rutinas más eficientes para algunos casos especiales:

```
>>mod = armax (z,[n_a n_b n_c n_k]);           (estructura ARMAX)
>>mod = oe (z,[n_b n_f n_k]);                 (estructura de Error de Salida)
>>mod = bj (z,[n_b n_c n_d n_f n_k]);         (estructura Box-Jenkins)
```

5. SELECCIÓN DE ESTRUCTURA Y VALIDACIÓN DEL MODELO

La elección de una estructura apropiada (por ejemplo: arx, armax, etc.) es un paso crucial en el camino de la identificación de sistemas dinámicos. Y dicha elección debe estar basada tanto en el entendimiento del proceso de identificación, como en un acabado conocimiento del sistema a identificar. Con anterioridad hemos descrito una serie de estructuras típicas para ser usadas en identificación, y lo que nos proponemos ahora es complementar dichas listas con una discusión detallada acerca de cómo arribar a un modelo, guiados tanto por el conocimiento del sistema como por el set de datos recolectados.

Una vez que la estructura ha sido escogida, el próximo paso consiste en seleccionar un modelo particular de dicha estructura, es decir, determinar un conjunto particular de parámetros para dicha estructura. Concluido este paso el modelo hallado puede llegar a ser el mejor disponible, sin embargo, más importante aún es saber a ciencia cierta si dicho modelo es suficientemente bueno para nuestros propósitos.

El proceso de evaluación de un modelo para determinar si es el apropiado, es conocido como *validación del modelo*.

5.1 Aspectos generales en la elección de una estructura modelo

El camino hacia una estructura particular involucra, al menos, tres pasos:

1º paso) Elección del *tipo* del set de modelos. (5a)

Esto involucra, por ejemplo, la selección entre modelos lineales y no lineales, entre caja negra y modelos en variables de estado parametrizados físicamente, y así siguiendo.

2º paso) Elección del *tamaño* del set de modelos. (5b)

Aquí se plantean varios temas como la selección de los grados de los polinomios en el modelo (4.23) y el problema de que tipo de variables incluir en la descripción del modelo.

3º paso) Elección de la parametrización del modelo. (5c)

Cuando un set de modelos ha sido seleccionado, todavía resta parametrizarlo, esto es, encontrar una estructura particular que se adecue al set de datos disponible y a la aplicación buscada.

El propósito fundamental de un proceso de identificación es, en pocas palabras, *obtener un modelo bueno y confiable con una cantidad razonable de trabajo*. Es por eso que la elección de una estructura, seguramente tendrá un considerable efecto tanto en la calidad del modelo resultante, como en el precio que significó arribar al mismo.

Calidad del Modelo

La calidad del modelo resultante puede, por ejemplo, ser medida por un criterio de reducción de *bias* (polarización de los parámetros), que consiste básicamente en la implementación de grandes estructuras para obtener una mayor flexibilidad. Pero a su vez, el hecho de aumentar el número de parámetros produce un incremento en la varianza (que sería algo así como un grado indeterminación) de los parámetros estimados. Es por todo esto que se llega a una situación conflictiva de requerimientos produciéndose un compromiso entre:

- *Flexibilidad*: que consiste en el empleo de estructuras modelo que ofrezcan una gran capacidad de descripción de posibles sistemas. La flexibilidad se puede incrementar de dos formas básicas, usando muchos parámetros o bien ubicando los mismos en posiciones estratégicas.
- *Parsimonia*: que es el no hacer uso de una gran cantidad de parámetros en forma innecesaria.

Precio del Modelo

El precio de un modelo está asociado con el esfuerzo realizado para arribar al mismo, esto es, tanto para realizar la minimización requerida por el método de cuadrados mínimos, como para calcular las correspondientes estimaciones. Este trabajo en particular es altamente dependiente de la estructura adoptada.

Otro precio asociado es el hecho de que un modelo complejo es más difícil de ser usado para hacer simulación y diseño de control. En cambio, el uso de un modelo marginalmente mejor a uno simple, implica por lo general, un buen precio de modelado.

Consideraciones generales

La elección final de la estructura será entonces un compromiso entre (5a), (5b) y (5c). Mientras que las técnicas y consideraciones utilizadas para la evaluación de dichos aspectos pueden ser divididas en cuatro diferentes categorías:

- *Consideraciones a priori:* Ciertos aspectos son independientes del set de datos z y pueden ser evaluados a priori, antes de que sean analizados los datos.
- *Técnicas basadas en análisis preliminares de los datos:* Con los datos disponibles, se pueden cumplimentar ciertos tests y evaluaciones que nos brindarán un panorama certero de las posibles estructuras a usar.
- *Comparación de diferentes estructuras:* Antes de que el modelo final sea elegido, es aconsejable informarse acerca de otras estructuras diferentes y comparar sus características.
- *Validación del modelo obtenido:* Independientemente de cómo ha sido obtenido el modelo, siempre es posible utilizar el set de datos z para evaluar y determinar si el mismo se comporta en forma útil y adecuada para nuestros propósitos finales.

5.2 Consideraciones a priori

Tipo del Modelo

La elección del *tipo* de modelo a utilizar es bastante subjetiva e involucra varias consideraciones independientes del set de datos z . Esta selección suele ser el resultado de un compromiso entre los items mencionados anteriormente, combinado con factores más racionales como la disponibilidad de softwares y la familiaridad que tengamos con ciertos modelos.

El compromiso entre parsimonia y flexibilidad es el corazón de un problema típico de identificación, en consecuencia, una pregunta inmediata que nos puede surgir sería: como podemos hacer un buen ajuste de los datos con pocos parámetros?. Siendo la respuesta más natural a este interrogante “usando conocimiento previo, intuición e ingenio”. Dichas circunstancias ponen en relieve que la Identificación está lejos de ser un proceso totalmente automático, y su complejidad guarda estricta dependencia con la aplicación buscada.

Para un sistema físico, es natural pensar que la mejor forma de incorporar datos al modelo sería si este estuviera modelado en forma continua. Pero esto significaría que los cálculos para hallar $\varepsilon(t, \theta)$ (ecuación 4.12a) y para su posterior minimización, se conviertan en una tarea excesivamente laboriosa. En consecuencia, es de esperar que dichos aspectos de complejidad en los algoritmos favorezcan el uso de modelos discretos tipo caja negra, e incluso induzcan al uso de modelos como el (4.23) que se adapten al set de datos, sin por ello tener que imponer alguna interpretación especial a los valores de los parámetros.

Un consejo general que suelen dar algunos autores es “en principio tratar con cosas simples”, y esto es porque uno debería, en lo posible, adoptar estructuras sofisticadas solo en el caso que las estructuras más sencillas no pasen los tests de validación. Por lo general, modelos que aceptan regresiones lineales como (4.11) conducen a simples y robustos esquemas de minimización como la de mínimos cuadrados. En consecuencia, las estructuras como ARX son a menudo muy buenas primeras elecciones para afrontar un problema de identificación.

Por último, vale destacar que es necesario contemplar alguna eventual no linealidad que pudiera presentar el sistema, ya que si la hubiera, resultaría muy difícil ajustar los datos a un sistema lineal. En particular, los efectos de no linealidad de actuadores y sensores pueden ser conocidos y luego ser usados para redefinir las señales de entrada y salida.

