



Identificación de Sistemas

Prof. Cesar de Prada

Dpto. Ingeniería de Sistemas y Automática

Universidad de Valladolid, España

Tlf + 34 983 423164

prada@autom.uva.es



Indice

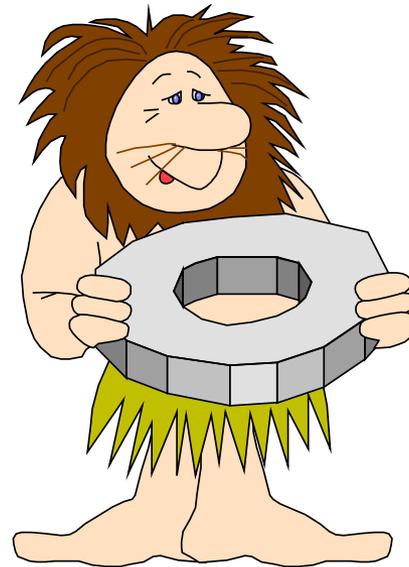
- ✓ Introducción
- ✓ Metodología
- ✓ Modelos
- ✓ Métodos de estimación
- ✓ Identificación práctica (HIDEN)
- ✓ Validación de modelos



¿Como obtener modelos?



Mediante razonamientos,
usando leyes de físicas,
químicas, etc



Mediante experimentación
y análisis de datos

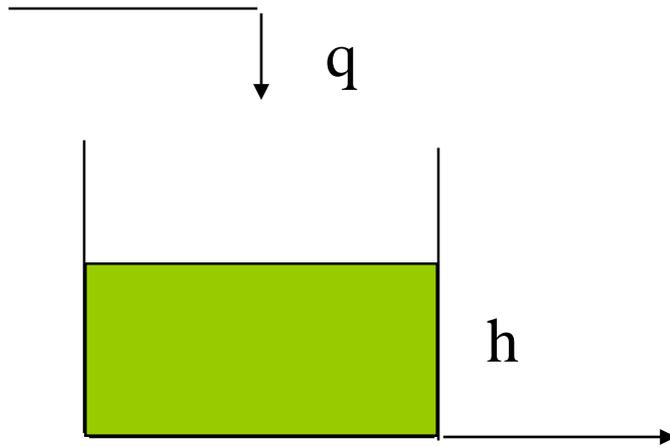


Modelos de conocimiento

- Se obtienen mediante razonamientos y la aplicación de principios de conservación de masa, energía, momento, etc y otras leyes particulares del dominio de aplicación
- Tienen validez general
- Requieren conocimiento profundo del proceso y de las leyes fisico-químicas



Modelos de conocimiento



Ejemplo:

Conservación de masa

Acumulación =
entrada - salida

$$A \rho \frac{d h}{d t} = q - k \rho \sqrt{h}$$

A sección del depósito
 ρ densidad, k constante

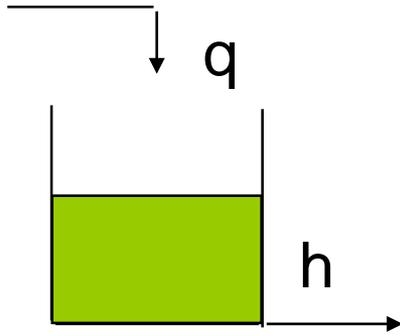


Modelos de conocimiento

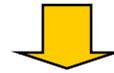
- Formados por conjuntos de ecuaciones diferenciales y algebraicas
- Útiles para muchos fines
 - ¿Qué pasa si?
 - Diseño
 - Pruebas antes de la implementación
- Difíciles de manipular matemáticamente
- Se resuelven mediante simulación
- Suelen tener algún parámetro desconocido



Linealización y discretización



$$A\rho \frac{d h}{d t} = q - k\rho\sqrt{h}$$



$$a \frac{d \Delta h}{d t} = b\Delta h + c\Delta q$$



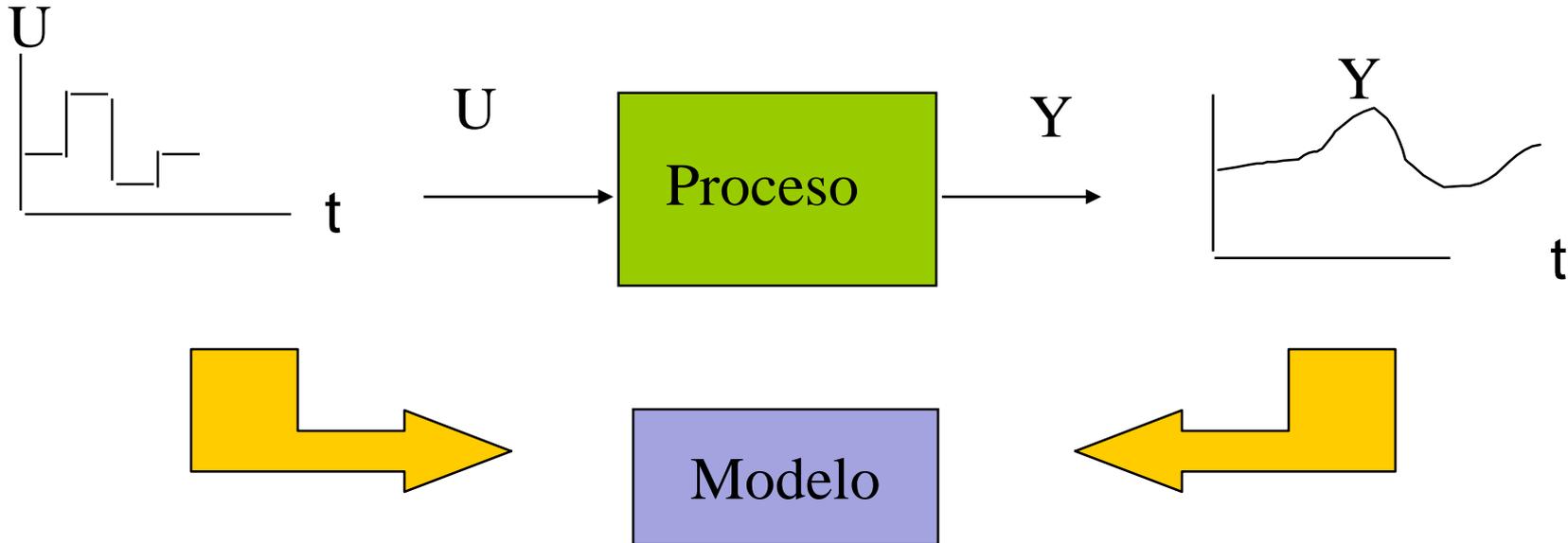
Δh y Δq son cambios de h y q sobre unos valores de equilibrio

$$\Delta h(kT) = \alpha\Delta h(kT - T) + \beta\Delta q(kT - T)$$



Identificación

El modelo se obtiene a partir de datos experimentales de entrada-salida del proceso

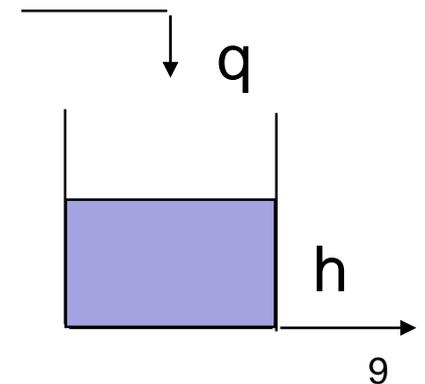




Modelos de caja negra

- Se postula una relación matemática entre la entrada y la salida que depende de unos parámetros que deben ser estimados mediante los datos experimentales, sin que dicha relación deba estar fundada directamente en leyes físico-químicas u otras.

$$h(kT) = \alpha h(kT - T) + \beta q(kT - T) + \gamma$$





Modelos grises

- Son modelos de conocimiento en los que una parte está modelada como un modelo de caja negra, la cual representa ciertos fenómenos complejos o difíciles de modelar de otra forma.
- Todos los tipos de modelos requieren de una etapa de estimación de los valores de sus parámetros en mayor o menor medida.



Metodología de la identificación

Objetivos, Proceso y diseño de experimentos

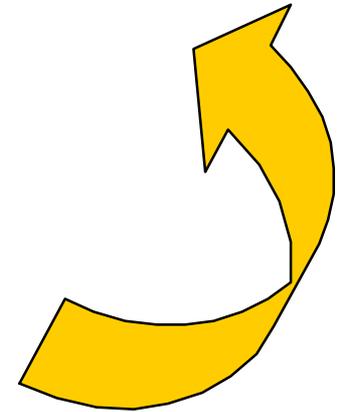
Toma de datos experimentales

Análisis y tratamiento de datos

Selección del tipo de modelo

Estimación de parámetros

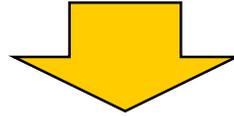
Validación del modelo





Objetivos

Teoría de estimación de parámetros

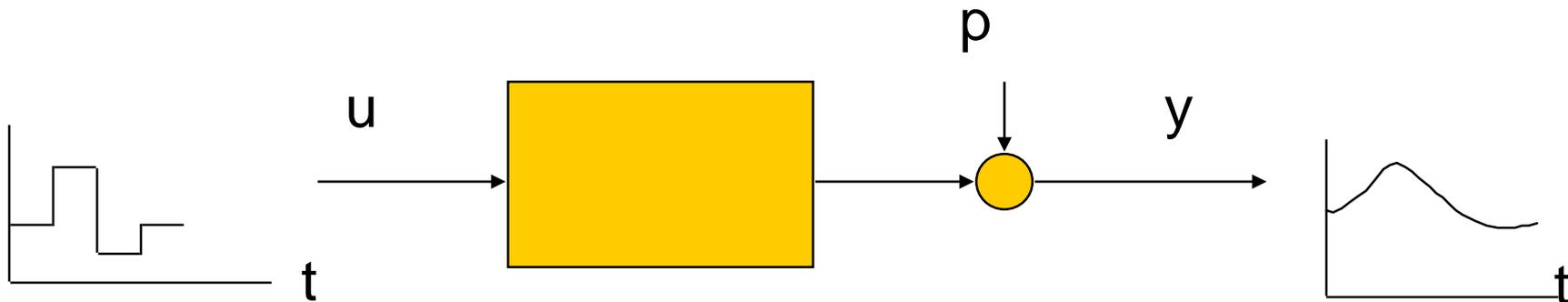


Importancia de los experimentos y
el tratamiento de datos



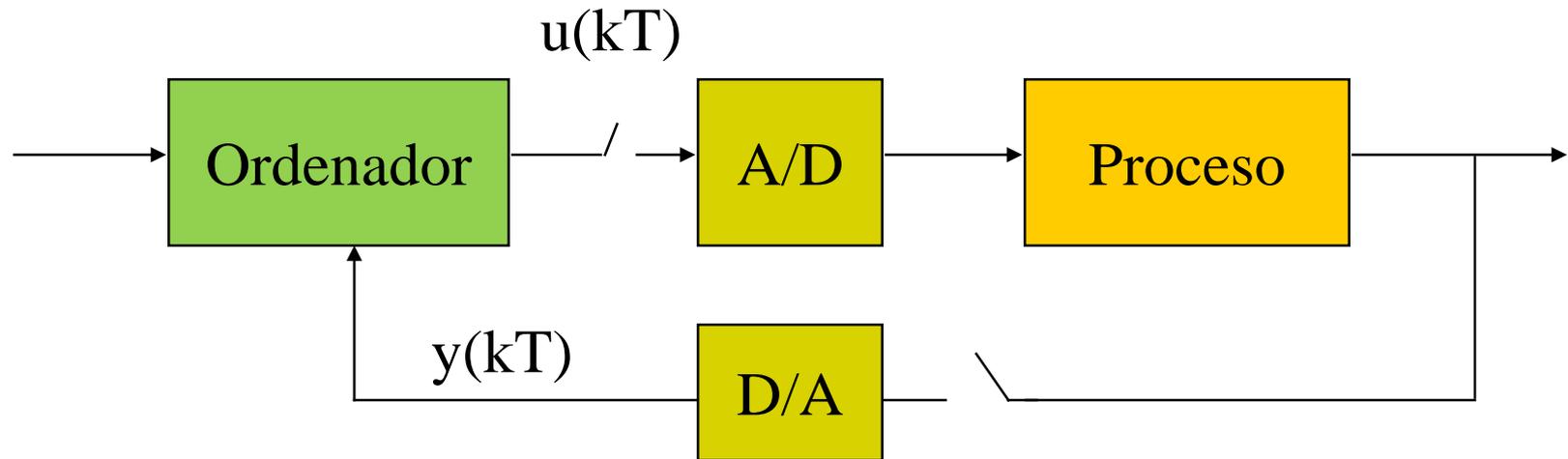
Modelos Dinámicos

- ✓ medida cuantitativa de los efectos de las entradas en las salidas a lo largo del tiempo
- ✓ deben representar los aspectos esenciales del proceso
- ✓ facilidad de manejo
- ✓ en el proceso siempre hay presentes perturbaciones





Modelos discretos



- modelos discretos
- la salida en $t = kT$ depende de las entradas y salidas en instantes de tiempo $(k-j)T$ anteriores



Tipos de modelos

- ✓ Modelos lineales:
 - Paramétricos (FT, SS,...)
 - No paramétricos (Resp. salto, Resp. Frec.)
- ✓ No lineales:
 - Hammerstein
 - Redes Neuronales
 - Modelos basados en leyes fisico-químicas
 - ...



Modelos Lineales

- ✓ la salida actual depende linealmente de las entradas y salidas en instantes de tiempo anteriores

$$y(t) = -a_1 y(t-1) - \dots - a_n y(t-n) + b_1 u(t-1) + \dots + b_m u(t-m)$$

las variables u e y son perturbaciones sobre un punto de operación

$$\begin{aligned} y(t) &= Y - Y_0 \\ u(t) &= U - U_0 \end{aligned}$$



Modelos lineales

las variables u e y son cambios sobre un punto de operación U_0 , Y_0

$$u(t) = U(t) - U_0(t)$$

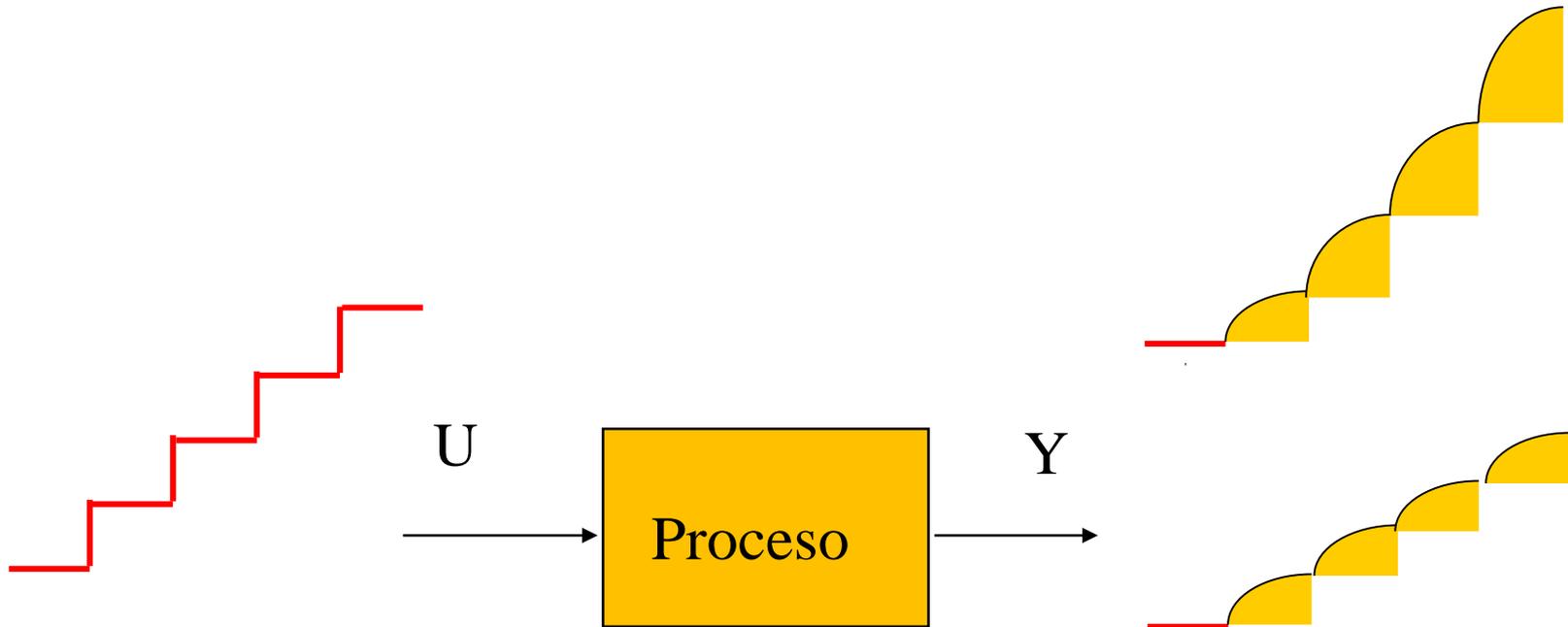
$$y(t) = Y(t) - Y_0(t)$$



El rango de validez está limitado a un entorno del punto de operación



¿Cuándo un proceso es lineal?





FUNCION DE TRANSFERENCIA

Operador retardo unitario

$$q^{-1}f(t) = f(t-1) \quad f(t+1) = qf(t) \quad \text{con } t = kT$$

$$y(t) = -a_1y(t-1) - \dots - a_ny(t-n) + b_1u(t-1) + \dots + b_mu(t-m)$$

$$(1 + a_1q^{-1} + \dots + a_nq^{-n})y(t) = (b_1q^{-1} + \dots + b_mu^{-m})u(t)$$

$$y(t) = \frac{B(q^{-1})}{A(q^{-1})} u(t)$$

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n}$$

$$B(q^{-1}) = b_1q^{-1} + \dots + b_mu^{-m}$$

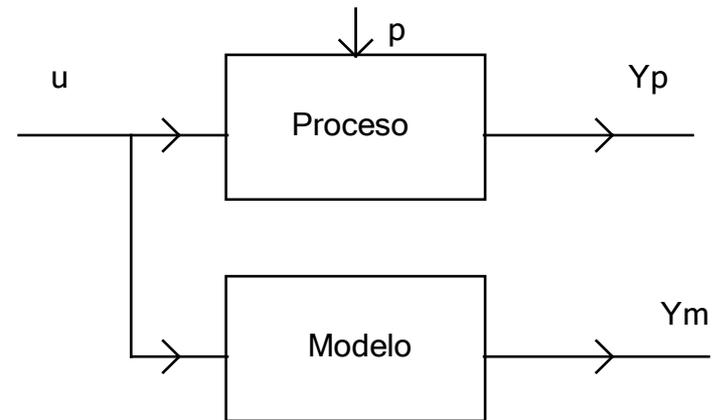
modelo DARMA



Modelos

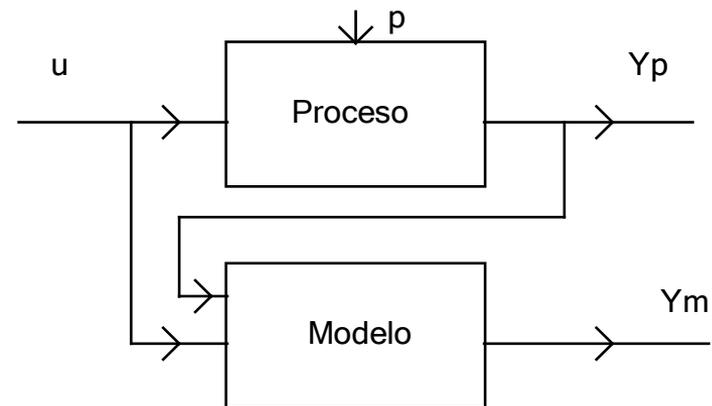
✓ Independientes

- y_m solo depende de u
- estables



✓ De regresión

- y_m depende también de p
- todo tipo





Modelo independiente/regresión

Modelo independiente

$$y_{\text{mod}}(t) = -a_1 y_{\text{mod}}(t-1) - \dots - a_n y_{\text{mod}}(t-n) + b_1 u(t-1) + \dots + b_m u(t-m)$$

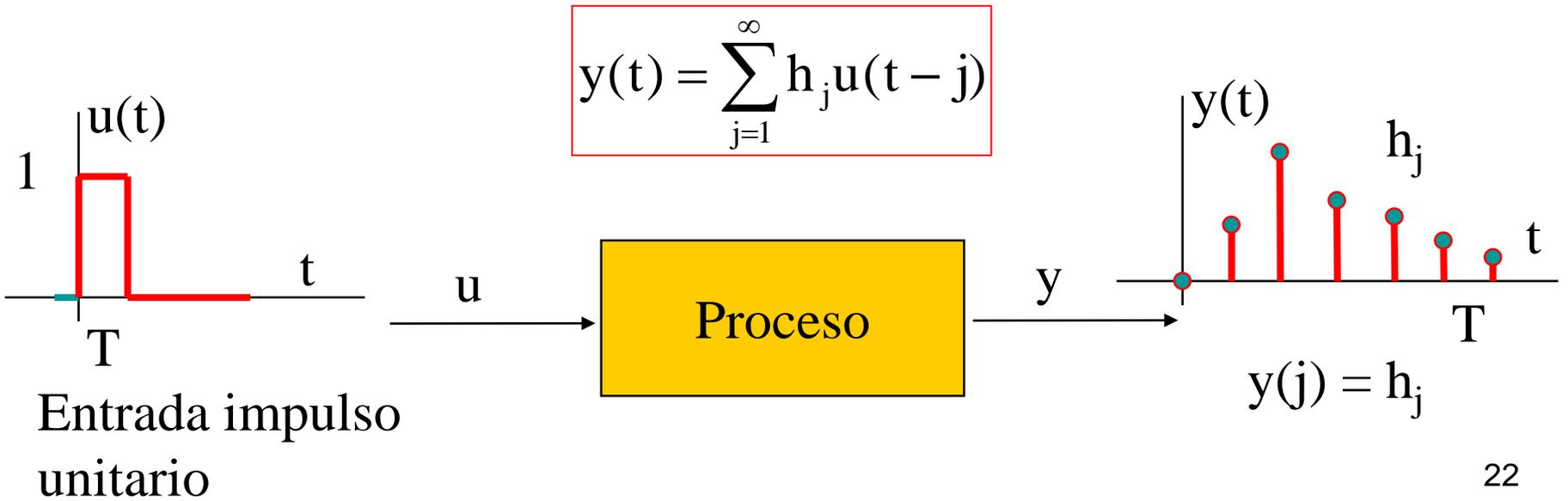
Modelo de regresión

$$y_{\text{mod}}(t) = -a_1 y_p(t-1) - \dots - a_n y_p(t-n) + b_1 u(t-1) + \dots + b_m u(t-m)$$