Orden del Modelo

La resolución del problema (5b) normalmente requiere hacer uso de los datos. Sin embargo, un entendimiento físico del sistema, puede en muchos casos indicarnos el rango de ordenes que deberían ser considerados para el modelo. Incluso cuando los datos ya han sido tomados, el conocimiento del número total de ellos (N) y su calidad, indicará cuantos parámetros son factibles de ser estimados. Por ejemplo si disponemos de unos pocos datos, no es razonable tratar de determinar un modelo a partir de una estructura compleja.

Normalmente el periodo de muestreo, además de tener que cumplir con el teorema de Nyquist, suele estar fuertemente influenciado por el diseño de control a utilizar. Y esto se debe a que el periodo de muestreo con el cual modelamos el sistema, debería ser el mismo que se usa para la aplicación de control, al menos que estemos dispuestos a recalcularlo durante los intervalos existentes entre muestras sucesivas.

Otro problema relacionado suele ser que, por problemas numéricos, sea difícil describir adecuadamente más de dos o tres décadas del rango de frecuencia con un mismo modelo. Por otro lado, consideraciones acerca de periodo de muestreo, excitación, y longitud del registro de datos, sugieren que por lo general uno no deba apuntar a cubrir más de tres décadas de las constantes de tiempo en un experimento. De modo que si el sistema contiene constantes de tiempo ampliamente separadas, el consejo es construir dos o más modelos, cada uno cubriendo una parte determinada del rango de frecuencias y siendo muestreado con un periodo adecuado. En un modelo de alta frecuencia, la dinámica de baja frecuencia se comporta, para la mayoría de los casos prácticos, como un bloque con tantos integradores como polos de baja frecuencia tenga el sistema; pensar en el diagrama de Bode. Análogamente, las altas frecuencias se ven como relaciones instantáneas (estáticas) en los modelos de baja frecuencia.

5.3 Selección de Estructura basada en análisis preliminares de los datos

Por análisis preliminar de los datos, queremos decir cálculos que no involucran la determinación de un modelo completo del sistema.

Estimación del Tipo del modelo

Generalmente, algunas técnicas no paramétricas pueden ser útiles para descubrir transformaciones no lineales de los datos, como así también para indicar el tipo de dependencias existentes entre las variables involucradas. Por lo general, dichos tests están basados en relaciones entre correlaciones de orden superior (mayor a segundo orden) y espectros que obedecen a descripciones lineales.

Estimación del Orden del modelo

El orden de un sistema lineal puede ser estimado de varias formas diferentes. En particular, los métodos basados en análisis previos de los datos caen dentro de las siguientes categorías.

1. Examinación de la función de transferencia estimada por el método de Análisis Espectral.
 2. Evaluación de rangos en matrices de covarianza.
 3. Correlación de variables.
1. *Estimación por análisis espectral:* Una estimación no paramétrica de la función de transferencia $G_N(e^{j\omega})$ nos entregará valiosa información acerca de los picos de resonancia, caídas (roll-off) y corrimientos de fase. De esta forma obtendremos valiosos indicios acerca del orden que requerirá el modelo para describir la dinámica del sistema (o al menos la parte que nos interesa). Habría que notar que por lo general un diagrama de bode de tiempo discreto presenta ciertas anomalías

si se lo compara con uno de tiempo continuo, es por eso que es conveniente tomar dichas observaciones con especial recaudo.

2. *Evaluación de la Matriz de Covarianza* : Supongamos que el sistema real está descrito por:

$$y(t) + a_1 y(t-1) + \dots + a_n y(t-n) = b_1 u(t-1) + \dots + b_n u(t-n) + v_0(t) \quad (5.1a)$$

para alguna secuencia de ruido $\{v_0(t)\}$. Supongamos incluso que n es el menor valor para el cual se sigue cumpliendo la ecuación (5.1a), o sea que n es el "orden verdadero" del modelo. Como siempre, tomemos:

$$\varphi_s(t) = [-y(t-1) \dots -y(t-s) \ u(t-1) \dots u(t-s)]^T \quad (5.1b)$$

Consideremos en principio que $v_0(t) = 0$. Esto implica que la matriz de covarianza, definida como:

$$R^s(N) = \frac{1}{N} \sum_{t=1}^N \varphi_s(t) \cdot \varphi_s^T(t) \quad (5.1c)$$

no será singular para $s \leq n$, lo que significaría que todavía es posible aumentar el orden del modelo, pues no hemos alcanzado aún el "orden verdadero" del mismo. Mientras que $R^s(N)$ será singular si $s \geq n+1$, o sea cuando nos hayamos excedido del número de parámetros natural del sistema. Dicho de otro modo esto quiere decir que es posible utilizar una función $k(s)=R^s(N)$ para hacer una evaluación cuantitativa del orden del modelo.

En el caso que se presente ruido $\{v_0(t)\}$ en la ecuación (5.1a), la matriz (5.1c) puede seguir siendo usada si es que se establece un umbral adecuado, y si la relación señal a ruido es lo suficientemente alta.

Pero cuando la influencia de $v_0(t)$ sobre el sistema deja de ser insignificante, una alternativa válida sería utilizar vectores de correlación. Esto es, si $\{v_0(t)\}$ y $\{u(t)\}$ no están correlacionadas, podríamos usar,

$$\zeta_s(t) = [u(t-1) \ u(t-2) \dots u(t-2s)]^T$$

y luego hallar:

$$R_\zeta^s(N) = \frac{1}{N} \sum_{t=1}^N \varphi_s(t) \cdot \zeta_s^T(t)$$

que será no singular para $s \leq n$, y singular para $s \geq n+1$.

3. *Correlación de variables*: El problema de la determinación del orden radica en el hecho de incluir o no una variable más en la estructura modelo. Esta variable podría ser $y(t-n-1)$ en (5.1a) o posiblemente una variable perturbación $w(t)$ que ha sido medida. En cualquier caso, lo importante es saber si esa nueva variable tiene algo que contribuir en la definición de la variable de salida $y(t)$. Esto se puede medir haciendo la correlación entre $y(t)$ y $w(t)$. Sin embargo, para desestimar la posible relación entre $y(t)$ y $w(t)$, calculada a partir de la estructura más pequeña, la correlación debe ser medida entre $w(t)$ y lo que resta por ser conocido, o sea los residuos $\mathcal{E}(t, \hat{\theta}_N) = y(t) - \hat{y}(t, \hat{\theta}_N)$. Esto se conoce en el análisis por regresión como *correlación canónica* o *correlación parcial*.

5.4 Comparación de estructuras modelo

Una aproximación muy natural para buscar una estructura adecuada, es simplemente la que surge de evaluar distintos modelos y comparar los resultados que ellos arrojen. A continuación se expondrán los lineamientos básicos para llevar a cabo las comparaciones mencionadas variando el grado de formalismo. Es importante aclarar que la etapa de validación del modelo (etapa final) se hace más efectiva si es que previamente se han descartado o filtrado los modelos de menor calidad.

Comparación de modelos utilizando sets de datos frescos (Validación Cruzada)

No habría que sorprenderse si es que a la hora de evaluar el modelo, este se ajusta bien al set de datos a partir del cual fue estimado. El test real sería verificar si es que el modelo también es capaz de describir datos frescos del proceso. Una sugestiva y a la vez atractiva forma de comparar dos modelos diferentes, obtenidos a partir de dos estructuras diferentes, es evaluar su desempeño cuando se emplean sets de datos que no han sido utilizados para ajustar ninguno de los dos modelos. De esta forma podríamos optar por el modelo que presente el mejor comportamiento o performance. A dicho procedimiento se lo conoce como *Validación Cruzada* y posee muchas variantes.