Respuesta Impulsional

$$y(t) = \frac{B(q^{-1})}{A(q^{-1})} u(t) = (h_1 q^{-1} + h_2 q^{-2} + \dots + h_m q^{-m} + \dots) u(t) = \\ = h_1 u(t-1) + h_2 u(t-2) + \dots + h_m u(t-m) + \dots$$





Respuesta impulsional

$$y(t) = \sum_{j=1}^{\infty} h_j u(t-j) \approx \sum_{j=1}^m h_j u(t-j) = h_1 u(t-1) + h_2 u(t-2) + \dots + h_m u(t-m)$$

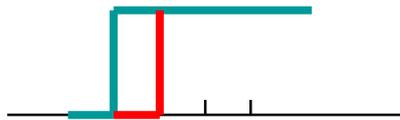
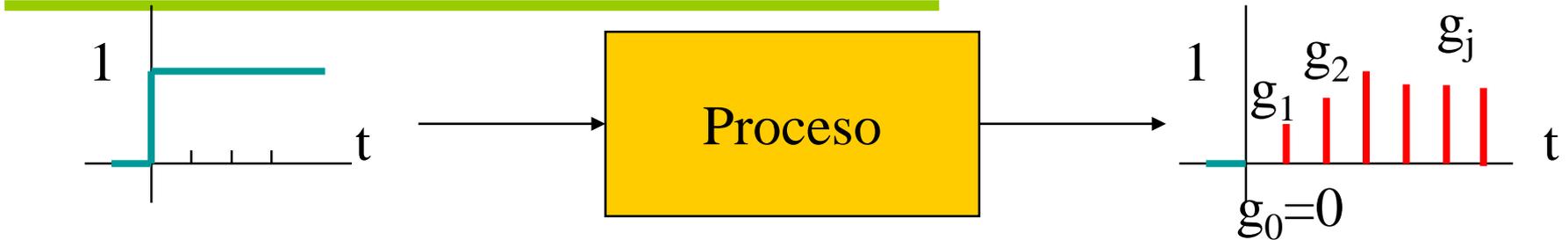
$$y(t) = (h_1 q^{-1} + h_2 q^{-2} + \dots + h_n q^{-n}) u(t) = H(q^{-1}) u(t)$$

- No requiere conocer la estructura del modelo
- Puede describir dinámicas no usuales
- Predicción simple y poco sensible a errores

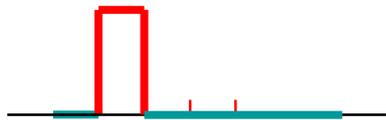
- Limitado a sistemas asintóticamente estables
- Contiene muchos parámetros
- Sistemas con estructura poco definida: temp,



Respuesta salto



Impulso = diferencia de dos saltos retrasados un periodo de muestreo T



$$h_j = g_j - g_{j-1}$$

$$y(t) = \sum_{j=1}^{\infty} h_j u(t-j) = \sum_{j=1}^{\infty} (g_j - g_{j-1}) u(t-j) =$$

$$= g_1 u(t-1) - g_0 u(t-1) + g_2 u(t-2) - g_1 u(t-2) + g_3 u(t-3) - g_2 u(t-3) + \dots =$$

$$= g_1 (u(t-1) - u(t-2)) + g_2 (u(t-2) - u(t-3)) + g_3 (u(t-3) - u(t-4)) + \dots =$$

$$= g_1 \Delta u(t-1) + g_2 \Delta u(t-2) + g_3 \Delta u(t-3) + \dots = \sum_{j=1}^{\infty} g_j \Delta u(t-j)$$



Respuesta salto

$$y(t) = \sum_{j=1}^{\infty} g_j \Delta u(t-j)$$

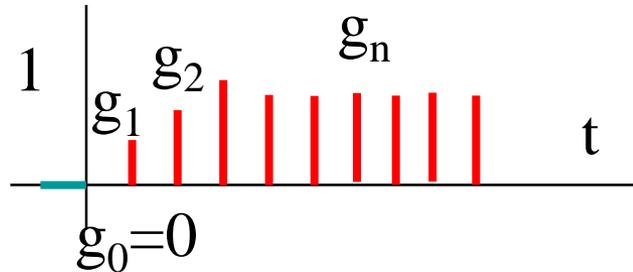
$$\Delta = 1 - q^{-1}$$

- No requiere conocer la estructura del modelo
- Puede describir dinámicas no usuales
- Predicción simple y poco sensible a errores

- Limitado a sistemas asintóticamente estables
- Contiene muchos parámetros
- Adecuado para sistemas de estructura poco definida



Respuesta salto



Si el sistema es
asintóticamente estable
los g_j serán constantes a
partir de uno dado g_n

$$\begin{aligned}y(t) &= \sum_{j=1}^{\infty} g_j \Delta u(t-j) = g_1 \Delta u(t-1) + g_2 \Delta u(t-2) + \dots + g_n \Delta u(t-n) + \\ &+ g_{n+1} [\Delta u(t-n-1) + \Delta u(t-n-2) + \dots] = \\ &= g_1 \Delta u(t-1) + \dots + g_{n+1} [u(t-n-1) - u(t-n-2) + u(t-n-2) - u(t-n-3) + \dots] = \\ &= g_1 \Delta u(t-1) + \dots + g_n \Delta u(t-n) + g_{n+1} [u(t-n-1)] \\ y(t) &= \sum_{j=1}^{\infty} g_j \Delta u(t-j) \approx \sum_{j=1}^n g_j \Delta u(t-j) + g_{n+1} u(t-(n+1))\end{aligned}$$



Variables de estado

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ y(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned}$$

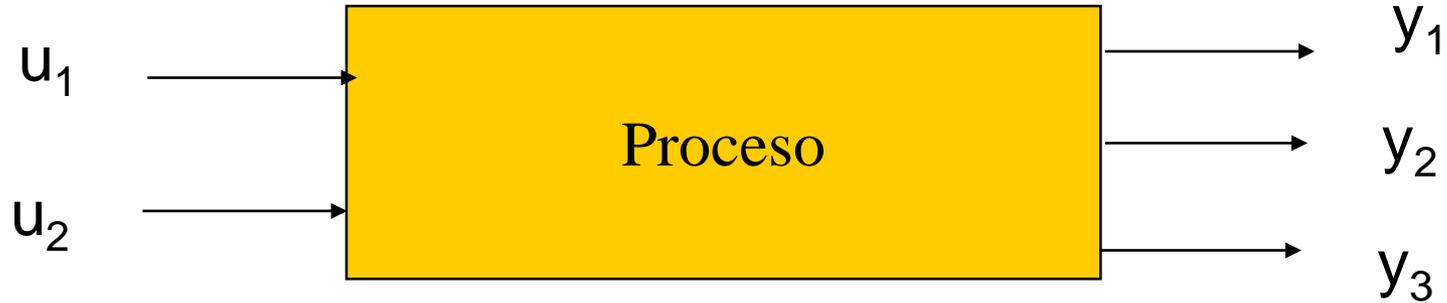
$\mathbf{x}(t)$ variable de estado del sistema

A, B y C matrices de ordenes adecuados

- Se obtienen de forma natural por linealización de los modelos dinámicos obtenidos mediante balances, etc
 - Describe bien la estructura interna del proceso
 - Extensión multivariable directa
 - Numero de parámetros no mínimo
 - Cálculos mas complejos
- Métodos de subespacios



Caso Multivariable



Las estructuras anteriores se pueden generalizar al caso MIMO

$$A(q^{-1})\mathbf{y}(t) = B(q^{-1})\mathbf{u}(t) \quad \mathbf{u}(t)' = [u_1(t), \dots, u_m(t)] \quad \mathbf{y}(t)' = [y_1(t), \dots, y_n(t)]$$

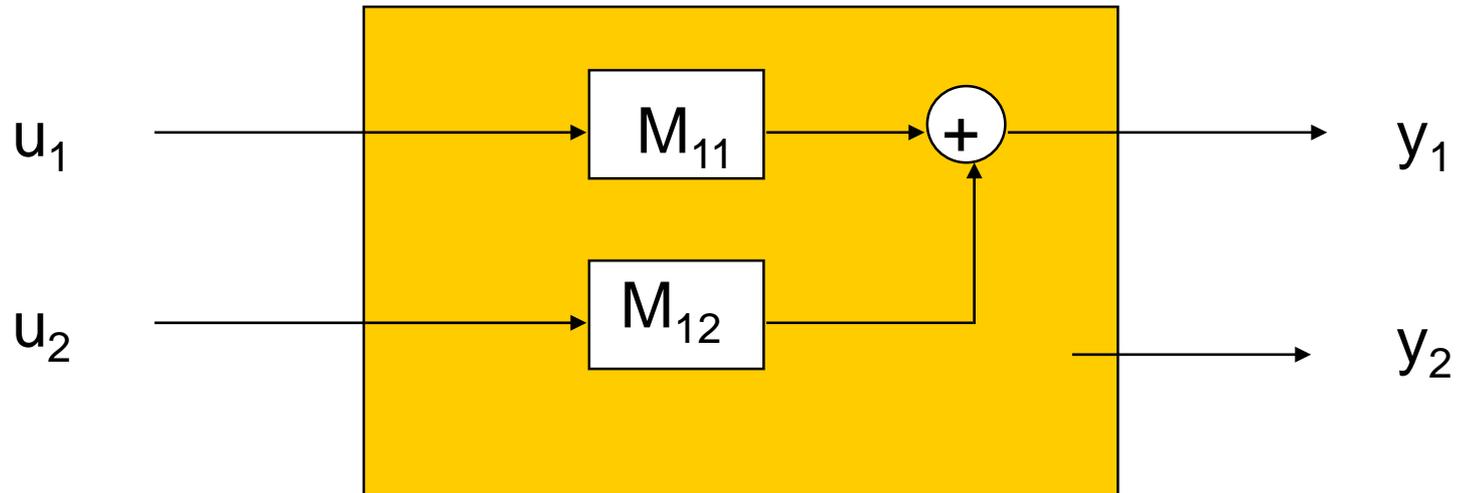
$$A(q^{-1}) = I + A_1q^{-1} + \dots + A_pq^{-p} \quad B(q^{-1}) = B_1q^{-1} + \dots + B_qq^{-q}$$

Modelo en fracciones matriciales: A_i matrices (n,n)

MFD matrix fraction description



Caso Multivariable



Si la Matriz A de la descripción MFD es diagonal:

$$y_1 = M_{11} u_1 + M_{12} u_2 + \dots$$



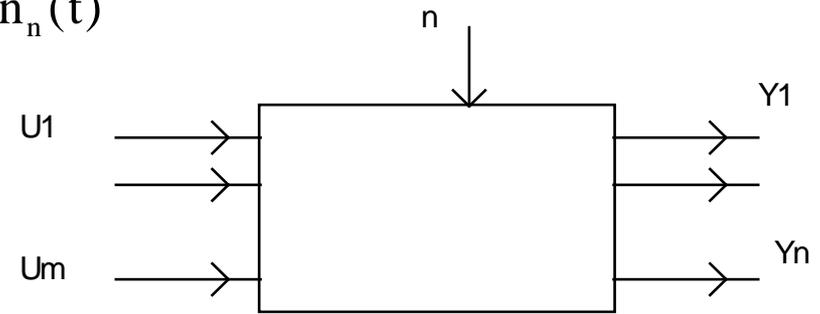
Caso Multivariable

$$y_1(t) = M_{11}u_1(t) + M_{12}u_2(t) + \dots + M_{1m}u_m(t) + n_1(t)$$

$$y_2(t) = M_{21}u_1(t) + M_{22}u_2(t) + \dots + M_{2m}u_m(t) + n_2(t)$$

...

$$y_n(t) = M_{n1}u_1(t) + M_{n2}u_2(t) + \dots + M_{nm}u_m(t) + n_n(t)$$

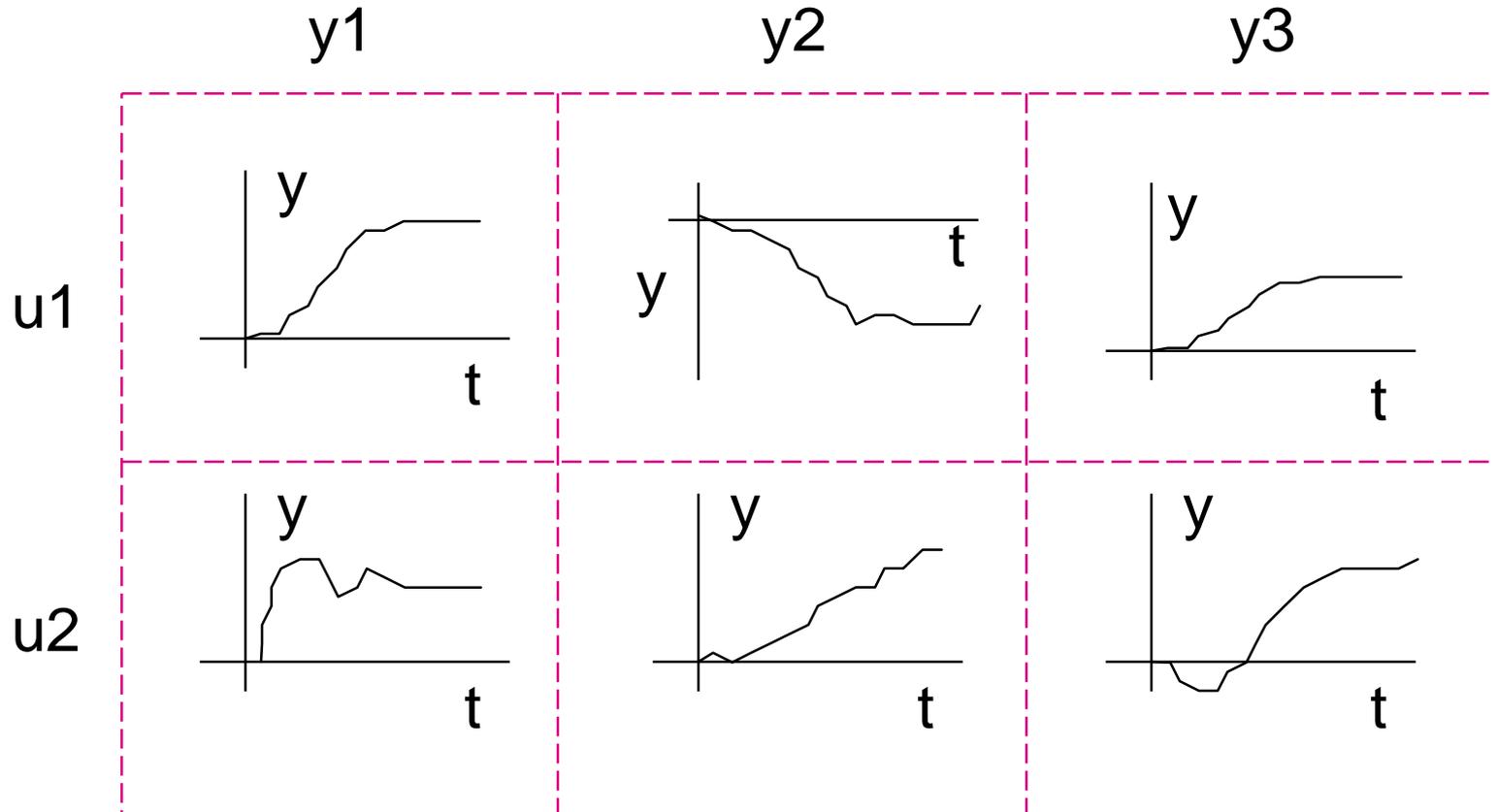


$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \dots \\ y_n(t) \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & \dots & M_{1m} \\ M_{21} & M_{22} & \dots & M_{2m} \\ \dots & \dots & \dots & \dots \\ M_{n1} & M_{n2} & \dots & M_{nm} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \\ \dots \\ u_m(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \\ \dots \\ n_n(t) \end{bmatrix}$$

$$y = Mu + n$$



Matriz Dinámica

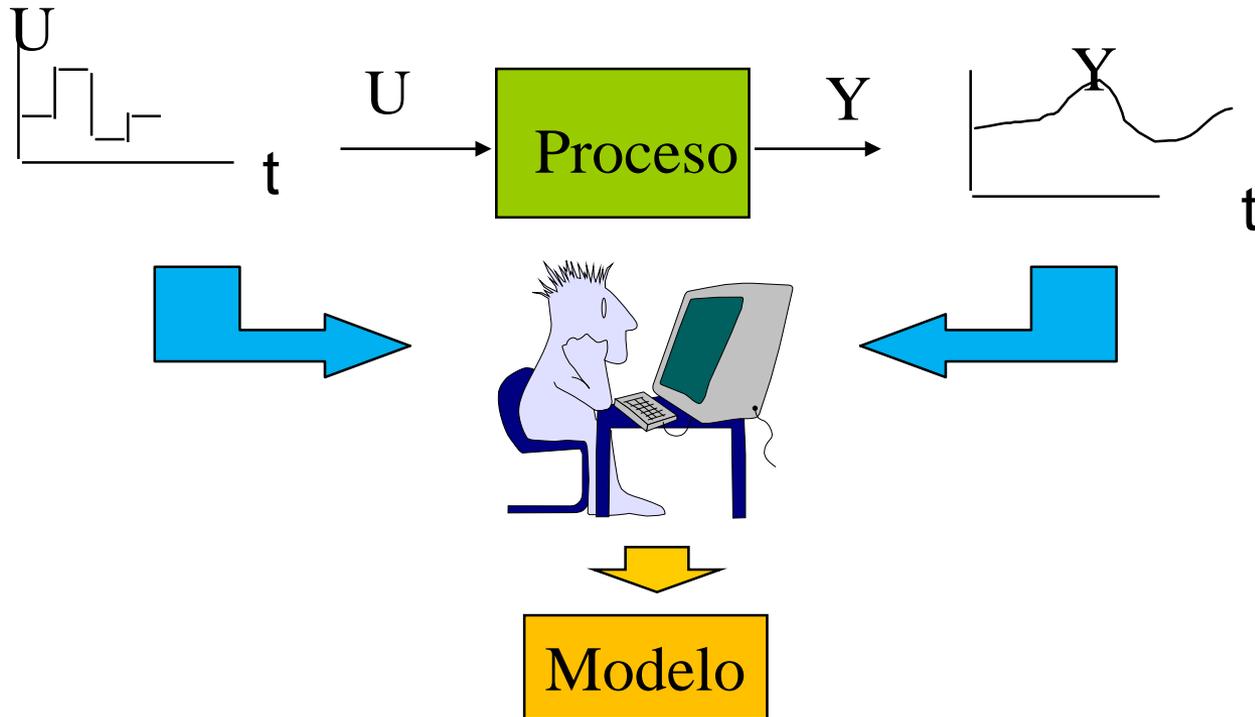


respuesta salto entre cada par entrada/salida



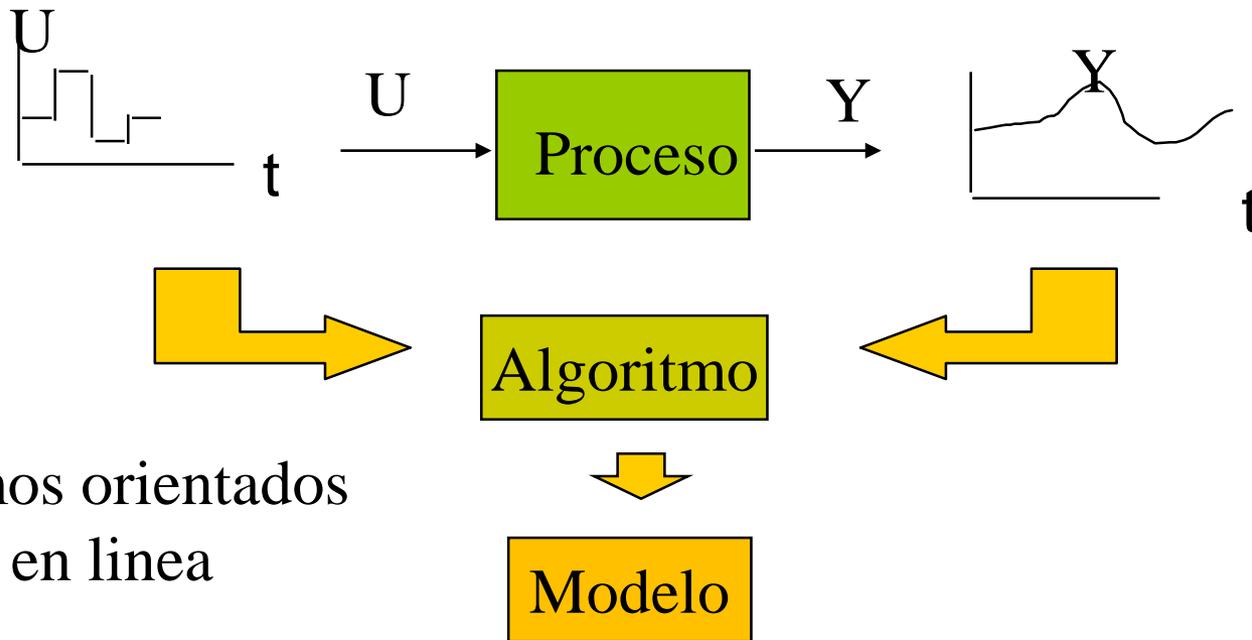
Algoritmos No recursivos

Orientados a identificación fuera de línea





Algoritmos Recursivos



Algoritmos orientados
a trabajo en línea

$$\theta(t) = \theta(t - 1) + K(t)[y(t) - M(t - 1)u(t)]$$

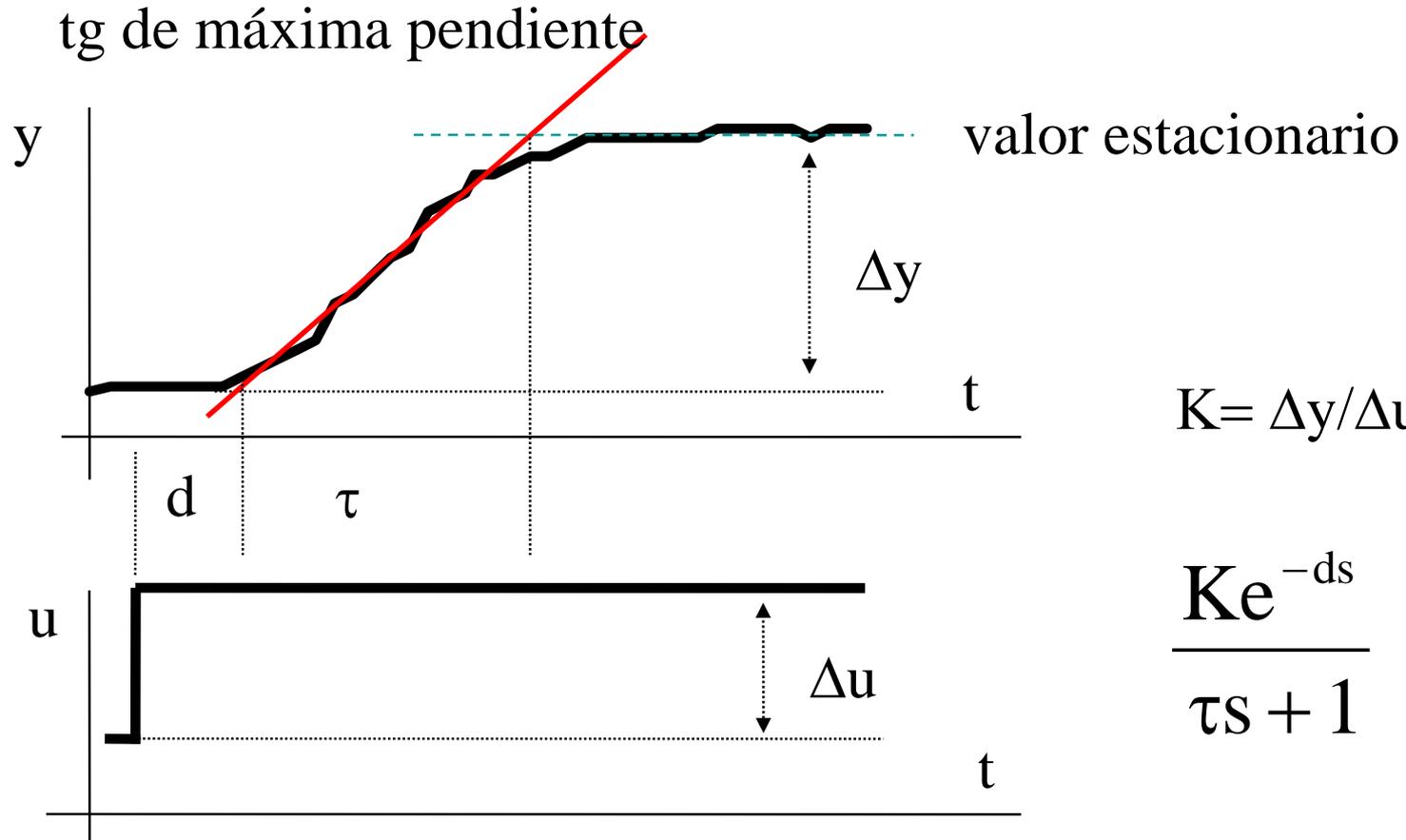


Algunos Métodos de Estimación

- ✓ Respuestas a entradas especiales
- ✓ Métodos basados en correlaciones
- ✓ Métodos de estimación espectral
- ✓ Métodos de minimización de error
- ✓ Métodos de variable instrumental
- ✓ Métodos de predicción de error
- ✓ Métodos de máxima verosimilitud
- ✓ Identificación global
- ✓ Minimización de la distancia paramétrica
- ✓ Subespacios invariantes



Identificación con un salto en u



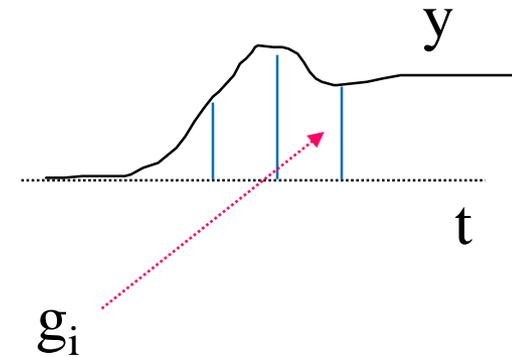
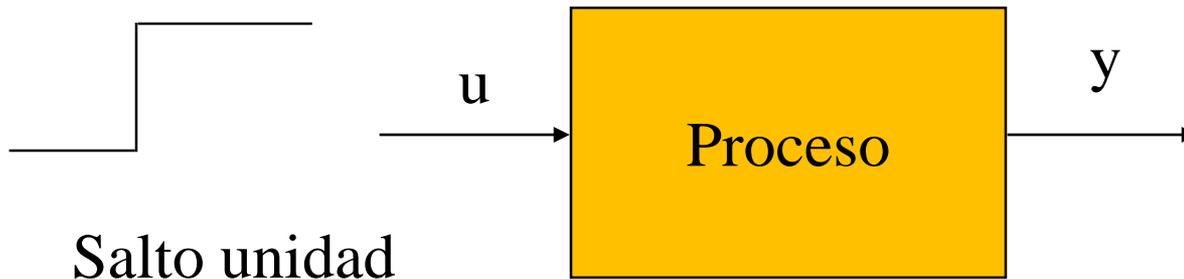


Ensayos de respuesta en salto

$$y(t) = \sum_{i=1}^{\infty} g_i \Delta u(t-i) + v(t)$$

Se parte de un estado estacionario

Para eliminar el efecto de $v(t)$ se toma la media de varios ensayos





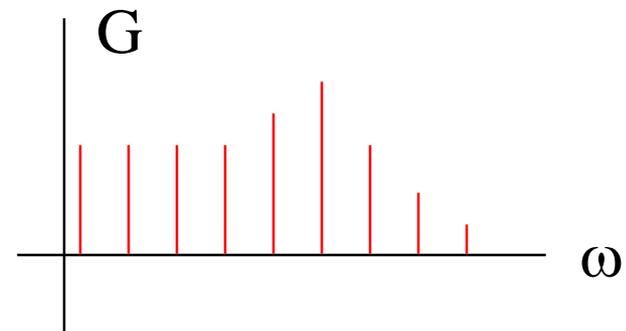
Respuesta en frecuencia

Calculo de los diagramas de amplitud y fase excitando al sistema con señales sinusoidales de distintas frecuencias



Modelo no paramétrico

Métodos especiales con correladores para eliminar el efecto de ruidos





Funciones de correlación

La autocovarianza de una serie de datos estima la dependencia interna entre el valor que toma en un instante t y el que toma en $t+k$

$$R_u(k) = E\{u(t)u(t-k)\} \cong \frac{1}{N-1} \sum_{i=k}^N u(i)u(i-k)$$

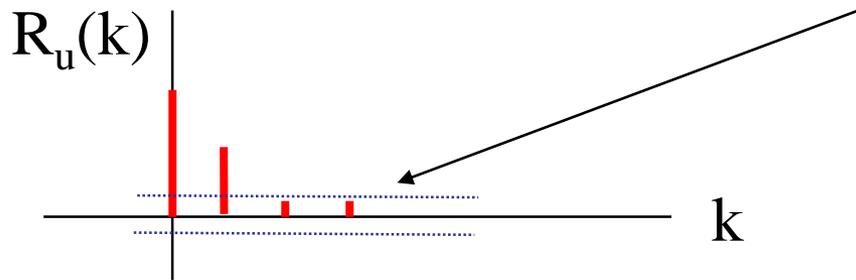
La covarianza cruzada entre dos series de datos estima la dependencia entre el valor que toma una en un instante t y el que toma la otra en $t+k$

$$R_{yu}(k) = E\{y(t)u(t-k)\} \cong \frac{1}{N-1} \sum_{i=k}^N y(i)u(i-k)$$

$$R_{yu}(k) = R_{uy}(-k)$$



Bandas de confianza



Si los valores calculados están dentro de las bandas de confianza se supone que no son significativamente diferentes de cero

Las bandas de confianza se calculan a menudo mediante la expresión:

$$\frac{2.009 \sqrt{1 - R(k)}}{n - 1}$$

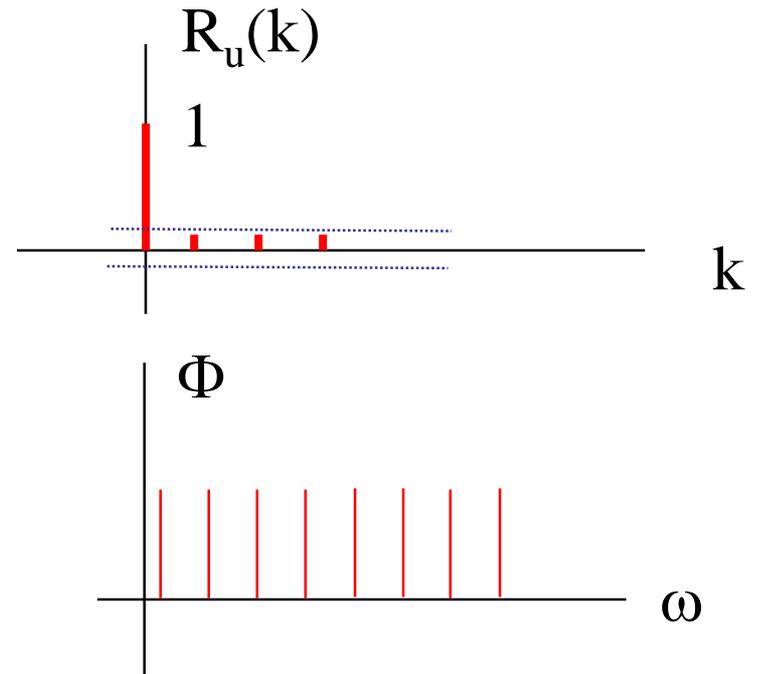


Ruido blanco

Señal aleatoria en la que sus valores en un instante t son independientes del valor que han tomado en instantes anteriores.

$$R_u(k) = 0 \quad k \neq 0$$

Su espectro de potencia es constante





Métodos de correlación



$$y(t) = \sum_{i=1}^m h_i u(t-i) + v(t) \quad y(t)u(t-j) = \sum_{i=1}^m h_i u(t-i)u(t-j) + v(t)u(t-j)$$

$$E\{y(t)u(t-j)\} = E\left\{\sum_{i=1}^m h_i u(t-i)u(t-j)\right\} + E\{v(t)u(t-j)\}$$

$$E\{y(t)u(t-j)\} = \sum_{i=1}^m h_i E\{u(t-i)u(t-j)\} + E\{v(t)u(t-j)\}$$

$$R_{yu}(j) = \sum_{i=1}^m h_i R_u(j-i) + R_{vu}(j)$$

Sistemas estables en lazo abierto



Métodos de correlación

$$R_{yu}(j) = \sum_{i=1}^m h_i R_u(j-i) + R_{vu}(j)$$

En **lazo abierto**, v y u están incorrelacionadas y

$R_{vu}(j)=0$, pero si $v(t)$ es un proceso ARMA:

$$v(t) = -d_1 v(t-1) - \dots - d_r v(t-r) + \xi(t) + c_1 \xi(t-1) + \dots + c_s \xi(t-s)$$

$$R_{vu}(j) \approx \frac{1}{N} \sum_{t=1}^{N-j} v(t)u(t-j) =$$

$$= \frac{1}{N} \sum_{t=1}^{N-j} [-d_1 v(t-1) - \dots - d_r v(t-r) + \xi(t) + c_1 \xi(t-1) + \dots + c_s \xi(t-s)] u(t-j)$$

En lazo cerrado $u(t-j)$ puede depender a través de $y(t-j-k)$ de $v(t-j-i)$ $E\{v(t)u(t-j)\} \neq 0$

$$R_{yu}(j) = \sum_{i=1}^m h_i R_u(j-i)$$



Métodos de correlación

$$R_{yu}(j) = \sum_{i=1}^{\infty} h_i R_u(j-i) \approx \sum_{i=1}^m h_i R_u(i-j)$$

Dando
valores
0....m-1,
a j:

$$\begin{bmatrix} R_{yu}(0) \\ R_{yu}(1) \\ \dots \\ R_{yu}(m-1) \end{bmatrix} = \begin{bmatrix} R_u(1) & R_u(2) & \dots & R_u(m) \\ R_u(0) & R_u(1) & \dots & R_u(m-1) \\ \dots & \dots & \dots & \dots \\ R_u(2-m) & \dots & \dots & R_u(1) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \dots \\ h_m \end{bmatrix}$$

Para que exista
solución

$$\begin{bmatrix} R_u(1) & R_u(2) & \dots & R_u(m) \\ R_u(0) & R_u(1) & \dots & R_u(m-1) \\ \dots & \dots & \dots & \dots \\ R_u(2-m) & \dots & \dots & R_u(1) \end{bmatrix}$$

debe estar
bien
condicionada

Condición de **excitación persistente**. Depende del experimento



PRBS

$$R_{yu}(j) = \sum_{i=1}^m h_i R_u(i-j)$$

Si la entrada es un ruido blanco: $R_u(k) = 0 \quad k \neq 0$

$$R_{yu}(j) = h_j R_u(0)$$

Una ruido blanco es difícil de implementar, por lo que, a menudo, se usan en su lugar otras señales con características similares a las de ruido blanco:

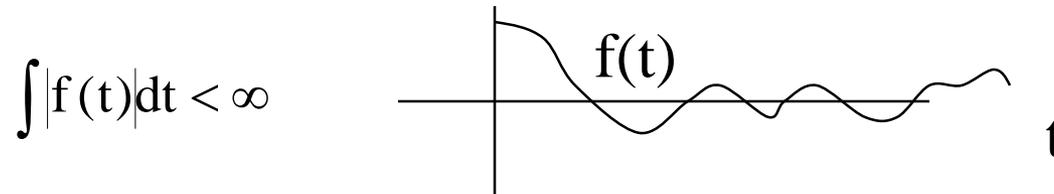
PRBS

DRBS



Espectro de potencia

Las condiciones teóricas de existencia de la Transformada de Fourier de una señal no se cumplen en el caso de procesos estocásticos



Se define entonces el espectro de potencia sobre $R(t)$

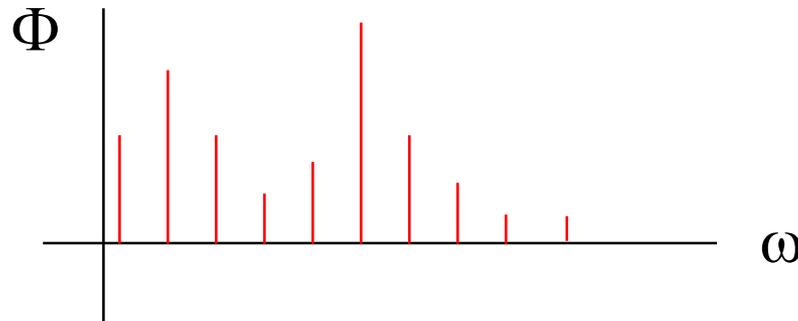
$$\Phi_u(\omega) = \int_{-\infty}^{\infty} R_u(t) e^{-j\omega t} dt$$

$$\Phi_{yu}(\omega) = \int_{-\infty}^{\infty} R_{yu}(t) e^{-j\omega t} dt$$



Espectro de potencia

$$\Phi_u(\omega) = \frac{1}{2\pi} E\left\{ |U(\omega)|^2 \right\}$$



El espectro de potencia de una señal es proporcional al valor esperado del cuadrado de la amplitud de la componente de la señal asociada a esa frecuencia, o sea a su potencia.

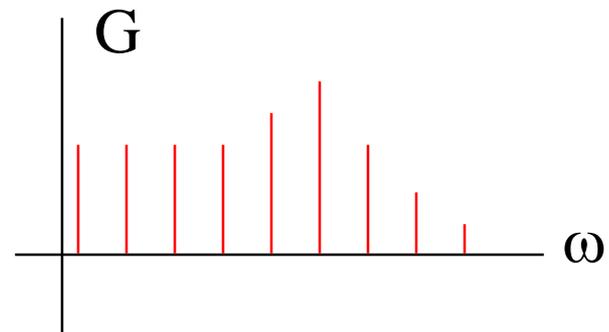


Estimación de espectros

Dominio de la frecuencia

Modelo no paramétrico: se estiman los valores de $G(j\omega)$ para distintos valores de la frecuencia

$$\hat{G}(e^{j\omega T}) = \frac{\Phi_{yu}(\omega)}{\Phi_{uu}(\omega)}$$

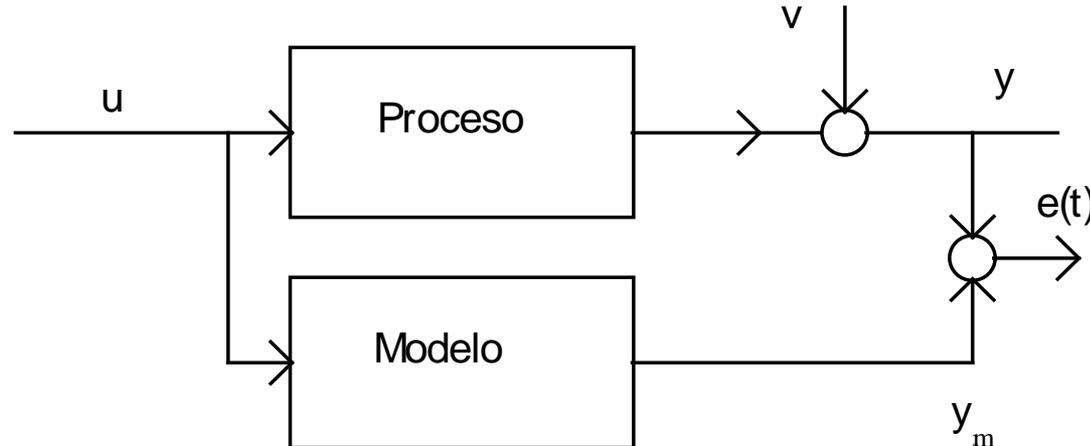




Minimización de error

Criterio de estimación: Dado un conjunto de datos experimentales $u(t)$, $y(t)$, buscar los parámetros θ que minimizan la función de coste V :

$$V = \frac{1}{N} \sum_{t=1}^N e(t)^2 = \frac{1}{N} \sum_{t=1}^N [(y(t) - y_m(t, \theta))]^2$$





Método de Mínimos cuadrados

Least Squares (LS)

Modelo

θ vector de parámetros

$$y_m(t) = \varphi(t)' \theta$$

φ vector de datos

Criterio de estimación: Dado un conjunto de datos experimentales $u(t)$, $y(t)$, minimizar respecto a los parámetros θ :

$$V = \frac{1}{N} \sum_{t=1}^N e(t)^2 = \frac{1}{N} \sum_{t=1}^N [(y(t) - y_m(t))]^2 = \frac{1}{N} \sum_{t=1}^N [(y(t) - \varphi(t)' \theta)]^2$$



Respuesta impulso

$$y_m(t) = h_1 u(t-1) + h_2 u(t-2) + \dots + h_m u(t-m)$$

$$y_m(t) = \varphi(t)' \theta$$

$$\varphi(t)' = [u(t-1), u(t-2), \dots, u(t-m)]$$

$$\theta' = [h_1, h_2, \dots, h_m]$$

La salida del modelo depende de valores pasados de las entradas del proceso



Respuesta salto

$$y_m(t) = g_1 \Delta u(t-1) + g_2 \Delta u(t-2) + \dots + g_m \Delta u(t-m) + g_{m+1} u(t-m-1)$$

$$y_m(t) = \varphi(t)' \theta$$

$$\varphi(t)' = [\Delta u(t-1), \Delta u(t-2), \dots, \Delta u(t-m), u(t-m-1)]$$

$$\theta' = [g_1, g_2, \dots, g_m, g_{m+1}]$$

La salida del modelo depende de valores pasados de las entradas del proceso



Modelo de regresión

$$y_m(t) = -a_1 y(t-1) - a_2 y(t-2) - \dots - a_n y(t-n) \\ + b_1 u(t-1) + \dots + b_m u(t-m)$$

$$y_m(t) = \varphi(t)' \theta$$

$$\varphi(t)' = [-y(t-1), \dots, -y(t-n), u(t-1), \dots, u(t-m)]$$

$$\theta' = [a_1, \dots, a_n, b_1, \dots, b_m]$$

La salida del modelo depende de valores pasados de las entradas y salidas del **proceso**



Función de transferencia

$$y_m(t) = \frac{B(q^{-1})}{A(q^{-1})} u(t)$$

$$y_m(t) = -a_1 y_m(t-1) - a_2 y_m(t-2) - \dots - a_n y_m(t-n) + b_1 u(t-1) + \dots + b_m u(t-m)$$