Una medición de la performance, podría ser por ejemplo la suma de los cuadrados de los errores de predicción o de la diferencia entre las salidas actuales y las salidas simuladas $y_M = \hat{G}_N(q).u(t)$. Una característica atractiva del proceso de Validación Cruzada es que la comparación tiene sentido sin tener que introducir ningún argumento probabilístico y sin hacer conjeturas acerca del sistema real. La única desventaja es que tenemos que rescatar un set de datos frescos para las comparaciones, o sea que no podemos hacer uso de toda la información que disponemos para construir el modelo.

Comparación de modelos con sets de datos de "Segunda mano"

Solo si no podemos separar los datos frescos para hacer una Validación Cruzada, debemos sacar conclusiones entre dos modelos usando el mismo set de datos que ha sido utilizado para ajustar los mismos. Este hecho complica la comparación considerablemente debido a que el modelo obtenido a partir de la estructura más grande, automáticamente va a entregar el menor resultado para un criterio de ajuste convencional. Mientras la estructura modelo se expande, este resultado va a ser aún más pequeño, ya que el valor de V_N ($V_N = \min\{V_N(\theta)\}$) disminuye a medida que el modelo incorpora más características relevantes del set de datos (Figura 5.1). Incluso después de hallarse una estructura que permita una correcta descripción del sistema, el valor V_N continúa descendiendo, pero ahora a causa de parámetros adicionales (innecesarios) que se ajustan a las distintas características de realizaciones particulares del ruido. Esto es conocido como "overfit", y el valor de este ajuste extra no lo podemos conocer debido a que aplicaremos el modelo a diferentes realizaciones del ruido. Sería razonable entonces que la disminución de V_N por efecto del overfit sea menos significativa que el descenso producido por la inclusión de más características relevantes al modelo. Es por eso que buscaremos el punto de quiebre o codo en la curva de la Figura (5.1). Ciertamente es una buena idea graficar dicha curva para sacar conclusiones y verificar si el ajuste fue valedero.

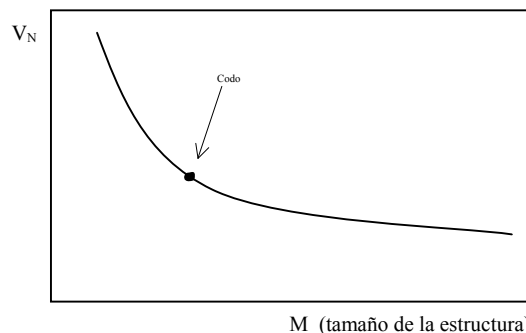


Figura 5.1 Valor mínimo de $V_N(\theta)$ en función del tamaño de la estructura modelo.

Se han desarrollado varios procedimientos con la intención de formalizar la búsqueda de dicho codo, y en la actualidad existen dos ideas básicas. La primera aproximación consiste en imitar la idea de la Validación Cruzada, esto es, usar el valor de V_N como base para el ajuste, pero aplicando al modelo el mismo set de datos a partir del cual fue calculado.

El otro procedimiento está basado en una filosofía diferente, donde se asume que la estructura más pequeña consigue realizar una correcta descripción del sistema, y luego se calcula como será la disminución esperada debido al overfit en una estructura más grande. Finalmente se evalúa si la disminución observada es significativamente mayor a la teórica o no. Esta idea se basa en el test de hipótesis estadístico.

Por último, es importante destacar que ambas aproximaciones, desde un punto de vista operacional, están estrechamente relacionadas.

5.5 Validación del modelo

El proceso de estimación de parámetros opta por el “mejor” modelo dentro de una estructura preseleccionada. Pero lo importante en todo esto es saber si dicho modelo es lo suficientemente bueno para nuestros propósitos. Este es el problema de la validación de un modelo y el asunto tiene varios aspectos a considerar como así también tres preguntas básicas que contestar:

- 1) Se adapta lo suficientemente bien el modelo a los datos observados ?
- 2) Es el modelo lo suficientemente adecuado para nuestro propósito ?
- 3) Describe el modelo al sistema real ?

Por lo general, la forma de contestar estas preguntas es examinar el modelo con la mayor cantidad posible de información acerca del sistema real. Esto incluye conocimiento previo, análisis de los datos y experiencia en el uso del modelo. En una aplicación de identificación el mayor problema a la hora de enfrentarse con el modelo lo presentan los datos en si mismos. Debido a esto, las técnicas de validación del modelo tienden a centrarse en responder las primeras dos preguntas.

Validación del modelo con respecto al propósito del modelado

Dado que es imposible dar respuesta a la pregunta 3), lo que importa en la práctica es echar luz a la segunda pregunta. Y en este aspecto, podemos afirmar que siempre existe un propósito particular que nos conduce a hacer el modelado, ya sea el diseño de un compensador, la predicción de futuras salidas, o bien la simulación del sistema. Es por eso que la última validación consiste en verificar si el problema que motivó el modelado puede ser resuelto usando el modelo obtenido. Por ejemplo, si un compensador diseñado en base al modelo hallado exhibe una satisfactoria respuesta desde el punto de vista del control, entonces el modelo será “válido”, independientemente de los aspectos formales que esto implique. A menudo será imposible, costoso, o peligroso evaluar todas los posibles modelos respecto de su propósito planeado. En lugar de eso, uno debe tomar otros caminos para formar un criterio que le permita discernir entre varios modelos.

Viabilidad de los Parámetros Físicos

Para una estructura que está parametrizada en términos de parámetros físicos, una validación natural y a la vez efectiva es comparar los valores estimados y sus varianzas, con valores razonables obtenidos en base a un conocimiento previo del sistema. Es una buena práctica el evaluar la sensibilidad que presenta el comportamiento entrada-salida, con respecto a estos parámetros para chequear su correcta identificabilidad (esto debería reflejarse en las varianzas estimadas).

Consistencia del Modelo Entrada-Salida

En un modelo de caja negra, nosotros centramos nuestro interés en sus propiedades de entrada-salida. Particularmente, para modelos lineales se suelen presentar dichas características a través de diagramas de Bode. Es por eso que siempre es aconsejable evaluar y comparar diferentes modelos lineales por intermedio de estas gráficas, incluso utilizando la varianza estimada de los parámetros, trasladada a intervalos de confianza de \hat{G} y \hat{H} . Comparaciones en diagramas de Bode, entre estimaciones por análisis espectral y otras derivadas de modelos paramétricos, son por lo general muy útiles ya que se forman a partir de suposiciones diferentes.

Generalmente, cuando el sistema real no se corresponde con el set de modelos, se obtiene una aproximación cuyas características dependen de las condiciones del experimento, los filtros usados y de la estructura seleccionada. Por lo tanto, este tipo la comparación en diagramas de Bode permite establecer si es que las características esenciales del sistema dinámico han sido capturadas correctamente.

Reducción del Modelo

Un procedimiento que evalúa si el modelo es una descripción simple y apropiada del sistema, consiste en aplicar alguna técnica de reducción en el modelo. Esto es, si el orden del modelo puede ser reducido sin afectar notablemente las propiedades de entrada-salida, entonces se dice que el modelo original era “innecesariamente complejo”. En esta oportunidad no daremos mayores detalles acerca estas técnicas, pero sí podemos comentar que el autor Söderström (1975) ha desarrollado varias ideas útiles acerca de este tema, especialmente para cancelaciones de polos y ceros.