Una función de transferencia no puede ponerse como función lineal de θ

$$y_m(t) = \varphi(t)' \theta$$

Al ser las $y_m(t-i)$ función de los parámetros θ

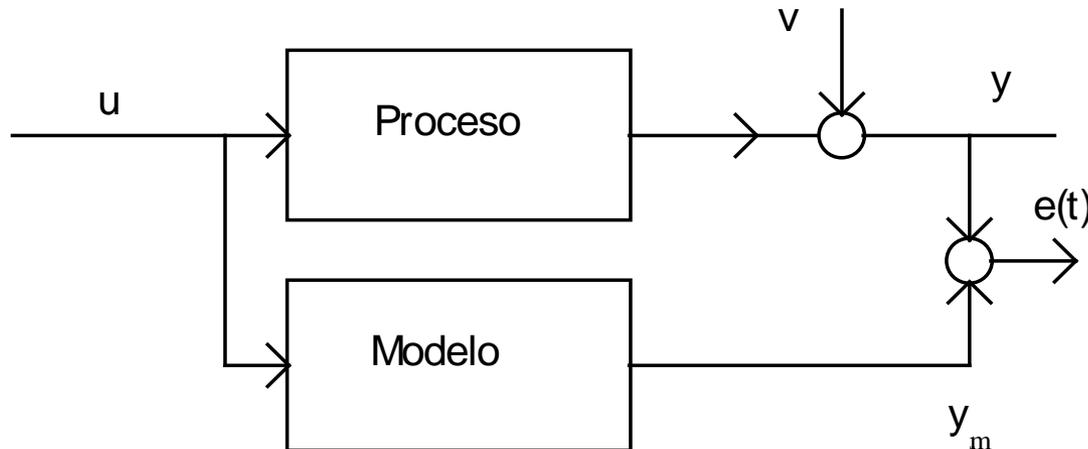


Método de Mínimos cuadrados

Least Squares (LS) $y_m(t) = \varphi(t)' \theta$

Criterio de estimación: Dado un conjunto de datos experimentales $u(t)$, $y(t)$, minimizar respecto a los parámetros θ :

$$V = \frac{1}{N} \sum_{t=1}^N e(t)^2 = \frac{1}{N} \sum_{t=1}^N [(y(t) - y_m(t))]^2 = \frac{1}{N} \sum_{t=1}^N [(y(t) - \varphi(t)' \theta)]^2$$





LS formulación matricial

$$V = \frac{1}{N} \sum_{t=1}^N [(y(t) - \varphi(t)' \theta)]^2$$

Definiendo la
matriz y el vector:

$$\Phi' = [\varphi(1) | \varphi(2) | \dots | \varphi(N)]$$

$$\mathbf{y}' = [y(1), y(2), \dots, y(N)]$$

$$\mathbf{y} - \Phi \boldsymbol{\theta} = \begin{bmatrix} y(1) \\ y(2) \\ \dots \\ y(N) \end{bmatrix} - \begin{bmatrix} \varphi(1)' \\ \varphi(2)' \\ \dots \\ \varphi(N)' \end{bmatrix} \boldsymbol{\theta} = \begin{bmatrix} y(1) - \varphi(1)' \theta \\ y(2) - \varphi(2)' \theta \\ \dots \\ y(N) - \varphi(N)' \theta \end{bmatrix}$$

$$V = \frac{1}{N} (\mathbf{y} - \Phi \boldsymbol{\theta})' (\mathbf{y} - \Phi \boldsymbol{\theta})$$



Ejemplo

$$y(t) = h_1 u(t-1) + h_2 u(t-2) + h_3 u(t-3)$$

Modelo resp. impulso
con 3 parametros h_i y
 $N=5$ (se usan $N+3-1$
datos experimentales)

$$y' = [y(1), y(2), y(3), y(4), y(5)]$$

$$\theta' = [h_1 \quad h_2 \quad h_3]$$

$$\varphi(t)' = [u(t-1), u(t-2), u(t-3)]$$

$$\Phi' = [\varphi(1) | \varphi(2) | \varphi(3) | \varphi(3) | \varphi(5)] =$$

$$= \begin{bmatrix} u(0) & u(1) & u(2) & u(3) & u(4) \\ u(-1) & u(0) & u(1) & u(2) & u(3) \\ u(-2) & u(-1) & u(0) & u(1) & u(2) \end{bmatrix}$$

$$\mathbf{y} - \Phi\boldsymbol{\theta} = \begin{bmatrix} y(1) \\ y(2) \\ y(3) \\ y(4) \\ y(5) \end{bmatrix} - \begin{bmatrix} u(0) & u(-1) & u(-2) \\ u(1) & u(0) & u(-1) \\ u(2) & u(1) & u(0) \\ u(3) & u(2) & u(1) \\ u(4) & u(3) & u(2) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$$



Mínimos cuadrados

$$\min_{\theta} V = \min_{\theta} \frac{1}{N} \sum_{t=1}^N [(y(t) - \varphi(t)' \theta)]^2$$

$$\Phi' = [\varphi(1) | \varphi(2) | \dots | \varphi(N)]$$
$$\mathbf{y}' = [y(1), y(2), \dots, y(N)]$$

$$V = \frac{1}{N} (\mathbf{y} - \Phi \theta)' (\mathbf{y} - \Phi \theta) = \frac{1}{N} [\mathbf{y}' \mathbf{y} - \theta' \Phi' \mathbf{y} - \mathbf{y}' \Phi \theta + \theta' \Phi' \Phi \theta]$$

$$\left. \frac{\partial V(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} = \frac{1}{N} [-2\Phi' \mathbf{y} + 2\Phi' \Phi \hat{\theta}] = \mathbf{0}$$

$$\hat{\theta} = [\Phi' \Phi]^{-1} \Phi' \mathbf{y}$$

$$[\Phi' \Phi] = \sum_{t=1}^N \varphi(t) \varphi(t)'$$

Debe ser invertible



Interpretación geométrica

$$\begin{aligned}\Phi' &= [\varphi(1) | \varphi(2) | \dots | \varphi(N)] \\ \mathbf{y}' &= [y(1), y(2), \dots, y(N)] \\ V &= \frac{1}{N} (\mathbf{y} - \Phi\theta)' (\mathbf{y} - \Phi\theta)\end{aligned}\quad \Phi = \begin{bmatrix} \frac{\varphi'(1)}{\varphi'(N)} \\ \frac{\varphi'(2)}{\varphi'(N)} \\ \dots \\ \frac{\varphi'(N)}{\varphi'(N)} \end{bmatrix} = [\phi_1 | \phi_2 | \dots | \phi_d]$$

Considerando las d columnas ϕ_i de Φ como vectores de dimensión N , el problema de encontrar θ para minimizar V se puede ver como el de encontrar una combinación lineal de los vectores ϕ_i , o sea, un vector del hiperplano definido por ellos, que esté lo más próximo posible al vector \mathbf{y} . Esto corresponde a la proyección ortogonal de \mathbf{y} sobre el hiperplano. $d =$ número de parámetros a estimar



Ejemplo

$$y(t) = h_1 u(t-1) + h_2 u(t-2) + h_3 u(t-3)$$

$$\mathbf{y} - \Phi \boldsymbol{\theta} = \begin{bmatrix} y(1) \\ y(2) \\ y(3) \\ y(4) \\ y(5) \end{bmatrix} - \begin{bmatrix} u(0) & u(-1) & u(-2) \\ u(1) & u(0) & u(-1) \\ u(2) & u(1) & u(0) \\ u(3) & u(2) & u(1) \\ u(4) & u(3) & u(2) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} =$$

Modelo resp. impulso
con 3 parametros h_i y
 $N=5$ (se usan $N+3-1$
datos experimentales)

$$= \begin{bmatrix} y(1) \\ y(2) \\ y(3) \\ y(4) \\ y(5) \end{bmatrix} - h_1 \begin{bmatrix} u(0) \\ u(1) \\ u(2) \\ u(3) \\ u(4) \end{bmatrix} - h_2 \begin{bmatrix} u(-1) \\ u(0) \\ u(1) \\ u(2) \\ u(3) \end{bmatrix} - h_3 \begin{bmatrix} u(-2) \\ u(-1) \\ u(0) \\ u(1) \\ u(2) \end{bmatrix}$$

θ_1 θ_2 θ_3

ϕ_1 ϕ_2 ϕ_3

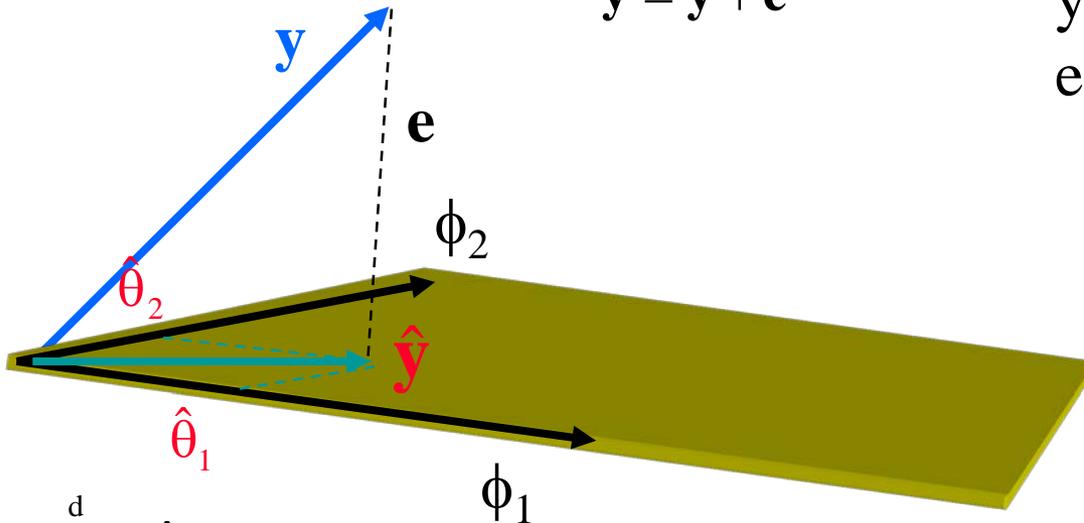
$$\mathbf{y} - \Phi \boldsymbol{\theta} = \mathbf{y} - \sum_{i=1}^d \phi_i \theta_i$$



Interpretación geométrica

$$\begin{aligned} \mathbf{e} &= \mathbf{y} - \hat{\mathbf{y}} \\ \mathbf{e} &\perp \phi_i \quad \mathbf{e}'\phi_i = 0 \\ \mathbf{y} &= \hat{\mathbf{y}} + \mathbf{e} \end{aligned}$$

El vector de errores \mathbf{e} es perpendicular a cualquier vector ϕ_i , y por tanto, a las estimas $\hat{\mathbf{y}}$



$$\hat{\mathbf{y}} = \sum_{i=1}^d \phi_i \hat{\theta}_i$$



Residuos

Errores entre los datos experimentales y los valores estimados con el modelo

$$e(t) = y(t) - \hat{y}(t) = y(t) - \varphi(t)' \theta$$

$$\mathbf{e} = \begin{bmatrix} e(1) \\ e(2) \\ \vdots \\ e(N) \end{bmatrix}$$

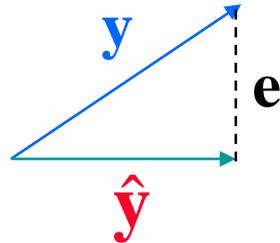
$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$$

$$\mathbf{e} \perp \hat{\mathbf{y}}$$

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{e}$$

$$|\mathbf{y}|^2 = |\hat{\mathbf{y}}|^2 + |\mathbf{e}|^2$$

$$\frac{1}{N} \sum_{t=1}^N y(t)^2 = \frac{1}{N} \sum_{t=1}^N \hat{y}(t)^2 + \frac{1}{N} \sum_{t=1}^N e(t)^2$$



Coefficiente de correlación múltiple: mide la proporción de la varianza de los datos que es “explicada” por el modelo

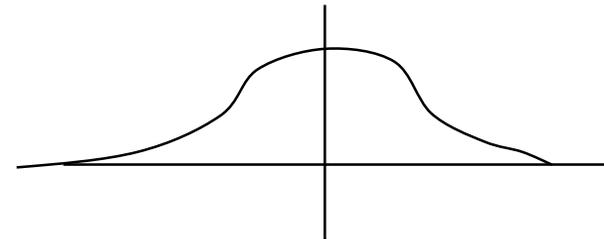
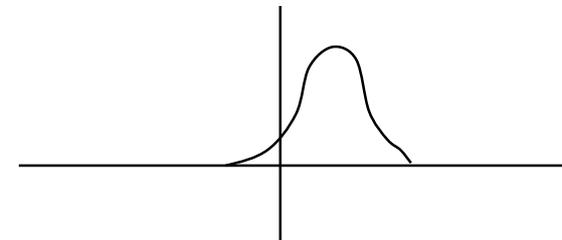
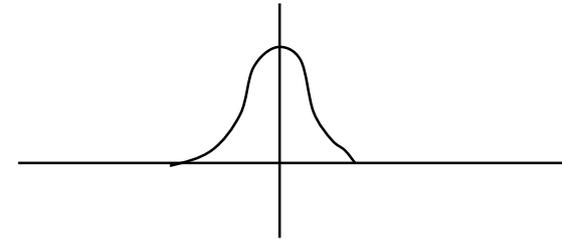
$$R_y^2 = \frac{\sum \hat{y}(t)^2}{\sum y(t)^2} = 1 - \frac{\sum e(t)^2}{\sum y(t)^2}$$



Propiedades

- ✓ Sesgo
- ✓ Consistencia
- ✓ Eficiencia

Los resultados estimados de θ dependen de los experimentos y son, por tanto variables estocásticas





Propiedades (1)

Suponiendo que el proceso puede ser representado de forma exacta por $y(t) = \phi(t)' \theta_0 + v(t)$

$$\hat{\theta} = [\Phi' \Phi]^{-1} \Phi' \mathbf{y}$$

$$\hat{\theta} = [\Phi' \Phi]^{-1} \Phi' [\Phi \theta_0 + \mathbf{v}]$$

$$\hat{\theta} = \theta_0 + [\Phi' \Phi]^{-1} \Phi' \mathbf{v}$$

$$E\{\hat{\theta}\} = \theta_0 + E\{[\Phi' \Phi]^{-1} \Phi' \mathbf{v}\}$$

Si $E\{[\Phi' \Phi]^{-1} \Phi' \mathbf{v}\} = 0$ la estima es no sesgada

$$E\{\hat{\theta}\} = \theta_0$$



Propiedades (2)

¿Cuando el término $E\{[\Phi'\Phi]^{-1}\Phi'\mathbf{v}\}$ es nulo?

$$E\{[\Phi'\Phi]^{-1}\Phi'\mathbf{v}\} = \left[\frac{1}{N} \sum_{t=1}^N \varphi(t)\varphi'(t) \right]^{-1} \left[\frac{1}{N} \sum_{t=1}^N \varphi(t)v(t) \right]$$

La inversa será no nula,
luego el término

$$\frac{1}{N} \sum_{t=1}^N \varphi(t)v(t) \approx R_{\varphi v}(0) \rightarrow 0$$

Para que la estima de θ no sea sesgada los datos $\varphi(t)$ no deben estar correlacionados con los ruidos $v(t)$



Caso Respuesta Impulso

$$y(t) = \sum_{j=1}^m h_{j0} u(t-j) + v(t)$$

$$\varphi(t)' = [u(t-1), u(t-2), \dots, u(t-m)]$$

$$\theta_0' = [h_{10}, h_{20}, \dots, h_{m0}]$$

Estimación no
sesgada
si u y v no están
correlacionados

$$E\{\hat{\theta}\} = \theta_0 + E\{[\Phi'\Phi]^{-1} \Phi'v\}$$

$$\Phi'v = \begin{bmatrix} u(0) & u(1) & u(2) & \dots & u(N-1) \\ u(-1) & u(0) & u(1) & \dots & u(N-2) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ u(1-m) & u(2-m) & u(3-m) & \dots & u(N-m) \end{bmatrix} \begin{bmatrix} v(1) \\ v(2) \\ v(3) \\ \vdots \\ v(N) \end{bmatrix} \propto \begin{bmatrix} R_{uv}(-1) \\ R_{uv}(-2) \\ \vdots \\ R_{uv}(-m) \end{bmatrix} = \begin{bmatrix} R_{vu}(1) \\ R_{vu}(2) \\ \vdots \\ R_{vu}(m) \end{bmatrix}$$



Caso Respuesta Impulso

$$y(t) = \sum_{j=1}^m h_{j0} u(t-j) + v(t)$$

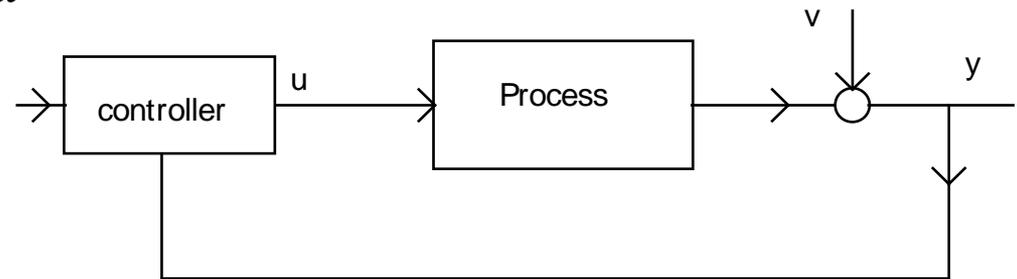
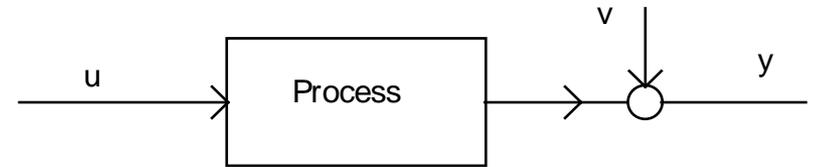
$$\varphi(t)' = [u(t-1), u(t-2), \dots, u(t-m)]$$

$$\theta_0' = [h_{10}, h_{20}, \dots, h_{m0}]$$

$$E\{\hat{\theta}\} = \theta_0 + E\{[\Phi' \Phi]^{-1} \Phi' \mathbf{v}\}$$

Estimación puede ser sesgada en lazo cerrado: v y u están correlacionadas a través de la y de realimentación

Estimación no sesgada en lazo abierto: u y v no correlacionados





Caso modelo de regresión

Representación exacta del proceso

$$y(t) = \frac{B_0(q^{-1})}{A_0(q^{-1})} u(t) + v(t)$$

$$A_0(q^{-1})y(t) = B_0(q^{-1})u(t) + A_0(q^{-1})v(t)$$

$$\varphi(t)' = [-y(t-1), \dots, -y(t-n), u(t-1), \dots, u(t-m)]$$

$$\theta_0' = [a_{10}, \dots, a_{n0}, b_{10}, \dots, b_{m0}]$$

$$y(t) = \varphi(t)' \theta_0 + A_0(q^{-1})v(t)$$

$$E\{\hat{\theta}\} = \theta_0 + E\{[\Phi' \Phi]^{-1} \Phi' A_0 v\}$$

En general los valores de Φ y $A_0 v$ estarán correlacionados pues la salida depende de v , y la estimación será sesgada



Caso modelo de regresión

$$\varphi(t)' = [-y(t-1), \dots, -y(t-n), u(t-1), \dots, u(t-m)]$$

$$\theta_0' = [a_{10}, \dots, a_{n0}, b_{10}, \dots, b_{m0}]$$

$$\Phi'v = \begin{bmatrix} -y(0) & -y(1) & -y(2) & \dots & -y(N-1) \\ -y(-1) & -y(0) & -y(1) & \dots & -y(N-2) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ -y(1-n) & -y(2-n) & -y(3-n) & \dots & -y(N-n) \\ u(0) & u(1) & u(2) & \dots & u(N-1) \\ u(-1) & u(0) & u(1) & \dots & u(N-2) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ u(1-m) & u(2-m) & u(3-m) & \dots & u(N-m) \end{bmatrix} \begin{bmatrix} A_0 v(1) \\ A_0 v(2) \\ A_0 v(3) \\ \vdots \\ A_0 v(N) \end{bmatrix} \propto \begin{bmatrix} -R_{yv}(-1) \\ -R_{yv}(-2) \\ \vdots \\ -R_{yv}(-n) \\ R_{uv}(-1) \\ R_{uv}(-2) \\ \vdots \\ R_{uv}(-m) \end{bmatrix}$$



Varianza de las estimaciones

Las estimaciones dependen de los datos experimentales y presentan una variabilidad

$$E\{\hat{\theta}\} = \theta_0 + E\{[\Phi'\Phi]^{-1}\Phi'v\}$$

$$\text{cov}\{\hat{\theta}\} = E\{(\hat{\theta} - \theta_0)(\hat{\theta} - \theta_0)'\} = [\Phi'\Phi]^{-1}\Phi'R_v\Phi'[\Phi'\Phi]^{-1}$$

La varianza de las estimaciones depende del ruido y de los datos experimentales

residuos: $e(t) = y(t) - \varphi(t)'\theta$

$$v(t) = y(t) - \varphi(t)'\theta_0$$

Si $\hat{\theta} \rightarrow \theta_0$ $e(t) = v(t)$



Interpretación en frecuencia (1)

$$V = \frac{1}{N} \sum_{t=1}^N \left[y(t) - \sum_{i=1}^n h_i u(t-i) \right]^2 \quad \text{Modelo respuesta impulso}$$

$$\begin{aligned} V &= \frac{1}{N} \sum_{t=1}^N \left[\sum_{k=1}^{\infty} h_{k0} q^{-k} u(t) + v(t) - \sum_{i=1}^n h_i q^{-i} u(t) \right]^2 = \\ &= \frac{1}{N} \sum_{t=1}^N \left[\left(\sum_{k=1}^{\infty} h_{k0} q^{-k} - \sum_{i=1}^n h_i q^{-i} \right) u(t) + v(t) \right]^2 \end{aligned}$$

Datos tomados en lazo abierto: u y v están incorrelacionados

$$V = \frac{1}{N} \sum_{t=1}^N \left(\sum_{k=1}^{\infty} h_{k0} q^{-k} - \sum_{i=1}^n h_i q^{-i} \right)^2 u(t)^2 + \frac{1}{N} \sum_{t=1}^N [v(t)]^2$$