Simulación

Una rutina comúnmente utilizada que puede ser tenida en cuenta como test de validación del modelo, es simular el sistema con la entrada actual y comparar la salida medida con la simulada. Formalmente, para un modelo general, esto significa que la salida simulada $y_M(t)$ es generada de acuerdo a:

$$y_M(t) = g(t, Z_M^{t-1}; \hat{\theta}_M) \quad (5.2)$$

$$Z_M^{t-1} = (y_M(t-1), u(t-1), y_M(t-2), u(t-2), \dots, y_M(1), u(1)) \quad (5.3)$$

Nota: no hay que confundir $y_M(t)$ con $\hat{y}(t, \hat{\theta}_N)$, ya que solo coinciden para el modelo de error de salida.

El modelo de parámetros θ_N puede entonces ser evaluado por una comparación entre $y_M(t)$ e $y(t)$, ya sea por una comparación visual en un gráfico o bien por alguna medida de distancia un poco más formal. Para la comparación se debe usar, preferiblemente, un set de datos diferente al usado para la estimación de θ_N (Validación Cruzada).

5.6 Conclusiones finales

El “sistema real” es una entidad que no puede ser dimensionada totalmente a través de un modelado práctico. En consecuencia, tenemos que conformarnos con descripciones parciales que sean adecuadas para nuestros propósitos. A veces esto significará trabajar con varios modelos del mismo sistema que serán usados para diferentes puntos de operación, para diferentes intervalos de muestreo (escalas de tiempo), y así siguiendo.

Es importante destacar el principio de “tratar con cosas simples en un primer momento”, lo que sugiere que uno debería comenzar evaluando regresiones lineales simples, tal como los modelos ARX en el caso lineal, y variantes con transformaciones no lineales de datos basadas en consideraciones físicas, cuando sea necesario.

En cuanto al tema de la validación de modelos se debería apuntar lo siguiente:

- Comparación en diagramas de Bode, de modelos lineales obtenidos bajo varias condiciones y utilizando diversas estructuras (incluyendo estimaciones por análisis espectral).
- Comparación de salidas medidas y simuladas.
- Monitoreo de intervalos de confianza de los parámetros estimados, para rastrear y localizar los ceros y polos en la función de transferencia. Como así también posibles pérdidas locales de identificabilidad.

Finalmente, es necesario tener en cuenta que dichas técnicas deberían ser tomadas como guías de trabajo, y no dejar que la selección de variables se haga exclusivamente por procedimientos estadísticos, desplazando de este modo al imprescindible ingrediente subjetivo que pueda aportar quien realice el análisis.

5.7 Aplicación en Matlab

Luego de haber analizado los datos, es probable que hayamos obtenido una extensa colección de modelos de diferentes órdenes y estructuras. Entonces, llegado este punto es aconsejable estar provisto de una gran variedad de herramientas que nos permitan evaluar las cualidades de los diferentes modelos. En este sentido podemos asegurar que el Toolbox de Identificación de Matlab es un poderoso instrumento capaz de llevar a cabo dichas tareas.

Comparación de Diferentes Estructuras

Es natural comparar los resultados obtenidos a partir de diferentes modelos de una misma estructura, esto es, variando los órdenes y retardos de la misma. Por ejemplo, para modelos del tipo ARX esto puede ser llevado a cabo eficientemente por el comando 'arxstruc'. Para ello, primero debemos definir una matriz NN con todas la estructuras ARX que deseamos investigar:

$$NN = \begin{pmatrix} a_{a1} & a_{b1} & a_{k1} \\ \vdots & \vdots & \vdots \\ a_{am} & a_{bm} & a_{km} \end{pmatrix}$$

luego,

```
>> V = arxstruc (ze,zv,NN);
```

Siendo el modelo ajustado al set de datos z_e , para cada conjunto de parámetros de NN. Posteriormente, a cada uno de los modelos se le aplica el set de datos z_e y se calcula la suma de los errores de predicción al cuadrado. A estos valores se los conoce como funciones de pérdida y son guardados en V con los modelos correspondientes.

Para seleccionar el modelo que posea la función de pérdida más pequeña para el set de validación z_e , se procede utilizando:

```
>> nn = selstruc (V,0);
```

Este no es ni más ni menos que el procedimiento de *Validación Cruzada* mencionado con anterioridad.

Por lo general, es también una buena idea inspeccionar visualmente como cambia el ajuste con el número de parámetros estimados. Un gráfico del ajuste versus el número de parámetros se obtiene con:

```
>> nn = selstruc (V);
```

Esta rutina lo incita a uno a elegir el número de parámetros a estimar, basándose en primer lugar, en la inspección visual del gráfico para luego escoger el modelo que mejor se ajuste a dicha cantidad de parámetros.

En estos casos, el comando 'struc' suele convertirse en una gran ayuda, ya que genera matrices NN con diferentes estructuras características. Una secuencia de comando típica sería:

```
>> V = arxstruc (ze,zv,struc(2,2,1:10));
>> nn = selstruc (V,0);
>> nk = nn(3);
>> V = arxstruc (ze,zv,struc(1:5,1:5,nk-1:nk+1));
>> nn = selstruc (V);
```

De esta forma hemos establecido un valor adecuado para el retardo nk, testeando modelos de segundo orden con retardos entre uno y diez. El mejor ajuste selecciona un retardo, y luego se evalúan todas las combinaciones posibles entre modelos ARX con un número de parámetros "a" y "b" entre cero y cinco, utilizando retardos en un entorno cercano al seleccionado previamente.

Si el modelo es validado sobre el mismo set de datos a partir del cual fue estimado, esto es $z_e=z_v$, entonces el ajuste siempre mejorará a medida que se incrementa el tamaño de la estructura. Por lo tanto es necesario compensar este decrecimiento automático de las funciones de pérdida debido al overfit (Figura 5.1), y para ello se han desarrollado diferentes procedimientos. Probablemente la mejor técnica conocida es la Predicción Final de Error de Akaike (FPE), que está estrechamente relacionada con el Criterio de la Teoría de la Información. Ambas simulan la validación cruzada cuando el modelo se evalúa con un set de datos frescos.

La FPE es formada así:

$$FPE = \frac{1 + \frac{n}{N}}{1 - \frac{n}{N}} \cdot V$$

donde n es el número total de parámetros estimados, N es el largo del registro de datos y V la función de pérdida (ajuste cuadrático) para la estructura en cuestión.

El AIC se forma así:

$$\text{AIC} = \log[(1 + 2n/N) \cdot V]$$

De acuerdo a la teoría de Akaike se debe seleccionar, dentro de una colección de diferentes modelos, aquel que posea la menor FPE (o AIC). La estructura que minimiza el AIC se obtiene con:

```
>> nn = selestruc (V,'AIC');
```

donde V fue previamente generada por 'arxstruc'.

Un criterio relacionado es la Aproximación de Mínima Distancia de Rissanen, el cual selecciona la estructura que permita hacer la descripción general más pequeña a partir de los datos observados. Esto se puede obtener con:

```
>> nn = selstruc (V,'MDL');
```

Chequeo de cancelaciones Polo-Cero

Una cancelación polo-cero en el modelo dinámico es una indicación de que el orden escogido puede ser muy elevado. Para juzgar si una cancelación es real, hay que tomar en consideración las incertidumbres que existen en la localización de polos y ceros. En particular la función 'th2zp' calcula las regiones de confianza para los polos y ceros, lo cual puede ser graficado a través de:

```
>> zpplot (th2zp(th),1);
```

donde 1 indica la cantidad de desviaciones standard de ancho que tendrá el intervalo de confianza. Entonces, si las regiones de confianza se superponen, habría que intentar utilizar órdenes menores para el modelo.

Este chequeo es de especial utilidad cuando los modelos han sido generados por el comando 'arx', ya que los órdenes pueden estar sobredimensionados debido a que $1/A(q)$ describe el comportamiento frente al ruido.

Análisis Residual

Los residuos asociados con los datos y el modelo dado (diferencia entre y y y_M), son independientes de la entrada y tienen, idealmente, características de ruido blanco. En particular, la siguiente función:

```
>> e = resid (z,th);
```

calcula los residuos e y lleva a cabo los análisis de independencia y "blancura". Se presenta entonces en pantalla la función de autocorrelación de e y la función de intercorrelación entre u y e hasta el retardo número 25. También son exhibidos los intervalos de confianza de estas funciones, asumiendo que e es blanco e independiente de u .

La regla dice que si la función de autocorrelación se escapa significativamente de este intervalo de confianza, entonces no se acepta el correspondiente modelo como una buena descripción del sistema. Son necesarias algunas aclaraciones para dar por válida esta afirmación:

1. Una estructura como la de Error de Salida se enfoca en la dinámica de G y deja un poco de lado las propiedades de H . De modo que si estamos interesados en G habría que centrarse en buscar la independencia entre e y u en lugar de la blancura de e .
2. La intercorrelación entre e y u para retardos negativos, o el hecho de que el ruido $e(t)$ actual afecte la futura entrada $u(t)$, es una indicación de que existe realimentación de la salida hacia la entrada. Pero esta no es una razón para descartar el modelo, ya que la correlación para valores de retardo negativos es de gran interés debido a que, a pesar de que existen ciertos métodos que no funcionan bien cuando se presentan realimentaciones en los datos de

entrada-salida, a partir de estos se logran realizar buenos gráficos de correlación cruzada para propósitos de validación.

3. Cuando usamos modelos ARX, el procedimiento de mínimos cuadrados hace automáticamente cero la correlación entre $e(t)$ y $u(t-k)$ para $k = nk, nk + 1, \dots, nk + nb - 1$.

Como parte del proceso de validación se suelen también graficar los residuos para poder hacer una simple inspección visual en búsqueda de posibles irregularidades y outliers. Esto se hace usando:

```
>> plot (e);
```

Nota: Outliers son observaciones infrecuentes o puntos que no parecen seguir la distribución característica del resto de los datos.

Simulaciones libres de ruido

Para saber si el modelo escogido es capaz de reproducir los datos que se observan, a partir de excitar al modelo con una cierta entrada $u(t)$, se puede correr una simulación como la siguiente:

```
>> yh = idsim (u,th);
>> plot ([y yh]);
```

El mismo resultado se obtiene con:

```
>> compare ([y,u],th);
```

Un test más exigente y relevante, sería llevar a cabo esta simulación (e incluso su análisis residual) a partir de un set de datos frescos que no haya sido utilizado para estimar el modelo (concepto de validación cruzada).

Evaluación de la incertidumbre del modelo

El modelo estimado es siempre incierto debido a las perturbaciones existentes en los datos observados (incertidumbre externa), y a la falta de una estructura absolutamente correcta (incertidumbre interna). La variabilidad del modelo que se produce debido a perturbaciones aleatorias en la salida, puede ser calculada por la mayoría de los procedimientos de estimación conocidos, e incluso puede ser presentada en un gran número de formas. Dicha variación responde la pregunta de cuan diferente puede llegar a ser un modelo si se repite el proceso de identificación, usando la misma estructura, pero con un set de datos diferente que contiene la misma secuencia de entrada. Sin embargo, este análisis no tiene en cuenta los errores sistemáticos que se producen debido a una inadecuada elección de la estructura. Por lo tanto no hay garantías de que el "sistema real" se sitúe dentro del intervalo de confianza.

Para hallar la incertidumbre en la respuesta frecuencial se comienza por pasar el modelo del formato *theta* al formato *freqfunc*, que es un formato que contiene información acerca de valores de frecuencia, amplitudes, como así también sus desviaciones estándar :

```
>> [g,fiv] = th2ff (th);
```

luego se pueden graficar los datos obtenidos usando el comando 'bodeplot', no sin antes especificar la cantidad de desviaciones estándar que se desean graficar, por ejemplo:

```
>> bodeplot (g,2);
```

Por otro lado, la incertidumbre en la respuesta temporal se puede visualizar utilizando:

```
>> idsimsd (u,th);
```

En este caso se grafica en un mismo diagrama, la respuesta a la entrada u , de diez posibles modelos obtenidos a partir de la distribución del modelo th .

La ambigüedad de estas respuestas tiene que ver principalmente con las propiedades externas del modelo, ya que refleja los efectos de una inadecuada excitación y la presencia de perturbaciones.

En cambio, la incertidumbre interna se manifiesta por lo general en la matriz de covarianza de los parámetros estimados. Esto se puede revelar con,

```
>> present (th);
```

Nota: Debido a la alta precisión numérica que implementa Matlab, es altamente improbable que encontremos mensajes acerca de matrices singulares o deficientes. En lugar de eso, los modelos de órdenes sobreestimados normalmente muestran grandes varianzas de sus parámetros. Consecuentemente, este suele ser un buen indicador de que uno ha ido demasiado lejos en el incremento del orden del modelo.

Otra forma de interiorizarse acerca de la incertidumbre interna que presenta el modelo, puede ser observando las desviaciones estándar existentes en las localizaciones de polos y ceros. Para hacer esto se suele comenzar por pasar el modelo al formato *zepo* (cero polo) :

```
>> ceropolo = th2zp (th);
```

luego es posible graficar el mapa polo-cero usando el comando *zpplot* de la misma forma que se usa el comando *bodeplot*. Grandes incertidumbres en las representaciones son causadas por modelos de órdenes excesivamente altos, excitación inadecuada, o malas relaciones señal a ruido en las ristas de datos.

Comparación de Diferentes Modelos

Suele ser por lo general una muy buena idea presentar las propiedades del modelo en términos de variables que tengan un mayor significado físico que los parámetros en si mismos.

Por otro lado, diagramas de Bode, mapas polo-cero y simulaciones del modelo, nos dan una idea de las propiedades del sistema que deben ser capturadas por nuestro modelo.

Si ahora varios modelos de diferentes características proporcionan diagramas de Bode similares en el mismo rango de frecuencias de interés, entonces podemos asegurar en justa medida que esto estará reflejando particularidades auténticas del sistema real. Así entonces podemos elegir el modelo más simple de entre todos estos.

Una típica sesión de identificación incluye la estimación en diferentes estructuras, y comparaciones de las propiedades de los modelos.