Dominio de la frecuencia (2)

$$V = \frac{1}{N} \sum_{t=1}^N \left(\sum_{k=1}^{\infty} h_{k0} q^{-k} - \sum_{i=1}^n h_i q^{-i} \right)^2 u(t)^2 + \frac{1}{N} \sum_{t=1}^N [v(t)]^2$$

Igualdad de Parserval

$$\int_0^{NT} x(t)^2 dt = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \Phi_x(\omega) d\omega$$

$$V = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \left(\sum_{k=1}^{\infty} h_{k0} e^{-jk\omega T} - \sum_{i=1}^n h_i e^{-ji\omega T} \right)^2 \Phi_u(\omega) d\omega + \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \Phi_v(\omega) d\omega$$

Los errores están pesados en cada frecuencia por el espectro de potencia de los datos de entrada. El modelo, a las frecuencias que no se exciten presentará mayor error que aquellas donde $\Phi_u(\omega)$ sea significativo



Dominio de la frecuencia (3)

Caso modelo de regresión

$$y(t) = \frac{B_0(q^{-1})}{A_0(q^{-1})} u(t) + v(t)$$

$$V = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \left(\frac{B_0(e^{-j\omega T})}{A_0(e^{-j\omega T})} - \frac{B(e^{-j\omega T})}{A(e^{-j\omega T})} \right)^2 A(e^{-j\omega T})^2 \Phi_u(\omega) d\omega +$$
$$+ \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} A(e^{-j\omega T})^2 \Phi_v(\omega) d\omega$$

Los errores están pesados en cada frecuencia no solo por $\Phi_u(\omega)$, sino también por $A(\omega)$. Como normalmente A es de tipo pasa alta, se tienden a favorecer los ajustes de alta frecuencia



Experimentos

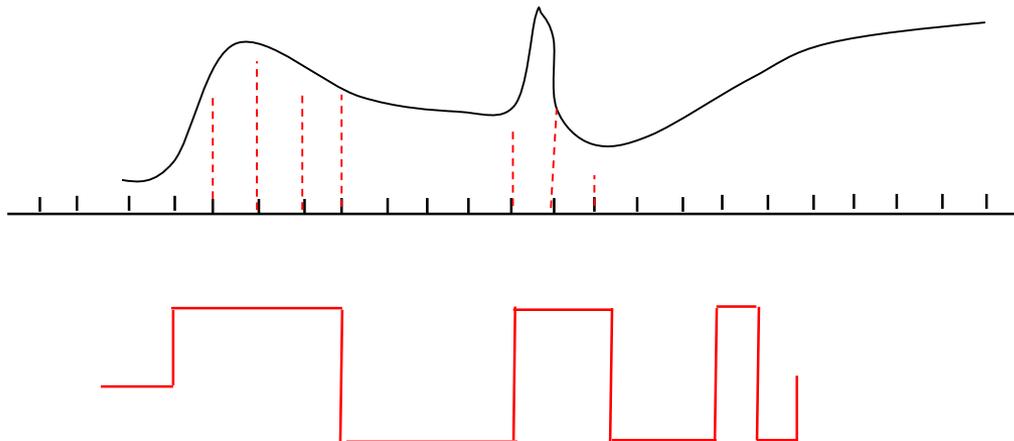
- ✓ Si se quiere determinar bien la dinámica de un proceso, los datos deben tener información sobre dicha dinámica
- ✓ Por tanto deben realizarse experimentos que correspondan a respuestas del sistema en distintas situaciones de funcionamiento
- ✓ Los registros históricos de variables reguladas, pueden no ser válidos al no tener información dinámica suficiente o estar correlacionada con el ruido.

Ensayos previos, tales como un salto a la entrada, pueden ayudar a determinar tiempos de asentamiento, amplitud adecuada de las señales, duración del experimento,.... (pre-test)



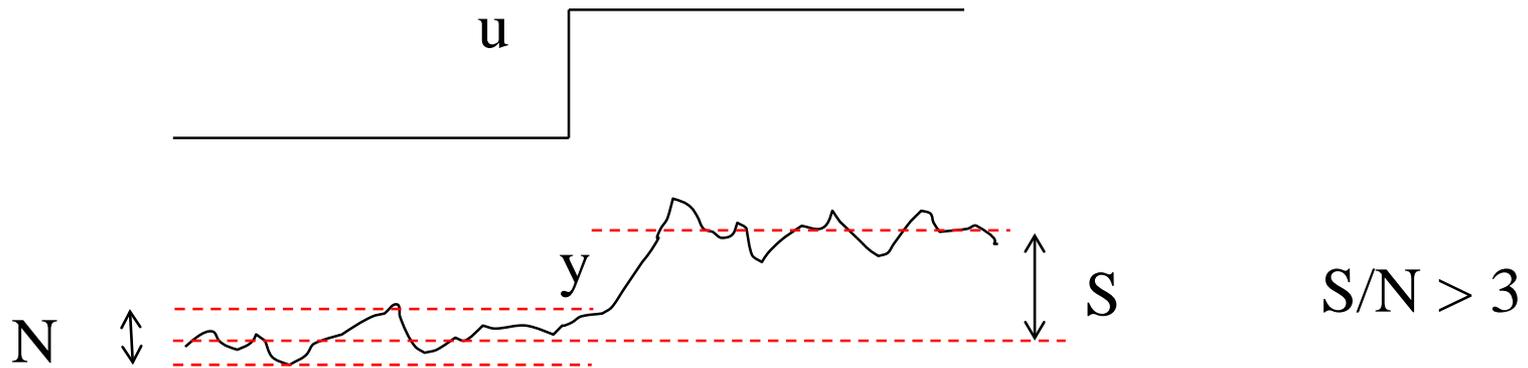
Experimentos, muestreo

- ✓ El periodo de muestreo debe fijarse en función del uso y la dinámica del proceso considerado. Para control, en función del tiempo de asentamiento deseado en lazo cerrado. La toma de datos suele ser mas frecuente que su uso posterior.
- ✓ La duración del experimento debe cubrir tanto datos para identificación como datos para validación (1000 →)





Experimentos, amplitud

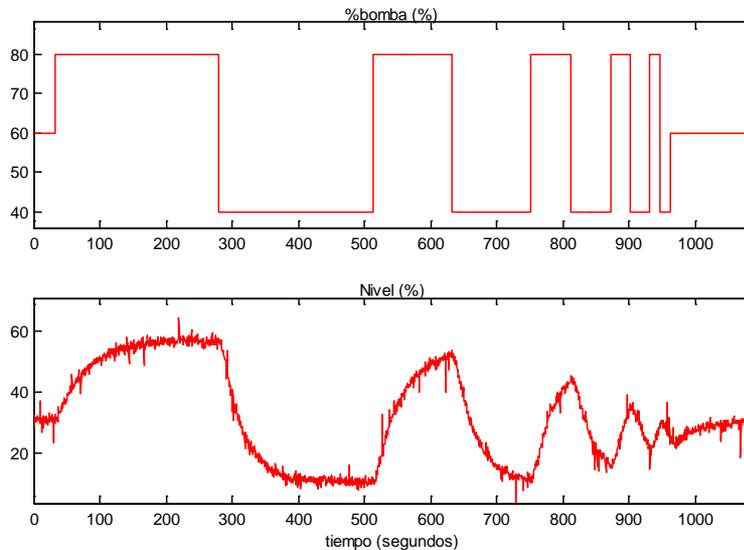


Dado que se están identificando modelos lineales, el cambio en las variables manipuladas debe ser lo mas pequeño posible para no alejarse de la zona de operación, pero también, si hay ruido, hay que mantener una adecuada relación señal/ruido, típicamente >3



Experimentos, frecuencia

Los experimentos deben hacerse con señales que cubran el rango de frecuencias de interés



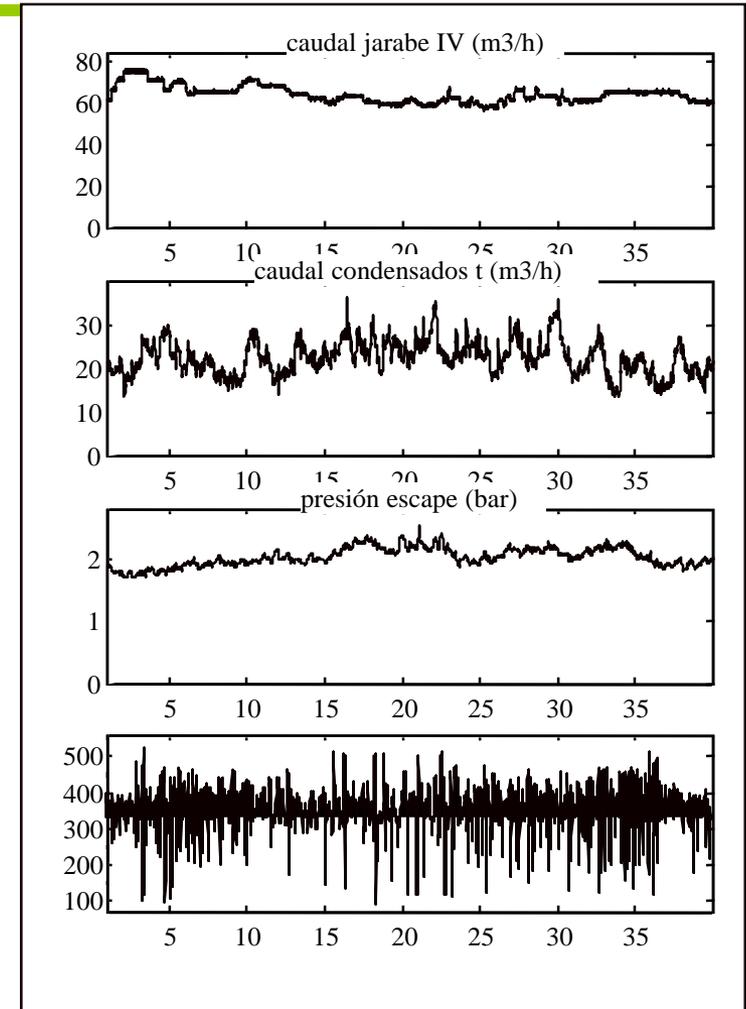
Señales típicas que cubren las frecuencias de interés

Los experimentos deben diseñarse cuidadosamente fuera de línea y ejecutar lo planificado para evitar realimentaciones a través del ejecutor del experimento



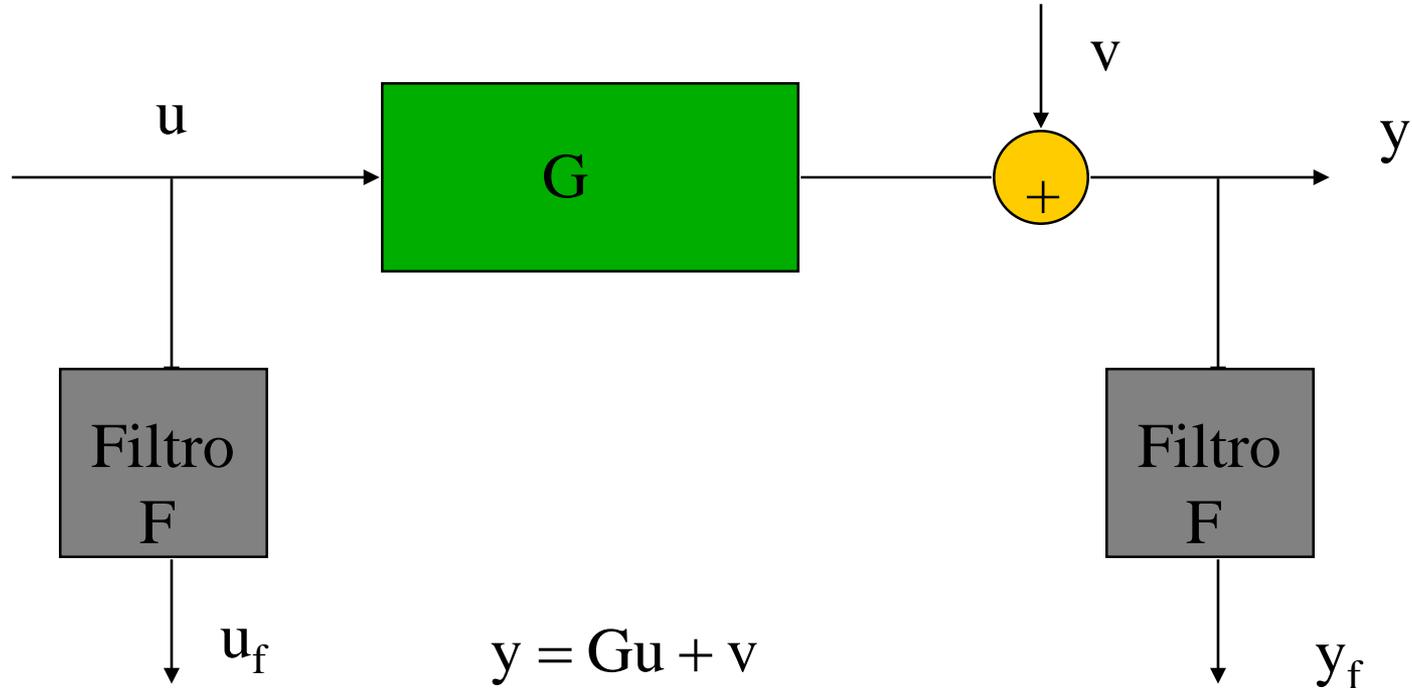
Experimentos

- ✓ Procesos tipo mecánico, electrónico,..
 - Cualquier ensayo en un tiempo corto
- ✓ Procesos tipo químico, biológico,..
 - Pueden hacerse durante mas tiempo pero sin perturbar demasiado
- ✓ **Planificación cuidadosa off-line y ejecución de lo planificado.**
- ✓ Experimentos para identificación y para validación
- ✓ El método de identificación puede imponer el tipo de experimento





Filtrado paralelo



$$y = Gu + v$$

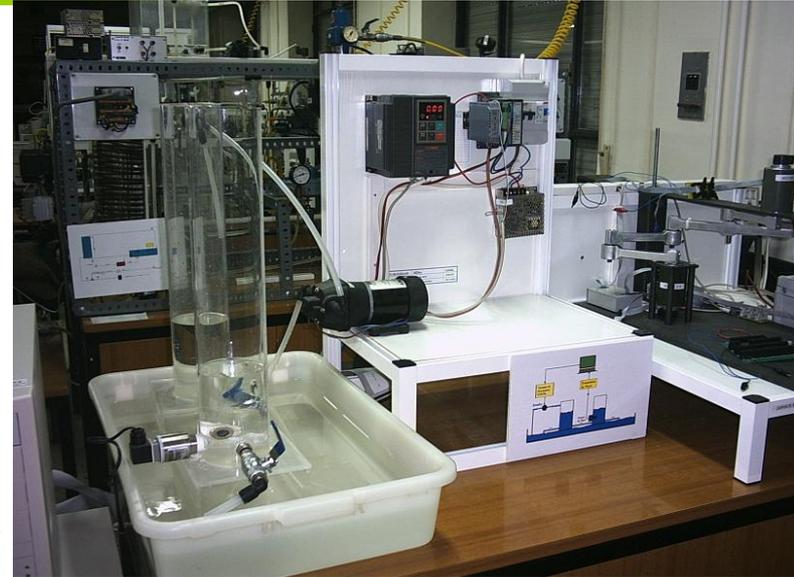
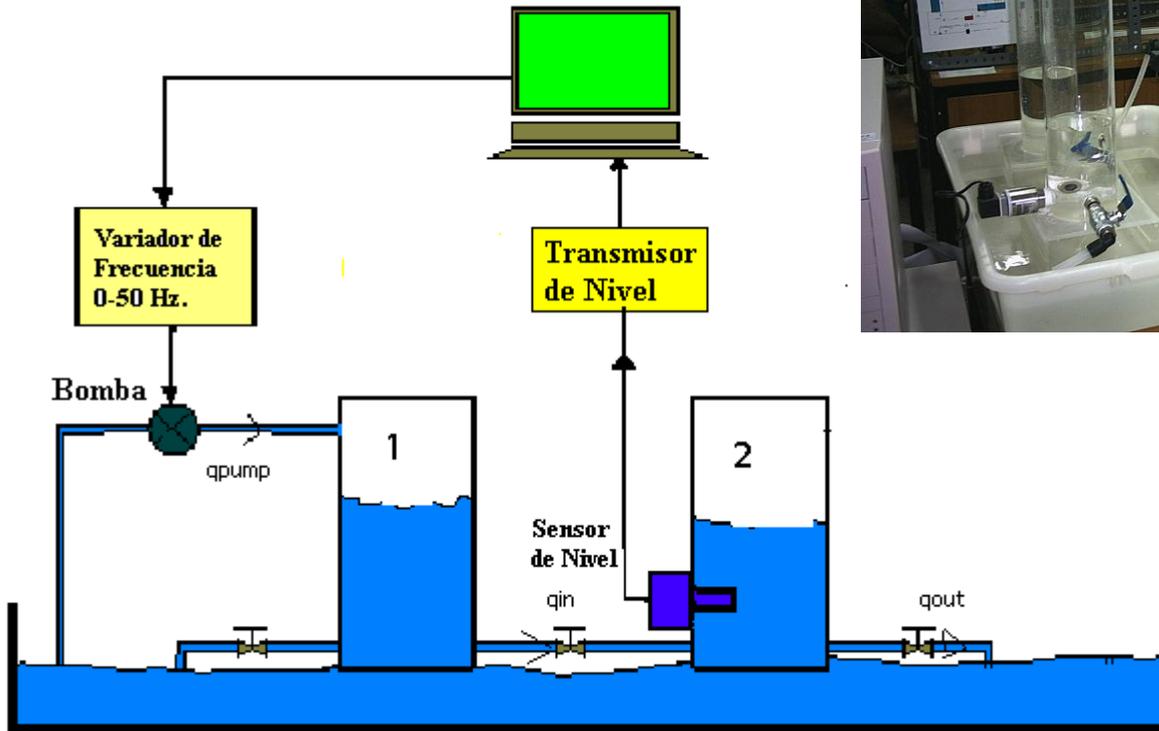
$$y_f = Fy = FG u + Fv = GF u + Fv$$

$$y_f = G u_f + Fv$$

Se conserva la dinámica
entre variables filtradas

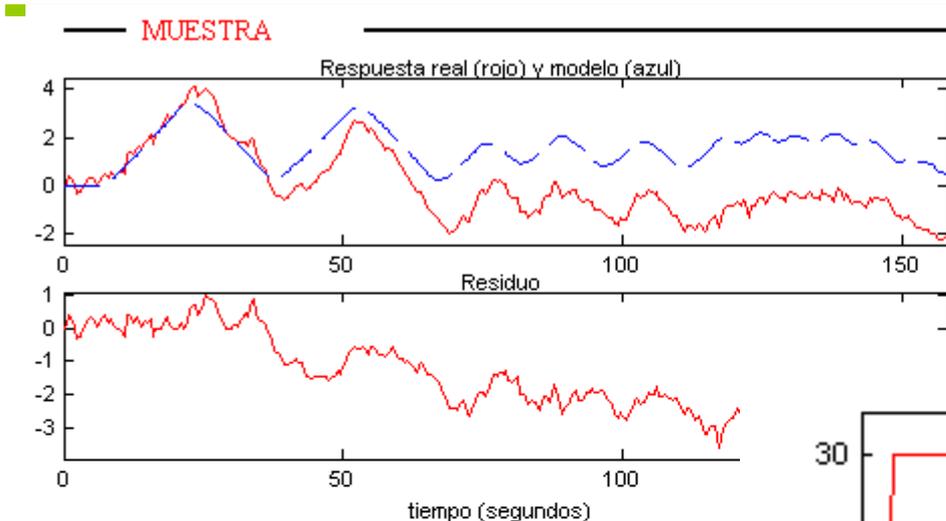


Ejemplo: Dos depósitos interconectados





Datos experimentales



- Aceptar
- Cancelar
- Identificar
- Guardar
- Ayuda

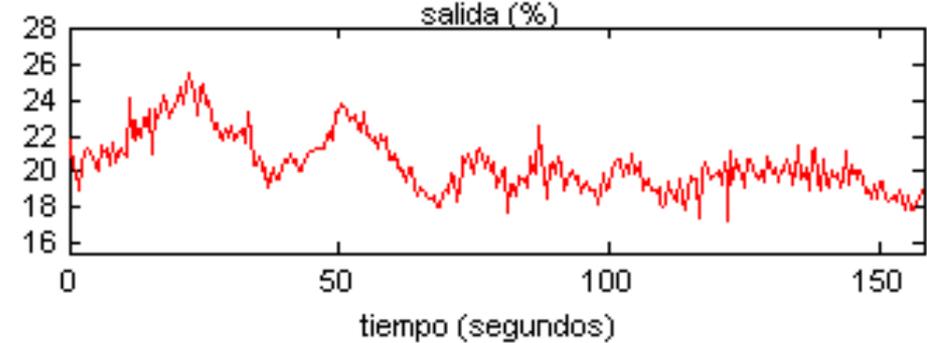
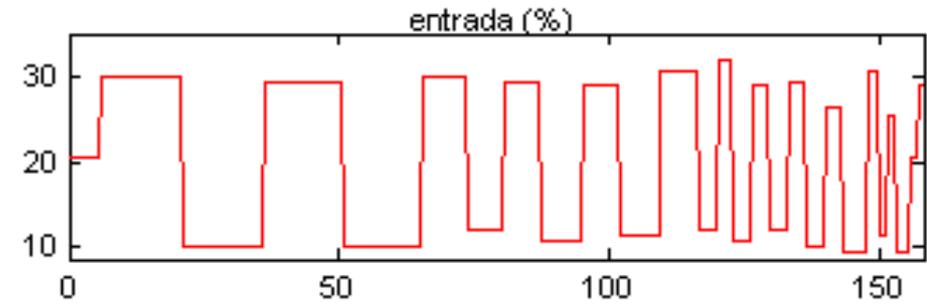
Poca
excitación a
bajas
frecuencias