6. ACONDICIONAMIENTO DE LOS DATOS

Una vez que los datos han sido recolectados del proceso físico, es probable que los mismos no estén en condiciones de ser usados por algoritmos de identificación. Esto se debe a que hay varias deficiencias posibles en los datos que deberían ser tenidas en cuenta, tales como:

1. Perturbaciones de alta frecuencia en la adquisición de los datos, por encima de las frecuencias de interés del sistema dinámico.
2. Bursts y/o outliers ocasionales.
3. Drift, offset y perturbaciones de baja frecuencia, posiblemente de carácter periódico.

Aclaraciones:

- *Burst*: es una secuencia o ráfaga de datos de valor unitario.
- *Outliers*: son observaciones infrecuentes o puntos que no parecen seguir la distribución característica del resto de los datos.
- *Drift*: variación lenta de una característica del sistema como la amplitud, frecuencia o potencia de salida. Esto se puede deber, por ejemplo, a cambios de temperatura o envejecimiento.

Es necesario acentuar el hecho de que en aplicaciones off-line, siempre es aconsejable graficar los datos, en orden de realizar una inspección minuciosa de las deficiencias mencionadas.

6.1 Perturbaciones de alta frecuencia

Este tipo de perturbaciones indican que la elección del intervalo de muestreo y de los filtros de pre-muestreo, no fueron correctamente escogidos. Si luego de hacer el experimento, resulta ser que el periodo de muestreo fue tomado innecesariamente pequeño, uno siempre podrá re-muestrear los datos tomando de a cada n muestras del registro original (proceso de diezmado). Sin embargo, sigue siendo necesario aplicar un filtro anti-aliasing antes de realizar el re-muestreo de los datos.

6.2 Burts y Outliers

Errores aislados o valores muy perturbados pueden tener una influencia considerable en la estimación resultante. Es por eso que resulta necesario proteger la estimación de los malos datos, y complementar esto con un robusto criterio de identificación. Cabe resaltar que los datos erróneos son muy fáciles de detectar en un gráfico residual, e igualmente sencillo cambiarlos manualmente por un valor proveniente de una interpolación.

6.3 Perturbaciones Lentas: Comentarios Generales

Perturbaciones de baja frecuencia, offsets, trends (valores medios), drift, y variaciones periódicas no son comunes en los sets de datos. Estos fenómenos provienen típicamente de fuentes externas que podríamos o no incluir en nuestro modelo. Y en este sentido podríamos decir que existen básicamente dos aproximaciones diferentes para lidiar con estos problemas:

1. Remover las perturbaciones a través de un pretratamiento explícito de los datos.
2. Dejar que el modelo del ruido se ocupe de las perturbaciones.

La primera aproximación involucra el hecho de remover trends y offsets por sustracción directa, mientras que la segunda confía en los modelos de ruido con polos sobre o cerca del círculo unitario, como la estructura ARIMA.

En principio es posible ilustrar ambas aproximaciones aplicadas al problema del offset, y el modelo lineal estándar que utilizamos será el siguiente:

$$A(q).y(t) = B(q).u(t) + v(t) \quad (6.1)$$

el cual nos informa tanto acerca de las propiedades dinámicas del sistema como así también sobre la relación estática entre una constante $u(t) = \bar{u}$ y el valor de estado estacionario $y(t) = \bar{y}$.

$$A(1).\bar{y} = B(1).\bar{u} \quad (6.2)$$

En la práctica, las medidas de entrada y salida tomadas en bruto, llamadas $u^m(t)$ e $y^m(t)$, son recolectadas y guardadas en unidades físicas, pudiendo ser los niveles de las mismas bastante arbitrarios. O sea que la ecuación (6.1), que escribe las propiedades dinámicas del sistema, puede tener muy poca relación con la ecuación (6.2) que relaciona los niveles de las señales a frecuencia cero. En otras palabras, (6.2) es realmente una restricción innecesaria para (6.1), y existen al menos tres formas posibles de enfrentar el problema:

1. Permitir que $y(t)$ y $u(t)$ sean desviaciones de un equilibrio físico: La aproximación más natural sería determinar el nivel de salida \bar{y} que corresponde a la entrada de un valor constante $u^m(t) \equiv \bar{u}$ cercano al punto de operación deseado. Entonces definimos:

$$y(t) = y^m(t) - \bar{y} \quad (6.3a)$$

$$u(t) = u^m(t) - \bar{u} \quad (6.3b)$$

como las desviaciones de este equilibrio. Estas variables corridas van a satisfacer automáticamente la ecuación (6.2), haciendo ambos miembros iguales a cero. De esta forma (6.2) no influenciará el ajuste que será realizado en (6.1). Este tipo de aproximación enfatiza la interpretación física de la ecuación (6.1) como una linealización alrededor del punto de equilibrio.

2. Extraer los valores principales o valores medios de las muestras: Una segunda aproximación es definir:

$$\bar{y} = \frac{1}{N} \sum_{t=1}^N y^m(t) \quad ; \quad \bar{u} = \frac{1}{N} \sum_{t=1}^N u^m(t) \quad (6.4)$$

y luego utilizar las ecuaciones (6.3). De modo que si una entrada $u^m(t)$, que varía alrededor de \bar{u} , conduce a un valor de salida que oscila alrededor de \bar{y} , entonces es probable que el par (\bar{u}, \bar{y}) esté cerca de corresponder al punto de equilibrio del sistema.

3. Estimar el offset en forma explícita: Uno podría también modelar el sistema usando variables en las unidades físicas originales, y luego agregar un valor constante que haga las veces de offset:

$$A(q).y^m(t) = B(q).u^m(t) + \alpha + v(t) \quad (6.5)$$

Si comparamos la ecuación (6.1) con (6.3), vemos que α corresponde a $A(1).\bar{y} - B(1).\bar{u}$.

De esta forma, el valor de α es incluido en el vector de parámetros θ y estimado a partir del set de datos. Es evidente que esta aproximación, es de hecho una variante de la segunda aproximación expuesta.

6.4 Evaluación de las Aproximaciones

En una aplicación off-line con offsets, la aproximación más recomendable es la primera, y si un experimento de estado estacionario resulta de difícil aplicación, entonces se suele utilizar la segunda aproximación. En cambio, la estimación explícita del offset es una forma innecesariamente complicada de extraer el valor medio de una muestra.

Nota:

Es de especial importancia remover los offsets cuando los métodos utilizados emplean un modelo de ruido fijo, tal como los modelos de error de salida.

6.5 Drift, Trends y Variaciones de Baja Frecuencia

Los métodos para tratar con otras perturbaciones lentas en los datos son bastante análogos a las aproximaciones discutidas con anterioridad. Esto se debe a que los *Drifts* y *Trends* pueden ser vistos como variaciones en el tiempo del punto de equilibrio. De esta forma, es posible ajustar líneas rectas o segmentos curvos, de la misma forma en que se consideran niveles constantes de offset y sus variaciones respecto a la media.

6.6 Aplicación en Matlab

La extracción de información de los datos no suele ser por lo general una tarea directa. Y las decisiones requeridas para la selección y validación del modelo, necesitan estar acompañadas por un cuidadoso tratamiento de los datos. Es por eso, que a continuación intentaremos introducir algunos conceptos básicos acerca de cómo lidiar con ciertas situaciones comunes presentes en la mayoría de los casos.

Niveles de Offset

Como comentamos previamente, los datos que han sido recolectados de una planta física real, están típicamente medidos en unidades físicas. Y por lo general los valores de entrada-salida en bruto no se suelen corresponder en forma consistente, lo cual forzaría a los modelos a gastar inevitablemente algunos parámetros para corregir dichos niveles.

En particular es bastante sencillo tratar con niveles de offset no deseados, ya que siempre intentaremos extraer el valor medio de la secuencia de entrada y de salida antes de realizar la estimación. El caso ideal sería si dichos valores medios corresponden a un equilibrio físico, pero si tales valores son desconocidos, entonces es aconsejable utilizar un comando para extraer el valor medio de cada muestra:

```
>> z = dtrend(z);
```

Nota: con este comando también es posible remover valores medios o medias lineales, ya sea en forma total o parcial (por tramos).

Outliers

Incluso los datos reales están sujetos a posibles perturbaciones. Esto es, largas perturbaciones poco habituales, fallas temporales del sensor o transmisor, etc. Es importante no permitir que estos *outliers* afecten en forma considerable a los modelos estimados.

En general, la robustificación del criterio de error suele ser de gran ayuda para estos casos. Sin embargo, siempre es buena práctica examinar los residuos para localizar muestras poco usuales de gran valor, para luego volver a los datos originales y poder evaluarlos en una forma más crítica. Obviamente que si los datos poseían algún tipo de anomalía, estos pueden ser suavizados y posteriormente repetido el procedimiento de estimación.

Filtrado de los Datos

Dependiendo del tipo de aplicación, es probable que el interés sobre el modelo esté focalizado en algunas bandas de frecuencia específicas. El filtrado previo de los datos, a través de filtros que resalten la información sobre dichas bandas, mejora el ajuste en las regiones de interés. En particular, el Toolbox de Procesamiento de Señales de Matlab contiene varias funciones para el diseño y aplicación de filtros digitales. Por ejemplo la función 'idfilt' (del Toolbox de Identificación) utiliza la función butter del Toolbox de Procesamiento de Señales, e implementa un filtro de Butterworth.

```
Ej:  
>> zf = idfilt(z,5, [10 100]);
```

Aquí hemos usado un filtro Butterworth de quinto orden tipo pasabanda, con frecuencia de corte inferior en $10 \times \pi$ Hz. y frecuencia de corte superior en $100 \times \pi$ Hz.

Otro comando bastante útil que posee el Toolbox de Procesamiento de Señales, es uno que realiza decimación en un factor variable para poder disminuir la frecuencia de muestreo, y observar los datos temporalmente más espaciados.

Ej:

```
>> y4 = decimate (y,4);  
>> u4 = decimate (u,4);
```

Aquí se ha reducido la frecuencia de muestreo en un factor 4.

Realimentación en los Datos

Si al momento de recolectar los datos el sistema estaba operando a lazo cerrado, entonces habrá que tomar ciertos recaudos a la hora de procesar la información.

Básicamente todos los métodos de predicción de error trabajan en forma correcta con datos de lazo cerrado. Cabe destacar, sin embargo, que el modelo de Error de Salida y el modelo Box-Jenkins que normalmente entregan una correcta descripción de la dinámica de G , incluso cuando H (que es igual a 1 en el modelo de error de salida) no permita una correcta descripción de las propiedades de ruido que presenta el sistema, no conserva estas cualidades cuando el sistema está funcionando a lazo cerrado. En estos casos es necesario modelar H en una forma más cuidadosa.

También el método de análisis espectral entrega resultados poco confiables cuando es aplicado en un conjunto de datos que proviene de un lazo cerrado.

En estos casos lo más aconsejable para detectar la presencia de realimentación, es utilizar el método de Análisis por Correlación y estimar la respuesta impulsiva del sistema. La existencia de valores significativos en la respuesta impulsional para retardos negativos es un claro indicio de la presencia de realimentación en el sistema.

Otra alternativa es, una vez que se ha estimado un modelo paramétrico, utilizar el comando *resid* para visualizar un gráfico de la correlación entre los residuos y la entrada. Si hallamos que la correlación es significativa para valores de retardo negativos, entonces es probable que exista algún proceso de realimentación en la generación de la entrada.

Dicho esto, queda de manifiesto que la evaluación en búsqueda de realimentaciones es parte natural del proceso de validación del modelo.

Retardos

La selección del retardo n_k en la estructura modelo, es considerada un paso muy importante en la obtención de buenos resultados de identificación. Sin ir más lejos, uno puede obtener una certera idea acerca de los retardos que presenta el sistema, a través del estudio de la respuesta impulsiva obtenida con el comando *cra*.

Es notable como la estimación incorrecta de los retardos en los modelos paramétricos, suele traducirse en visibles cambios del modelo obtenido. En particular, retardos subestimados (n_k muy chicos) conducen a pequeños valores en las estimaciones de los parámetros b_k . Por otro lado, retardos sobrestimados (n_k muy grandes) producen una significativa correlación entre los residuos y la entrada, en especial sobre los retardos o lags correspondientes a los términos b_k faltantes.

Un buen procedimiento consiste en empezar utilizando la sentencia *arxstruc* para evaluar todos los posibles retardos en conjunto, en especial con un modelo de segundo orden. Luego, habría que utilizar el retardo que entregue el mejor ajuste. Una vez encontrado un modelo apropiado, es aconsejable variar n_k alrededor de su valor nominal y realizar los ajustes finales.

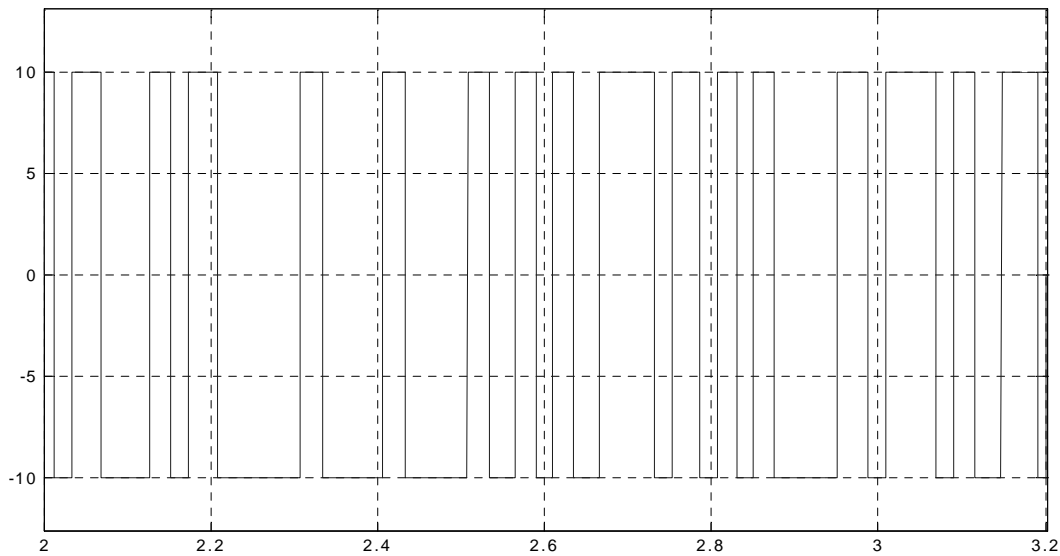
7. NOCIONES BÁSICAS PARA EL DISEÑO DE LA SEÑAL DE EXCITACIÓN

Si bien existe una gran variedad de excitaciones posibles para identificar un sistema, en esta sección nos centraremos en el particular problema que implica diseñar una señal de excitación persistente de características pseudoaleatorias PRBS (pseudorandom binary sequence).