Salvar gráficas al salir

RESULTADOS

salida / entrada:

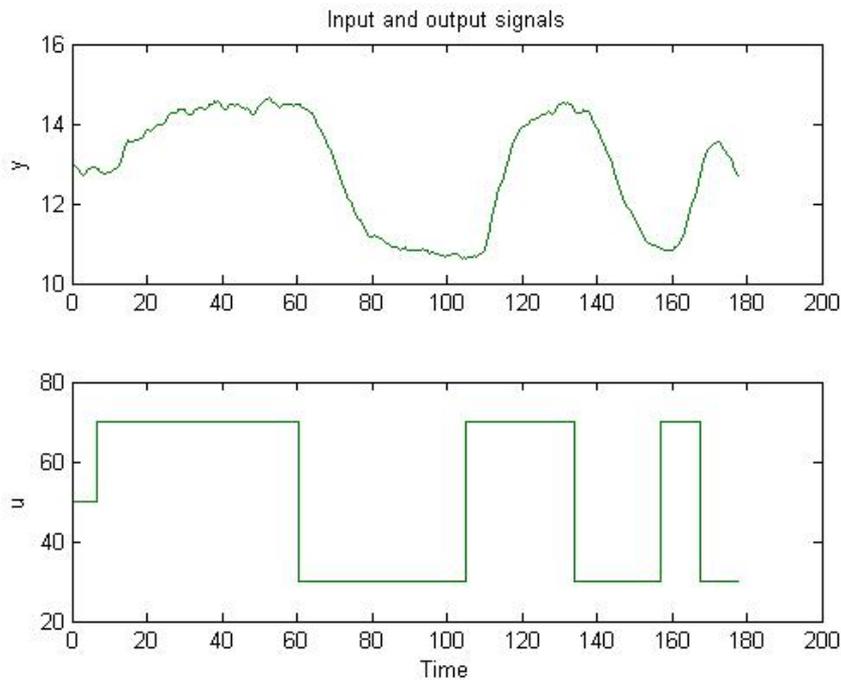
num1 =	0 0.012096
den1 =	1 -1.0278 0.026311
ganancial =	-8.3161
ceros1 =	
polos1 =	1.0015 0.026272



Identificar: calcula el modelo según el método elegido
Guardar: salva en un archivo tipo mat el modelo calculado

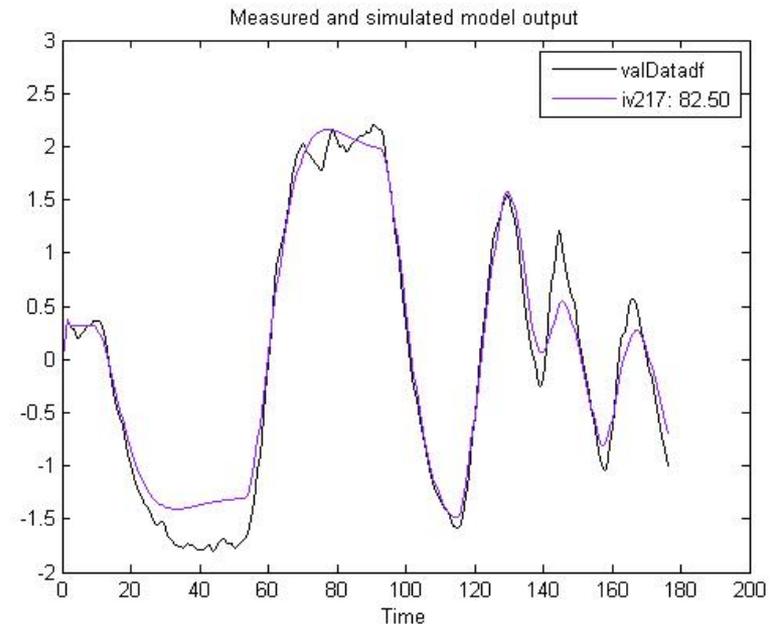


Datos experimentales



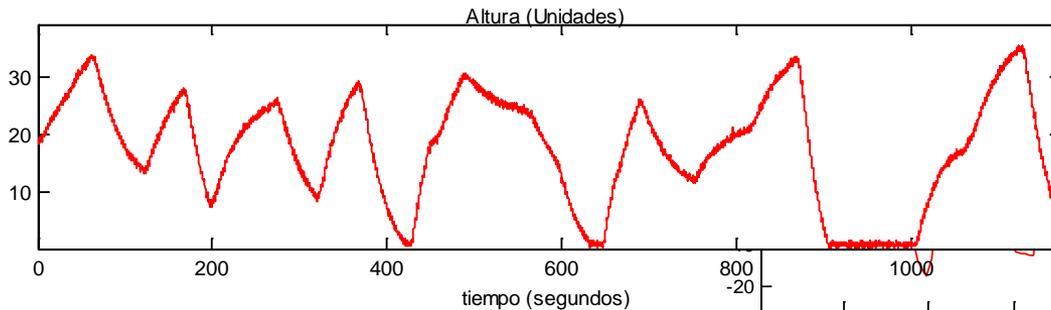
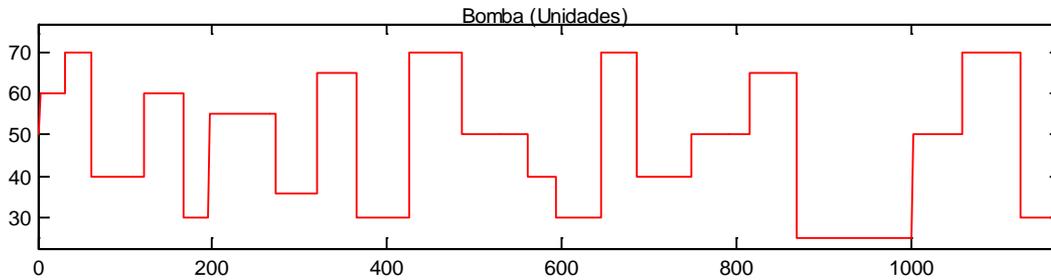
Datos identificación
con excesiva amplitud
30-70% :no linealidad

Datos validación:
respuesta modelo
/proceso error en
ganancia

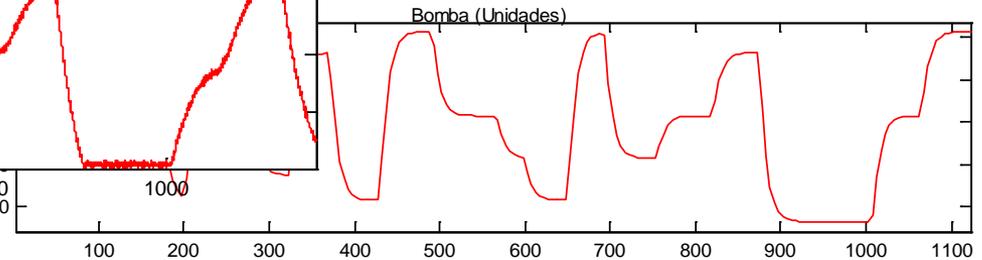




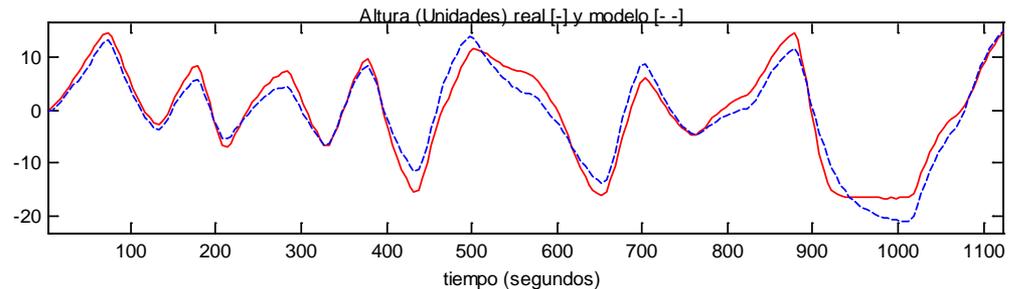
Datos experimentales



bolImpulsionalLargo01.mat



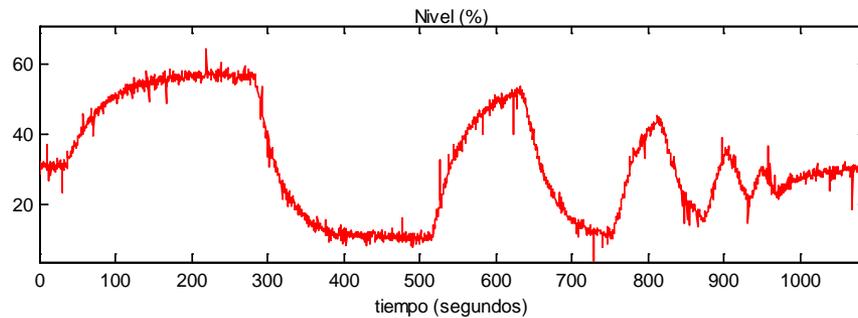
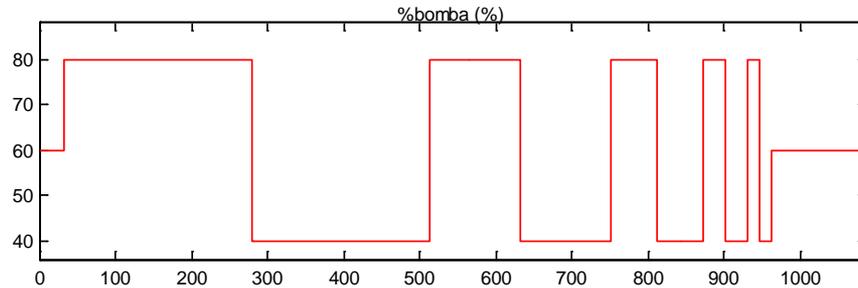
Datos de validación
con un tramo en
que el nivel esta
saturado a cero



Errores en esa
zona de
validación

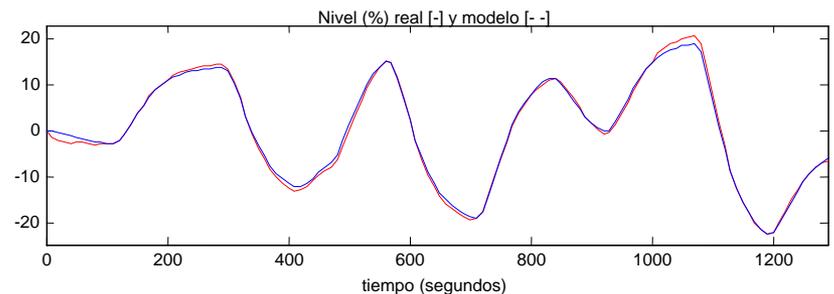
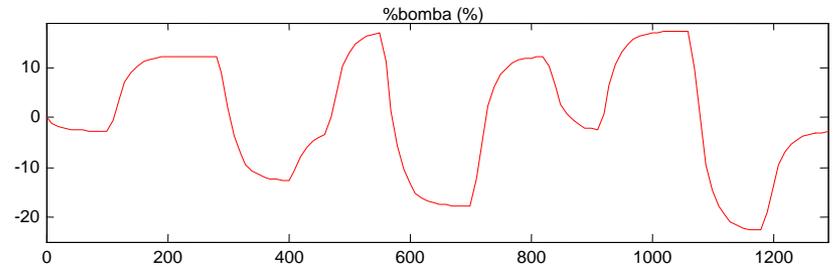


Cubriendo el rango de frecuencias de interés



Datos identificación

Datos validación
+respuesta modelo



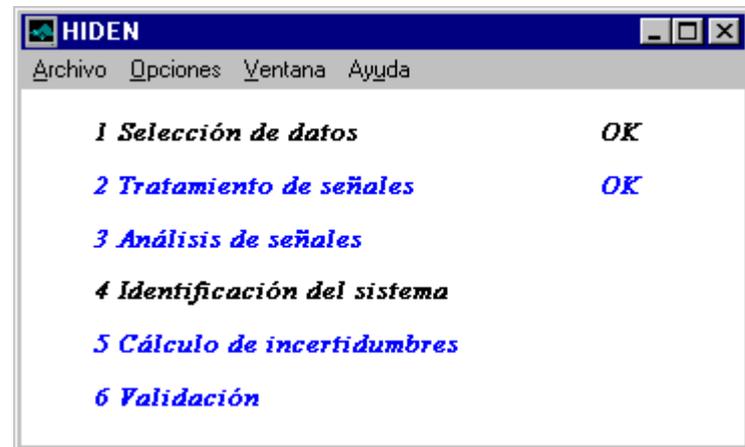


HIDEN

- ✓ Herramienta gráfica orientada a facilitar la identificación de modelos dinámicos lineales multivariados.
- ✓ Toolbox de Matlab



Implementa una metodología de identificación y ayuda al usuario a seguirla





RLS

Recursive Least Squares

RLS Mínimos cuadrados recursivos

$$\hat{\theta}(t) = [\Phi(t)' \Phi(t)]^{-1} \Phi(t)' \mathbf{y}(t) \quad \mathbf{P}(t) = [\Phi(t)' \Phi(t)]^{-1}$$

Inconvenientes: Memoria, tiempo de cálculo, fuera de línea,... Forma recursiva de procesar los datos

$$\hat{\theta}(t) = \mathbf{P}(t) \Phi(t)' \mathbf{y}(t)$$

$$\hat{\theta}(t-1) = \mathbf{P}(t-1) \Phi(t-1)' \mathbf{y}(t-1)$$

$$\hat{\theta}(t) - \hat{\theta}(t-1) = \mathbf{P}(t) \Phi(t)' \mathbf{y}(t) - \mathbf{P}(t-1) \Phi(t-1)' \mathbf{y}(t-1)$$

Operando:



RLS

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mathbf{K}(t)\varphi(t)[y(t) - \varphi(t)'\hat{\theta}(t-1)]$$

$$\mathbf{K}(t) = \mathbf{P}(t-1)[1 + \varphi(t)'\mathbf{P}(t-1)\varphi(t)]^{-1}$$

$$\mathbf{P}(t) = \mathbf{P}(t-1) - \mathbf{K}(t)\varphi(t)\varphi(t)'\mathbf{P}(t-1)$$

La matriz $\mathbf{P}(t)$ controla la velocidad de corrección
Típicamente es decreciente:

$$\mathbf{P}(t)^{-1} = \mathbf{P}(t-1)^{-1} + \varphi(t)\varphi(t)'$$

Tras un cierto tiempo no hay corrección significativa
Poco adecuado para identificación en tiempo real



Factor de olvido

$$V = \sum_{t=1}^N \lambda^{N-t} [y(t) - \varphi(t)' \theta]^2$$

λ es el factor de olvido 0.95 ... 0.999

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mathbf{K}(t) \varphi(t) [y(t) - \varphi(t)' \hat{\theta}(t-1)]$$

$$\mathbf{K}(t) = \mathbf{P}(t-1) [\lambda + \varphi(t)' \mathbf{P}(t-1) \varphi(t)]^{-1}$$

$$\mathbf{P}(t) = [\mathbf{P}(t-1) - \mathbf{K}(t) \varphi(t) \varphi(t)' \mathbf{P}(t-1)] / \lambda$$

Problemas de “explosión” con baja excitación

$$\mathbf{P}(t) \varphi(t) \rightarrow 0 \quad \Rightarrow \quad \mathbf{P}(t) = [\mathbf{P}(t-1)] / \lambda$$

Modificaciones para evitar la explosión: métodos de traza de \mathbf{P} constante, factor de olvido variables, etc.



VARIABLES INSTRUMENTALES IV

La estimación LS es sesgada si el ruido \mathbf{v} está correlacionado con Φ

$$\hat{\theta} = [\Phi' \Phi]^{-1} \Phi' \mathbf{y}$$

$$E\{\hat{\theta}\} = \theta_0 + E\{[\Phi' \Phi]^{-1} \Phi' \mathbf{v}\}$$

El método IV usa la fórmula de estimación:

$$\hat{\theta} = [\mathbf{W}' \Phi]^{-1} \mathbf{W}' \mathbf{y}$$

Con \mathbf{W} escogida incorrelacionada con \mathbf{v} , y $\mathbf{W}' \Phi$ invertible

$$E\{\hat{\theta}\} = E\{[\mathbf{W}' \Phi]^{-1} \mathbf{W}' \mathbf{y}\} = E\{[\mathbf{W}' \Phi]^{-1} \mathbf{W}' (\Phi \theta_0 + \mathbf{v})\} = \theta_0 + E\{[\mathbf{W}' \Phi]^{-1} \mathbf{W}' \mathbf{v}\} = \theta_0$$

La estimación IV es no sesgada para $N \rightarrow \infty$
Consistente



VARIABLES INSTRUMENTALES IV

$$\hat{\theta} = [W' \Phi]^{-1} W' y$$

Existen varias formas de escoger el vector w de variables instrumentales

$$\Phi' = [\varphi(1), \varphi(2), \dots, \varphi(N)]$$

$$W' = [z(1), z(2), \dots, z(N)]$$

$$z(t) = \frac{\hat{B}(q^{-1})}{\hat{A}(q^{-1})} u(t) \quad \text{o} \quad z(t) = [u(t-1), \dots, u(t-1), \dots, u(t-2N)]$$

Puede ser necesario iterar para obtener mejores $z(t)$

Hay muchas elecciones equivalentes: TW con T invertible .



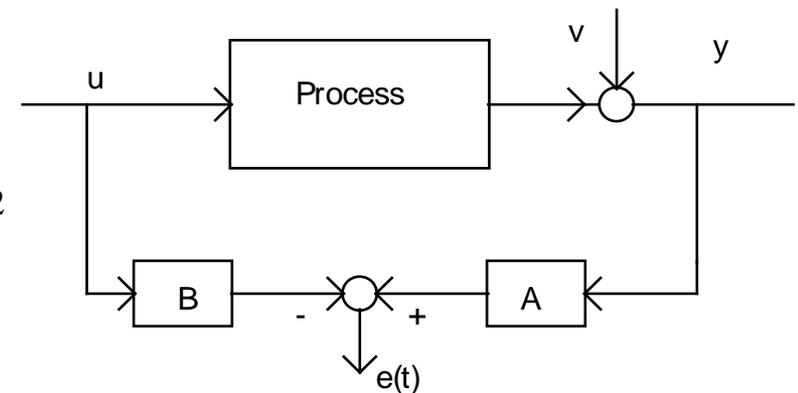
Modelo de regresión

$$y_m(t) = -a_1 y(t-1) - a_2 y(t-2) - \dots - a_n y(t-n) \\ + b_1 u(t-1) + \dots + b_m u(t-m)$$

$$e(t) = y(t) - y_m(t) = \\ = y(t) + a_1 y(t-1) + a_2 y(t-2) + \dots + a_n y(t-n) - \\ - b_1 u(t-1) - \dots - b_m u(t-m) \\ = A(q^{-1})y(t) - B(q^{-1})u(t)$$

Función de costo

$$V = \frac{1}{N} \sum_{t=1}^N [A(q^{-1})y(t) - B(q^{-1})u(t)]^2$$





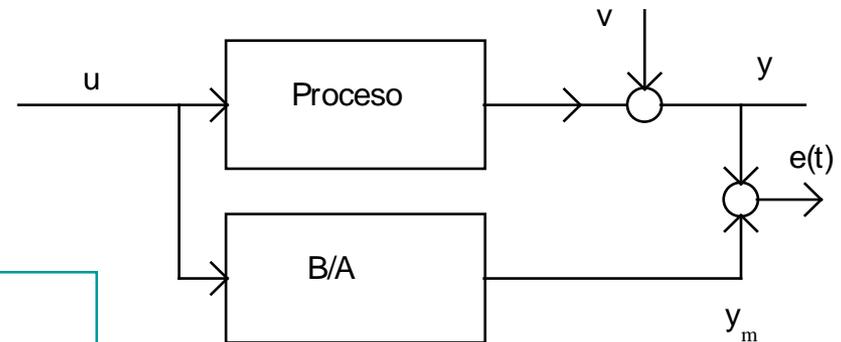
Método del Error de la salida

Modelo $y_m(t) = \frac{B(q^{-1})}{A(q^{-1})} u(t)$

Output Error (OE)

Modelo independiente

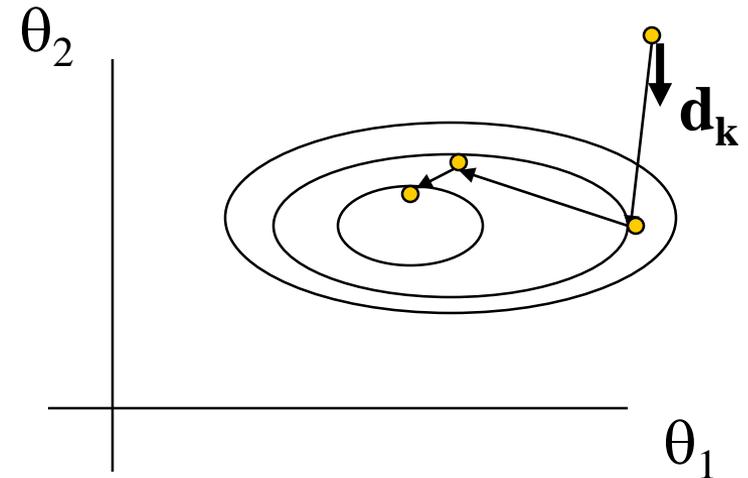
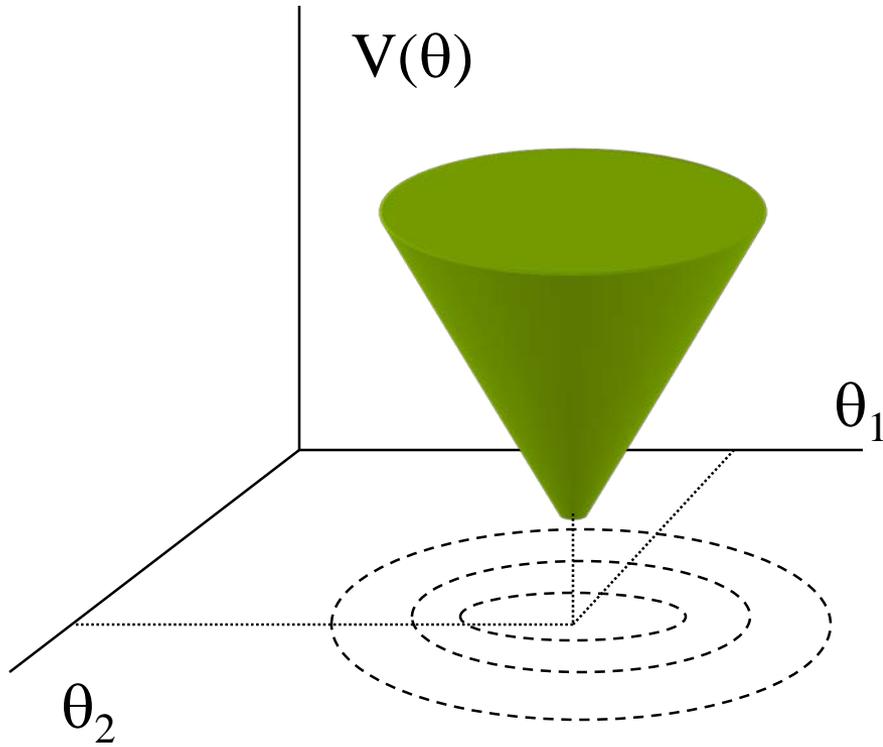
$$V = \frac{1}{N} \sum_{t=1}^N e(t)^2 = \frac{1}{N} \sum_{t=1}^N [y(t) - y_m(t)]^2 =$$
$$= \frac{1}{N} \sum_{t=1}^N \left[y(t) - \frac{B(q^{-1})}{A(q^{-1})} u(t) \right]^2$$