7.1 Secuencia Binaria de Alternancia Pseudoaleatoria

Los parámetros que describen enteramente las características de este tipo de señales son los siguientes:

- Ancho del pulso más corto (T_c).
- Probabilidad de cambio (p).
- Duración del ensayo (T_t).
- La amplitud del pulso (ΔP).



1) El ancho del pulso más corto (T_c) está directamente relacionado con el modo más rápido que se desea identificar. Esto es, a fin de tener una idea estimada de los modos rápidos (no dominantes) del sistema se debe optar por diseñar una PRBS con una distribución espectral de potencia que presente considerable energía en la zona que comprende las frecuencias más altas y más bajas a identificar.

Por lo tanto, a fin de no dejar aspectos importantes del proceso sin identificar sería adecuado definir el ancho T_c de modo tal que se puedan excitar hasta las dinámicas cercanas a τ_m (constante de tiempo más rápida del sistema). Tal es así, que se toma:

$$T_c = \tau_m / 2$$

2) Una vez elegido T_c , la selección del número de elementos de la PRBS (n_s), está directamente relacionado con el tiempo total de cada ensayo ($T_t = n_s \cdot T_c$), y este a su vez está altamente condicionado por la constante de tiempo más larga del sistema (τ_M). Con el fin de tener un valor más realista de la dinámica dominante del proceso, suele ser aconsejable efectuar un ensayo preliminar que consista en aplicar un escalón a la entrada del sistema, para así tener un valor estimativo de τ_M y poder definir T_t de modo de poder excitar la planta en forma adecuada. Un criterio válido podría ser tomar $T_t \approx 3 \cdot \tau_M$.

3) Para seleccionar la probabilidad de cambio de la PBRS (p) debe observarse que para valores de p tendiendo a cero, los cambios se hacen más infrecuentes, por consiguiente los pulsos (ristras de unos o ceros consecutivos) más largos y menos alternantes, pareciéndose la señal cada vez más a un escalón. En estas circunstancias se observa que la energía tiende a concentrarse en las bajas frecuencias, perdiéndose excitación para los modos rápidos. Cuando p tiende a uno, la PBRS conmuta en forma permanente, concentrándose la energía en las altas frecuencias (cerca de $f_{\text{máx}} = 1/2T_c$). Una secuencia de este tipo aplicada a una planta de características pasabajo, no produciría ningún efecto, ya que el sistema no podría responder suficientemente rápido a las conmutaciones de la entrada.

Por lo tanto, si se pudieran efectuar ensayos sin ninguna limitación de tiempo, lo ideal sería usar una $p = 0.5$, que tiene una distribución de energía bastante pareja en todo el rango de frecuencias de interés.

4) La amplitud del pulso (ΔP) dependerá de la zona de trabajo que se desee identificar.

Nota: Si bien la probabilidad de cambio de la PBRS y la amplitud del pulso suelen ser seleccionadas en forma independiente. Cabe resaltar que en conjunto determinan el grado de apartamiento de la salida respecto de su valor de equilibrio, de modo que una incorrecta elección de alguno de estos parámetros podría repercutir en un indeseado alejamiento de la zona de trabajo.

8. CONCLUSIONES ACERCA DE LA IDENTIFICACIÓN

Que nos puede ofrecer la identificación de sistemas?

Las técnicas de identificación de sistemas forman una versátil herramienta para la solución de varios problemas de la ciencia y la ingeniería. El valor de esta herramienta ha sido evidenciado por numerosas aplicaciones en diversos campos.

Es por eso que para finalizar con este estudio vamos a hacer un breve comentario acerca de las posibilidades y limitaciones prácticas que ofrece la identificación.

Posibilidades

Las técnicas discutidas en este apunte son independientes de la aplicación, ya que cualquier proceso de características dinámicas representa una aplicación potencial.

Ciertamente, se han registrado exitosas aplicaciones de identificación de sistemas en muy diferentes campos del conocimiento, tales como procesos de control, biomedicina, sismología, sistemas ambientales, dinámicas de aviones, macroeconomía, procesamiento de señales y muchas más.

Siendo que la clave para arribar a buenos resultados reside en el hecho de obtener datos de buena calidad y tener un buen criterio acerca de las características de la estructura modelo que debería ser utilizada.

Limitaciones: Calidad de los Datos

A esta altura resulta obvio que la limitación en el uso de las técnicas de identificación está estrechamente relacionada con la disponibilidad de un buen set de datos y de eficientes estructuras modelo. Ya que, sin un razonable registro de datos no se puede hacer mucho, y existen varias razones por las cuales no se puede obtener, en ciertas aplicaciones, registros de este tipo. Un ejemplo clásico es cuando la escala de tiempo del proceso es tan lenta que cualquier registro de datos que tome termina siendo corto. Este es el caso de los sistemas ecológicos y económicos que claramente sufren de este tipo de problemas. Otra razón es que la entrada no debe estar abierta para manipulaciones, sino se podría empeorar la relación señal a ruido y no se podría garantizar la total identificabilidad del sistema.

Es importante remarcar que las malas relaciones señal a ruido pueden, en teoría, ser compensadas por registros de datos más largos. Pero, incluso si la planta admite largos tiempos de experimentación, puede ser que no siempre este sea un buen camino de salida, debido a las variaciones temporales, drift y perturbaciones de baja frecuencia que pueda sufrir el proceso.

Finalmente, aún cuando tengamos permitida la manipulación de la entrada, podemos medir por largos periodos, y tengamos buenas relaciones señal a ruido, todavía puede ser difícil obtener un buen registro de datos. La principal razón de esto suele ser la presencia de perturbaciones que no se ajusten bien al modelo de "proceso estocástico estacionario". En la sección de Acondicionamiento de los Datos ya hemos discutido la forma de tratar con estas perturbaciones lentas para que las mediciones sean exitosas. Por lo tanto no queda más por recordar que: la calidad de los datos debe ser de primordial atención en aplicaciones de identificación de sistemas. Lo cual también determinará el costo del experimento.

Limitaciones: Estructuras Modelo

Es bastante trivial que una mala estructura modelo no puede ofrecer un buen modelo final, independientemente de la cantidad y calidad de datos disponibles.

Entonces el primer problema al que nos enfrentamos, es saber si el proceso (alrededor de su punto de operación) admite una descripción estándar preexistente ("caja negra"), o si es necesario construir un set de modelos hecho a medida. En el primer caso, nuestras chances de tener éxito son grandes; en el segundo, tendremos que recurrir a un análisis físico del sistema antes de estimar el modelo. Este problema es claramente dependiente de la aplicación y por lo tanto no muy discutido dentro de la literatura de identificación. Sin embargo, es necesario acentuar que la clave para arribar a un buen resultado reside en lo siguiente: El razonamiento, la intuición y el entendimiento del problema no pueden volverse obsoletos por la construcción automática de un modelo.

BIBLIOGRAFÍA

- [1] Lennart Ljung, *System identification (the teory for the user)*. Pentice Hall, 1987.
- [2] Lennart Ljung, *System Identification Toolbox (user guide)*. The Math Works, 1991.
- [3] Papoulis, *Probability, Random Variables, and Stochastic Processes*. McGraw Hill, 1965.
- [4] Charles S. Williams, *Deisgning Digital Filters*. Prentice Hall, 1986.