No-lineal en los parámetros.
Ha de resolverse por minimización numérica



Minimización de funciones



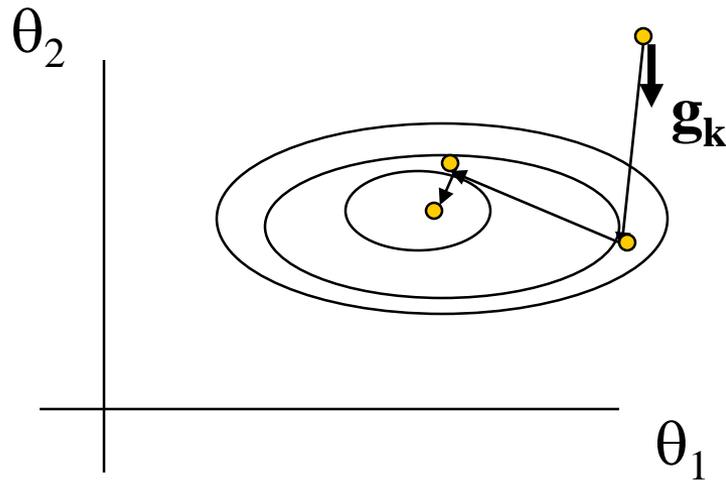
$$\theta_{k+1} = \theta_k + \alpha \mathbf{d}_k$$

Métodos iterativos

\mathbf{d} dirección de búsqueda
 α longitud del paso



Minimización de funciones



Método de Newton-Raphson
Método simplex

$$\theta_{k+1} = \theta_k + \alpha \mathbf{d}_k$$

Método del gradiente
mas pronunciado

$$\mathbf{d}_k = \mathbf{g}_k = \left. \frac{\partial V}{\partial \boldsymbol{\theta}} \right|_{\theta_k}$$

$$\alpha_k = \min_{\alpha} V(\theta_k + \alpha \mathbf{g}_k(\theta_k))$$

$$\theta_{k+1} = \theta_k + \alpha_k \mathbf{g}_k(\theta_k)$$



Algoritmo de Gauss-Newton

Fundamento:

- 1 Partir de una estima inicial θ_k
- 2 Linealizar el error e sobre θ_k
- 3 Aplicar LS al error linealizado para mejorar la estima θ_k
- 4 Iterar hasta que V no mejore

$$V = \frac{1}{N} \sum_{t=1}^N e(t)^2$$

$$e(t, \theta) \cong e(t, \hat{\theta}_k) + \left. \frac{\partial e(t, \theta)}{\partial \theta'} \right|_{\theta=\hat{\theta}_k} (\theta - \hat{\theta}_k) =$$

$$= e(t, \hat{\theta}_k) + \varphi_k(t)(\theta - \hat{\theta}_k) = [e(t, \hat{\theta}_k) - \varphi_k(t)\hat{\theta}_k] + \varphi_k(t)\theta$$

$$\varphi_k(t) = \left. \frac{\partial e(t, \theta)}{\partial \theta'} \right|_{\theta=\hat{\theta}_k} =$$

$$= \left[\frac{\hat{B}_k(q^{-1})}{\hat{A}_k(q^{-1})^2} u(t-1), \dots, \frac{\hat{B}_k(q^{-1})}{\hat{A}_k(q^{-1})^2} u(t-n), \frac{-1}{\hat{A}_k(q^{-1})} u(t-1), \dots, \frac{-1}{\hat{A}_k(q^{-1})} u(t-m) \right]$$



Algoritmo de Gauss-Newton

$$e(t, \theta) \cong [e(t, \hat{\theta}_k) - \varphi_k(t)\hat{\theta}_k] + \varphi_k(t)\theta$$

En LS: $e(t) = y(t) - \varphi(t)'\theta$

$[e(t, \hat{\theta}_k) - \varphi_k(t)\hat{\theta}_k]$ juega el papel de $y(t)$

$\varphi_k(t)$ juega el papel de $\varphi(t)$

LS:
$$\hat{\theta} = [\Phi'\Phi]^{-1}\Phi'y = \left[\sum_{t=1}^N \varphi(t)'\varphi(t) \right]^{-1} \sum_{t=1}^N \varphi(t)'y(t)$$

Gauss-Newton:

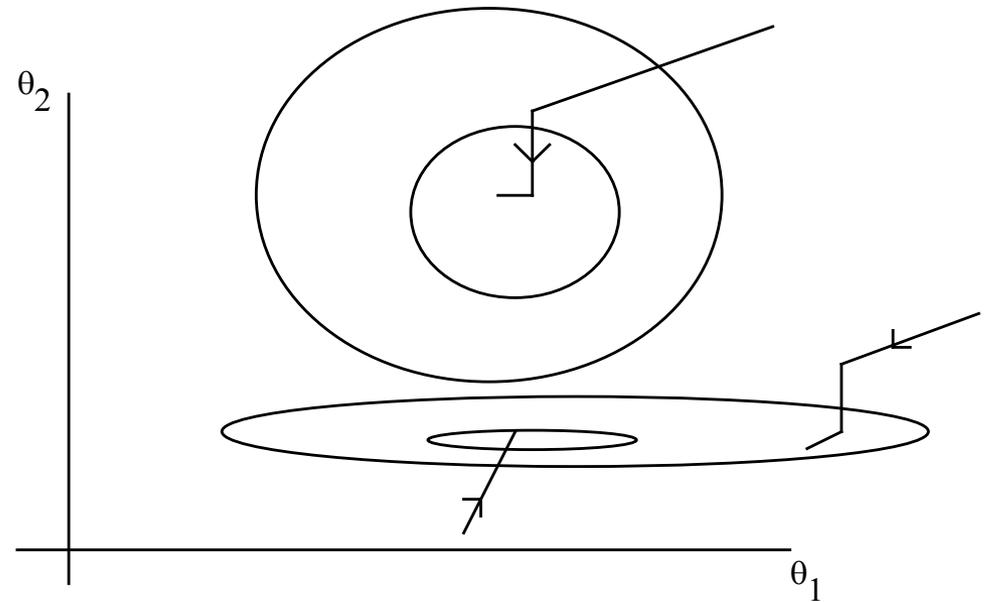
$$\begin{aligned} \hat{\theta}_{k+1} &= \left[\sum_{t=1}^N \varphi_k(t)'\varphi_k(t) \right]^{-1} \sum_{t=1}^N \varphi_k(t)' [e(t, \hat{\theta}_k) - \varphi_k(t)\hat{\theta}_k] = \\ &= \hat{\theta}_k - \left[\sum_{t=1}^N \varphi_k(t)'\varphi_k(t) \right]^{-1} \sum_{t=1}^N \varphi_k(t)' e(t, \hat{\theta}_k) \end{aligned}$$



Error de la salida

$$\begin{aligned} V &= \frac{1}{N} \sum_{t=1}^N e(t)^2 = \frac{1}{N} \sum_{t=1}^N [y(t) - y_m(t)]^2 = \\ &= \frac{1}{N} \sum_{t=1}^N \left[y(t) - \frac{B(q^{-1})}{A(q^{-1})} u(t) \right]^2 \end{aligned}$$

La forma de los contornos depende de los datos experimentales, y por tanto la calidad de la identificación





Propiedades OE (1)

Representación exacta $y(t) = \frac{B_0(q^{-1})}{A_0(q^{-1})}u(t) + v(t)$

$$V = \frac{1}{N} \sum_{t=1}^N e(t)^2 = \frac{1}{N} \sum_{t=1}^N [y(t) - y_m(t)]^2$$

$$\begin{aligned} V &= \frac{1}{N} \sum_{t=1}^N \left[\frac{B_0(q^{-1})}{A_0(q^{-1})}u(t) + v(t) - \frac{B(q^{-1})}{A(q^{-1})}u(t) \right]^2 = \\ &= \frac{1}{N} \sum_{t=1}^N \left[\left(\frac{B_0(q^{-1})}{A_0(q^{-1})} - \frac{B(q^{-1})}{A(q^{-1})} \right) u(t) + v(t) \right]^2 \end{aligned}$$



Propiedades OE (2)

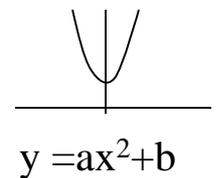
$$V = \frac{1}{N} \sum_{t=1}^N \left[\left(\frac{B_0(q^{-1})}{A_0(q^{-1})} - \frac{B(q^{-1})}{A(q^{-1})} \right) u(t) + v(t) \right]^2$$

Si los datos se toman en **lazo abierto** de modo que u y v estén incorrelacionados, y si N tiende a infinito:

$$E\{V\} = \left(\frac{B_0(q^{-1})}{A_0(q^{-1})} - \frac{B(q^{-1})}{A(q^{-1})} \right)^2 E\{u(t)^2\} + E\{v(t)^2\}$$

El mínimo alcanza el verdadero valor de los parámetros y la estimación es consistente

$$\left(\frac{B_0(q^{-1})}{A_0(q^{-1})} - \frac{B(q^{-1})}{A(q^{-1})} \right) \rightarrow 0$$



$$y = ax^2 + b$$



Dominio frecuencial

$$E\{V\} = \left(\frac{B_0(q^{-1})}{A_0(q^{-1})} - \frac{B(q^{-1})}{A(q^{-1})} \right)^2 E\{u(t)^2\} + E\{v(t)^2\}$$

Utilizando el Teorema de Parserval:

$$V = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \left(\frac{B_0(e^{-j\omega T})}{A_0(e^{-j\omega T})} - \frac{B(e^{-j\omega T})}{A(e^{-j\omega T})} \right)^2 \Phi_u(\omega) d\omega \\ + \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \Phi_v(\omega) d\omega$$

La estimación se mejora en el rango de frecuencias donde el espectro de potencia de u es grande



Modelo de ruido

$$v(t) = \frac{C(q^{-1})}{A(q^{-1})} \xi(t)$$

Proceso ARMA

$\xi(t)$ ruido blanco de media nula

Las características estadísticas de $v(t)$ dependen de la función de transferencia C/A

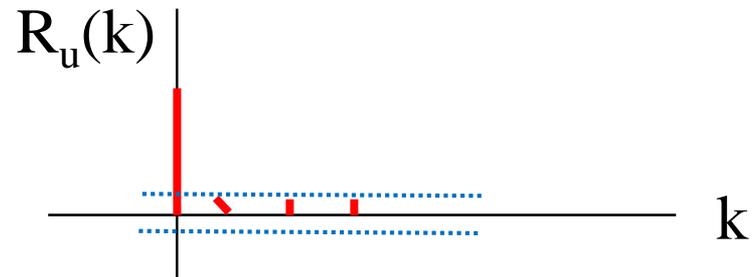
Alternativamente puede suponerse un ruido $v(t)$ sin estructura de modelo predefinida, y estimarlo en función de las medidas experimentales.



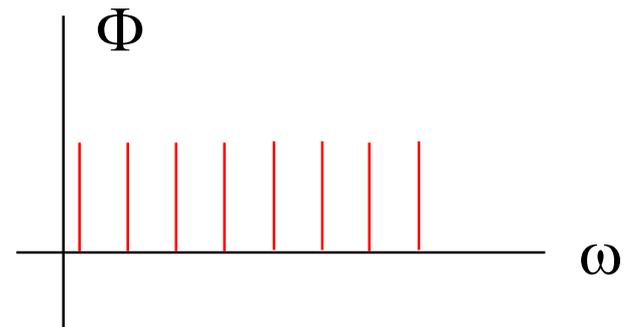
Ruido blanco

Señal aleatoria en la que sus valores en un instante t son independientes del valor que han tomado en instantes anteriores.

$$R_u(k) = 0 \quad k \neq 0$$



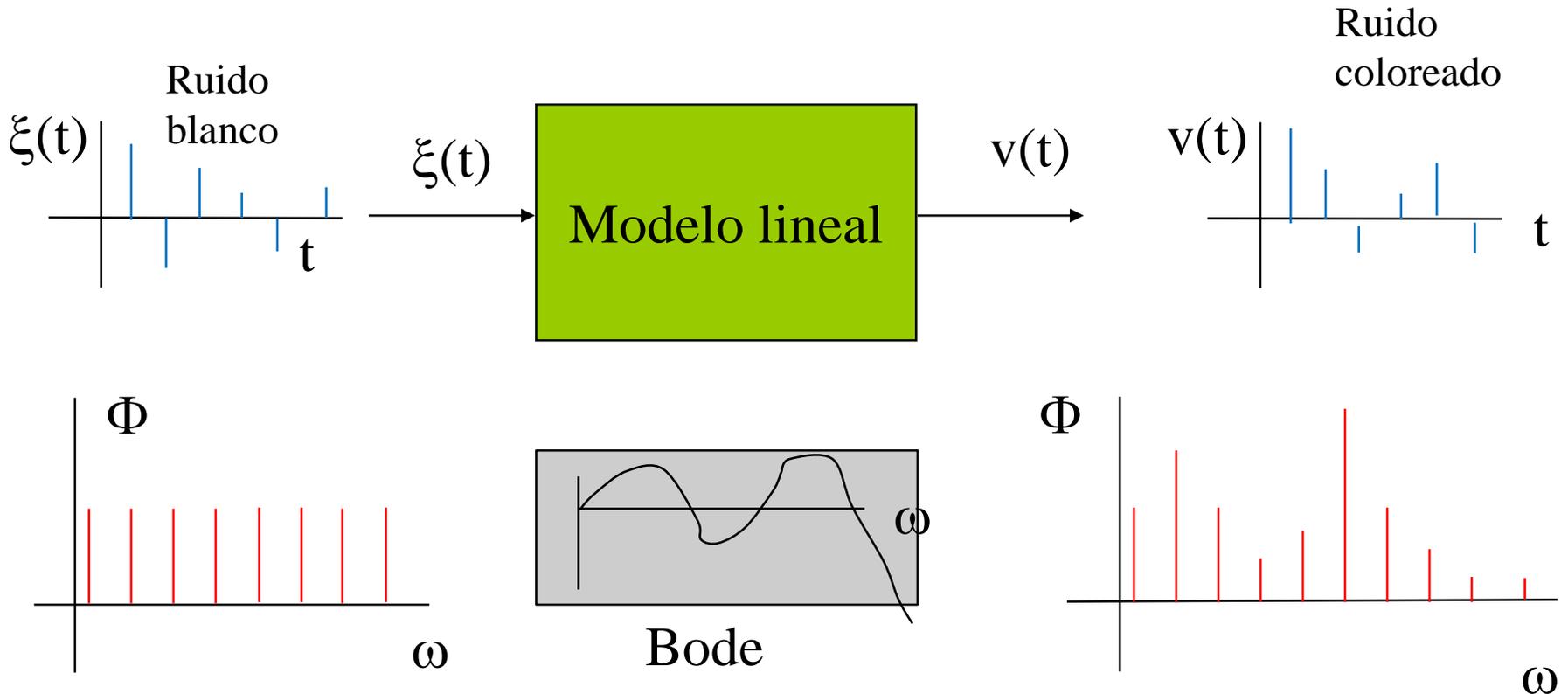
Su espectro de potencia es constante





Modelo de ruido

$$v(t) = \frac{C(q^{-1})}{A(q^{-1})} \xi(t) \quad \text{ARMA}$$





Métodos del Error de Predicción

Prediction Error Methods (PEM)

Se modelan explícitamente las perturbaciones

Modelo Box-Jenkins
$$y_m(t) = \frac{B(q^{-1})}{A(q^{-1})} u(t) + \frac{C(q^{-1})}{D(q^{-1})} \xi(t)$$

Modelo CARMA
ARMAX
$$y_m(t) = \frac{B(q^{-1})}{A(q^{-1})} u(t) + \frac{C(q^{-1})}{A(q^{-1})} \xi(t)$$

G y H funciones de
$$y_m(t) = G(q^{-1})u(t) + H(q^{-1})\xi(t)$$

transferencia,

H estable, mónica y
de fase mínima

Sobre el valor de y_m
solo pueden hacerse predicciones



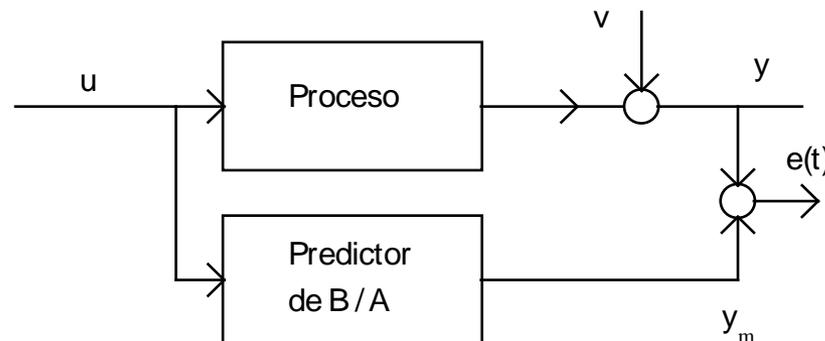
Métodos del Error de Predicción

Prediction Error Methods (PEM)

Criterio: Minimizar la suma de los cuadrados de los errores de predicción, respecto a los parámetros θ

$$\min_{\theta} V = \min_{\theta} \frac{1}{N} \sum_{t=1}^N \hat{e}(t|\theta)^2 = \min_{\theta} \frac{1}{N} \sum_{t=1}^N [y(t) - \hat{y}(t|\theta)]^2$$

$\hat{y}(t|\theta)$ Predicción de la salida del modelo con los parámetros θ





PEM: Predicciones

$$y_m(t) = G(q^{-1})u(t) + H(q^{-1})\xi(t)$$

$$y_m(t) = G(q^{-1})u(t) + [H(q^{-1}) - 1]\xi(t) + \xi(t)$$

$$\xi(t) = \frac{1}{H(q^{-1})} [y(t) - G(q^{-1})u(t)]$$

$$y_m(t) = G(q^{-1})u(t) + [H(q^{-1}) - 1] \frac{1}{H(q^{-1})} [y(t) - G(q^{-1})u(t)] + \xi(t) =$$

$$= G(q^{-1})u(t) + \left[1 - \frac{1}{H(q^{-1})} \right] y(t) - G(q^{-1})u(t) + \frac{G(q^{-1})}{H(q^{-1})} u(t) + \xi(t) =$$

$$= \frac{G(q^{-1})}{H(q^{-1})} u(t) + \left[1 - \frac{1}{H(q^{-1})} \right] y(t) + \xi(t)$$



PEM

Modelo lineal $y_m(t) = G(q^{-1})u(t) + H(q^{-1})\xi(t)$

Predicción óptima $\hat{y}(t|\theta) = \frac{G(q^{-1})}{H(q^{-1})}u(t) + \left[1 - \frac{1}{H(q^{-1})}\right]y(t)$

$$\hat{e}(t|\theta) = y(t) - \hat{y}(t|\theta) = y(t) - \frac{G(q^{-1})}{H(q^{-1})}u(t) - \left[1 - \frac{1}{H(q^{-1})}\right]y(t) =$$

$$= \frac{1}{H(q^{-1})} \left[y(t) - G(q^{-1})u(t) \right]$$

$$V = \frac{1}{N} \sum_{t=1}^N \left(\frac{1}{H(q^{-1})} \left[y(t) - G(q^{-1})u(t) \right] \right)^2$$

Debe resolverse
por minimización
numérica



PEM: Propiedades

Representación exacta $y(t) = G_0(q^{-1})u(t) + H_0(q^{-1})\xi(t)$

$$V = \frac{1}{N} \sum_{t=1}^N \left(\frac{1}{H(q^{-1})} [y(t) - G(q^{-1})u(t)] \right)^2 \quad \text{Si } N \rightarrow \infty$$

$$\begin{aligned} V &= E \left\{ \left(\frac{1}{H(q^{-1})} [G_0(q^{-1})u(t) + H_0(q^{-1})\xi(t) - G(q^{-1})u(t)] \right)^2 \right\} = \\ &= E \left\{ \left(\frac{G_0(q^{-1}) - G(q^{-1})}{H(q^{-1})} u(t) + \frac{H_0(q^{-1})}{H(q^{-1})} \xi(t) \right)^2 \right\} \end{aligned}$$



PEM: Propiedades

$$V = E \left\{ \left(\frac{G_0(q^{-1}) - G(q^{-1})}{H(q^{-1})} u(t) + \frac{H_0(q^{-1})}{H(q^{-1})} \xi(t) \right)^2 \right\}$$

Al ser H mónico, y G tener al menos un retardo:

$$\begin{aligned} V &= E \left\{ \left(\xi(t) + \text{a term independant of } \xi(t) \right)^2 \right\} = \\ &= E \left\{ \xi(t)^2 \right\} + E \left\{ \left(\text{term independant of } \xi(t) \right)^2 \right\} \geq E \left\{ \xi(t)^2 \right\} = \sigma_\xi^2 \end{aligned}$$

La varianza del ruido es una cota inferior que se alcanza si:

$$G_0(q^{-1}) - G(q^{-1}) \rightarrow 0 \quad \text{and} \quad \frac{H_0(q^{-1})}{H(q^{-1})} \rightarrow 1$$

Por tanto, el óptimo global proporciona una estimación no sesgada, incluso con datos tomados en lazo cerrado. Estimación consistente



PEM: dominio de la frecuencia

Si $N \rightarrow \infty$

$$V = E \left\{ \left(\frac{G_0(q^{-1}) - G(q^{-1})}{H(q^{-1})} u(t) + \frac{H_0(q^{-1})}{H(q^{-1})} \xi(t) \right)^2 \right\}$$

En **lazo abierto** u y el ruido están incorrelacionados, luego:

$$V = \frac{(G_0(q^{-1}) - G(q^{-1}))^2}{H(q^{-1})^2} E\{u(t)^2\} + \frac{H_0(q^{-1})^2}{H(q^{-1})^2} E\{\xi(t)^2\}$$



PEM: dominio de la frecuencia

$$V = \frac{(G_0(q^{-1}) - G(q^{-1}))^2}{H(q^{-1})^2} E\{u(t)^2\} + \frac{H_0(q^{-1})^2}{H(q^{-1})^2} E\{\xi(t)^2\}$$

Aplicando el teorema de Parserval:

$$V = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \left(\frac{B_0(e^{-j\omega T})}{A_0(e^{-j\omega T})} - \frac{B(e^{-j\omega T})}{A(e^{-j\omega T})} \right)^2 \frac{\Phi_u(\omega)}{H(e^{-j\omega T})} d\omega + \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \frac{H_0(e^{-j\omega T})}{H(e^{-j\omega T})} \Phi_v(\omega) d\omega$$

La estimación es mejor en aquel rango de frecuencias en el que $1/H$ y la señal de excitación son grandes



Residuos

Representación exacta $y(t) = G_0(q^{-1})u(t) + H_0(q^{-1})\xi(t)$

$$\xi(t) = \frac{1}{H_0(q^{-1})} [y(t) - G_0(q^{-1})u(t)] \quad \text{ruido blanco}$$

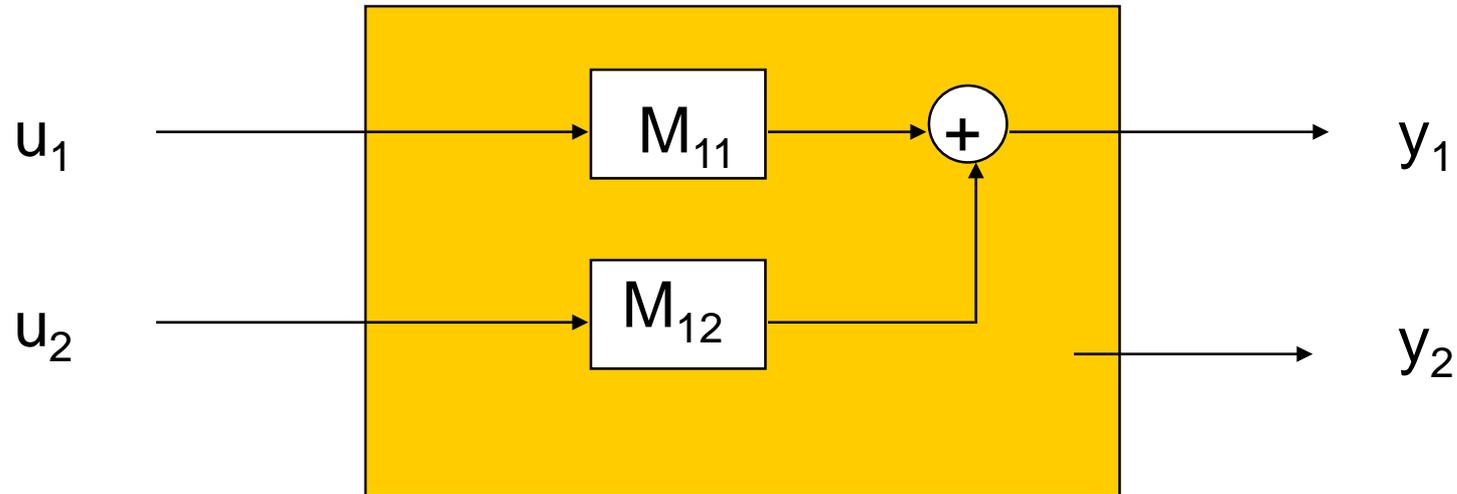
$$y(t) - \hat{y}(t|\theta) = e(t|\theta) = \frac{1}{H(q^{-1})} [y(t) - G(q^{-1})u(t)] \quad \text{Residuos}$$

$$\text{Si } \left. \begin{array}{l} H \rightarrow H_0 \\ G \rightarrow G_0 \end{array} \right\} \rightarrow e(t|\theta) \rightarrow \xi(t)$$

Test: Los residuos deben tener características de ruido blanco



Caso Multivariable



$$y_1 = M_{11} u_1 + M_{12} u_2 + \dots$$

Se hace una identificación para cada salida y todas las entradas m sistemas MISO



Caso multivariable

Modelo de respuesta
impulsional:

$$y_1(t) = \sum_{j=1}^{n_1} h_{11j} u_1(t-j) + \sum_{j=1}^{n_2} h_{12j} u_2(t-j)$$

$$y_2(t) = \sum_{j=1}^{n_1} h_{21j} u_1(t-j) + \sum_{j=1}^{n_2} h_{22j} u_2(t-j)$$

Puede formularse en la forma estandar:

$$y_1(t) = \varphi(t)' \theta_1$$

$$y_2(t) = \varphi(t)' \theta_2$$

$$\varphi(t)' = (u_1(t-1), \dots, u_1(t-n_1), u_2(t-1), \dots, u_2(t-n_2))$$

$$\theta'_i = [h_{i11}, \dots, h_{i1n_1}, h_{i21}, \dots, h_{i2n_2}]$$

Se resuelve un problema de identificación para cada salida



Caso Multivariable

Modelo de regresión:
$$y_1(t) = \frac{B_{11}(q^{-1})}{A_1(q^{-1})} u_1(t) + \frac{B_{12}(q^{-1})}{A_1(q^{-1})} u_2(t)$$

$$y_2(t) = \frac{B_{21}(q^{-1})}{A_2(q^{-1})} u_1(t) + \frac{B_{22}(q^{-1})}{A_2(q^{-1})} u_2(t)$$

$$y_i(t) = \varphi_i(t)' \theta_i$$

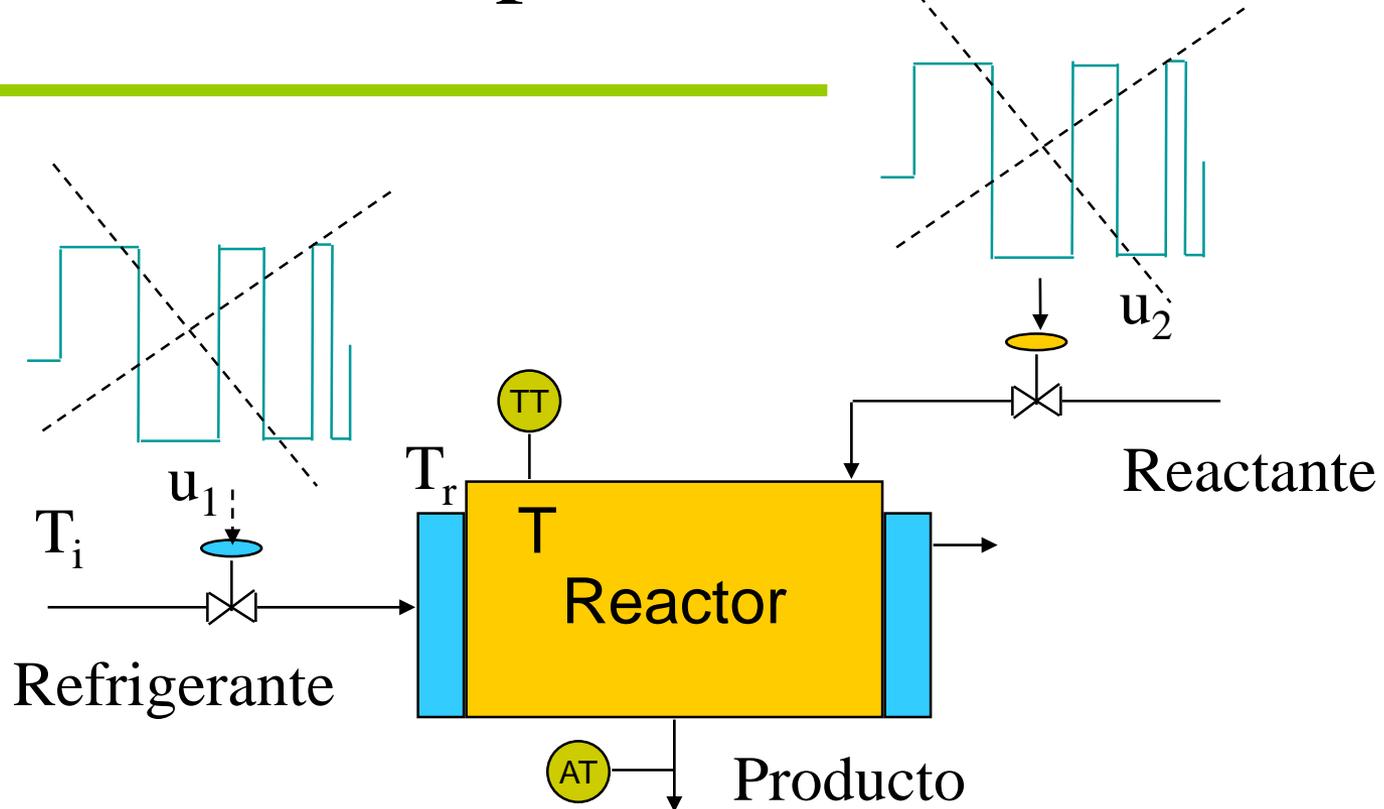
$$\varphi(t)' = [-y_i(t-1), \dots, -y_i(t-n), u_1(t-1), \dots, u_1(t-m_1), u_2(t-1), \dots, u_2(t-m_2), \dots]$$

$$\theta' = [a_{i1}, \dots, a_{in}, b_{i11}, \dots, b_{i1m_1}, b_{i21}, \dots, b_{i2m_2}, \dots]$$

$$V = \frac{1}{N} \sum_{t=1}^N \left[y_i(t) - \sum_{j=1}^2 \frac{B_{ij}(q^{-1})}{A_i(q^{-1})} u_j(t) \right]^2$$



Experimentos



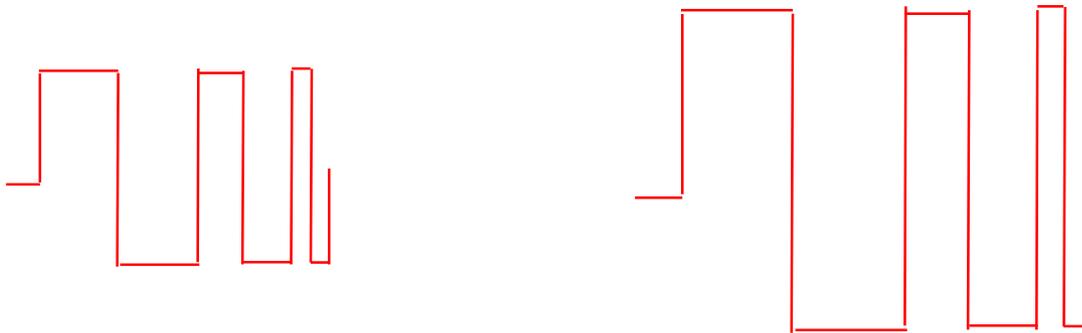
Si hay varias entradas, los cambios en las mismas deben de estar incorrelacionados para que el algoritmo de identificación pueda distinguir los efectos de cada entrada en las salidas



Experimentos

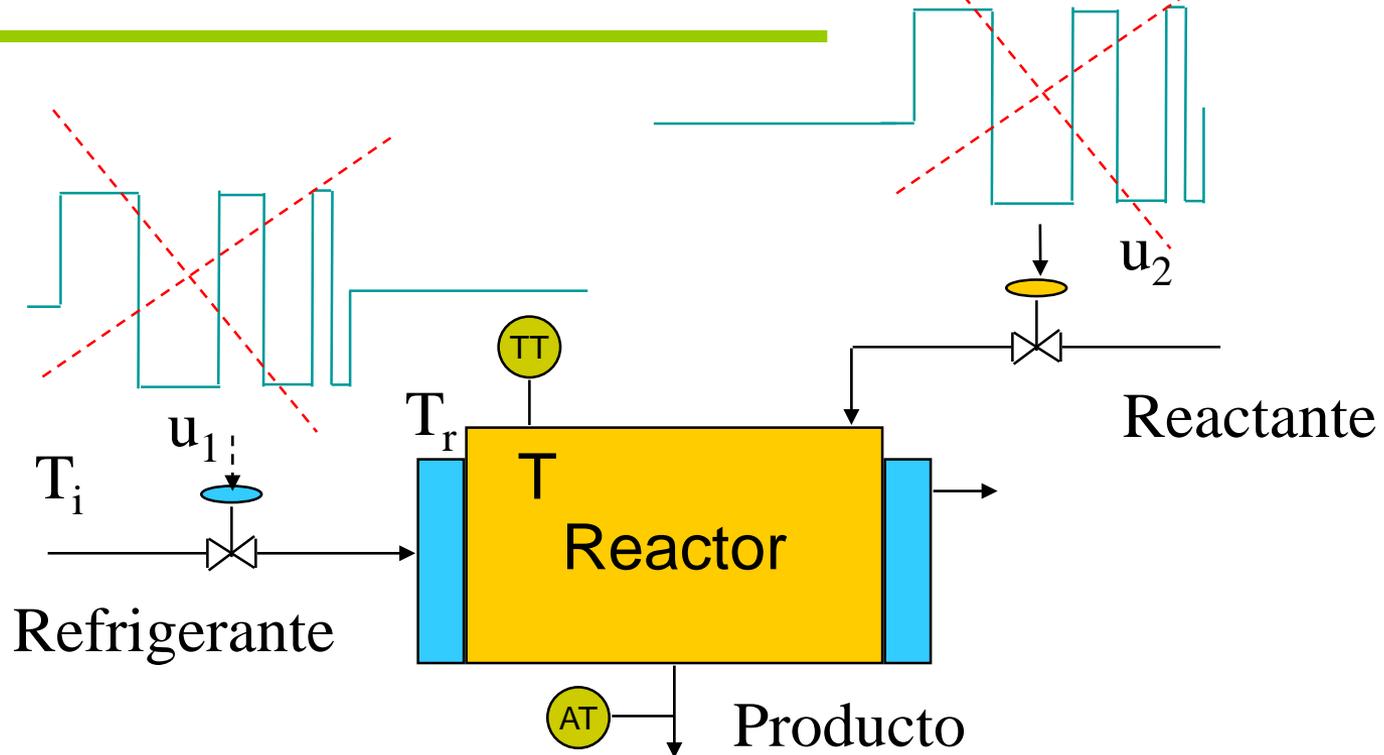
Por ejemplo, si se aplican señales proporcionales $u_1 = \alpha u_2$ a las entradas simultáneamente, hay muchas combinaciones de los modelos M_1 y M_2 que se ajustan a las mismas parejas de entradas y salidas

$$y(t) = M_1 u_1(t) + M_2 u_2(t) = (\alpha M_1 + M_2) u_2(t)$$





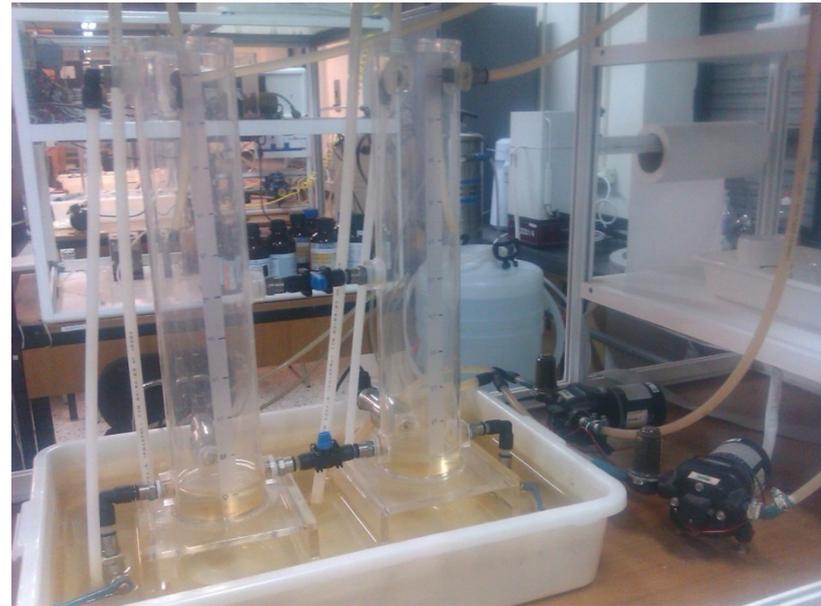
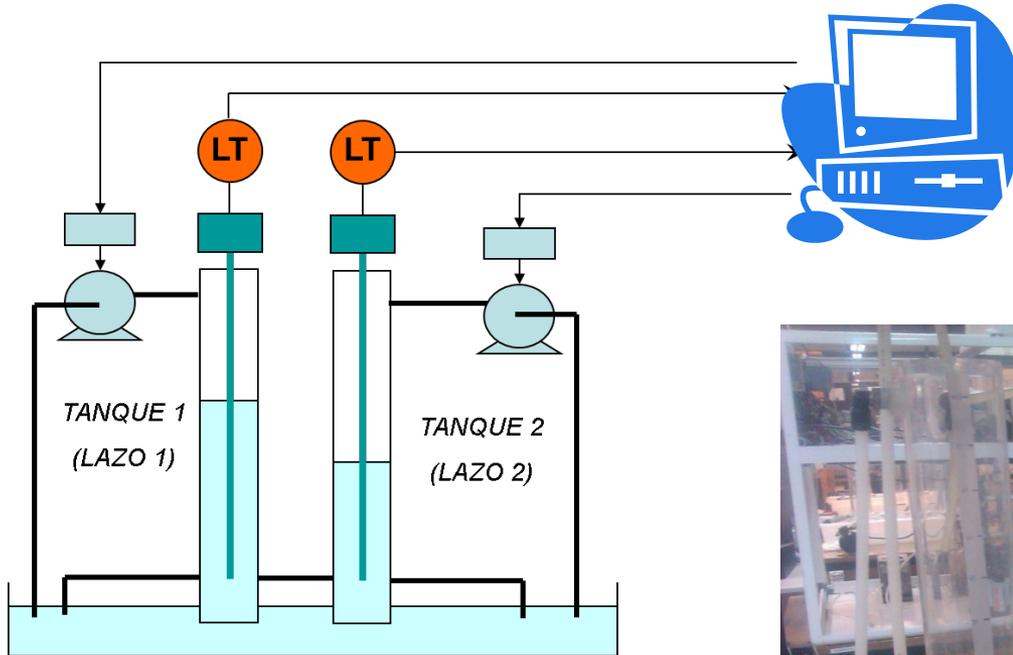
Experimentos



Excitar solo una entrada manteniendo las otras constantes y luego hacer lo mismo con las otras, no es una buena política: alarga el tiempo del experimento y no proporciona datos en condiciones realistas de funcionamiento



Dos depósitos acoplados





Modelos no lineales

- ✓ Muchas de las técnicas de modelos lineales pueden aplicarse directamente a modelos no-lineales, pero lineales en los parámetros que se desea estimar, p.e.

$$y(t) = \alpha e^{u(t-1)} + \beta \text{sen}(2u(t-2))$$

Ya que pueden formularse en el formato:

$$y(t) = \varphi(t)' \theta = (e^{u(t-1)}, \text{sen}(2u(t-2))) (\alpha, \beta)'$$

- ✓ El caso general requiere usar técnicas directas de optimización de una función de costo.



Validación

- ✓ El conjunto de test positivos y negativos permiten establecer el rango de validez del modelo
- ✓ Distintas situaciones implican distintas técnicas:
 - El proceso existe y se trata de construir un modelo que reproduzca la realidad
 - El proceso no existe aún y se trata de construir un modelo para predecir su comportamiento y optimizarlo cara al diseño



Validación

- ✓ Conjunto de pruebas de diferente naturaleza que ayudan a validar el modelo
- ✓ La validación ha de hacerse a todos los niveles de construcción del modelo:
 - Hipótesis
 - Formulación
 - Codificación
 - Resolución numérica (verificación)
 - Aplicación



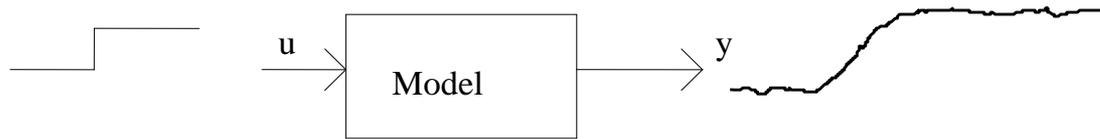
Validación

- ✓ Respuestas cualitativas del modelo
- ✓ Comprobación con datos experimentales
- ✓ Test estadísticos
- ✓ Análisis de sensibilidad
- ✓ Capacidad de predicción del modelo
- ✓ Distorsión de parámetros



Respuestas cualitativas

Respuestas cualitativas del modelo a entradas tipo

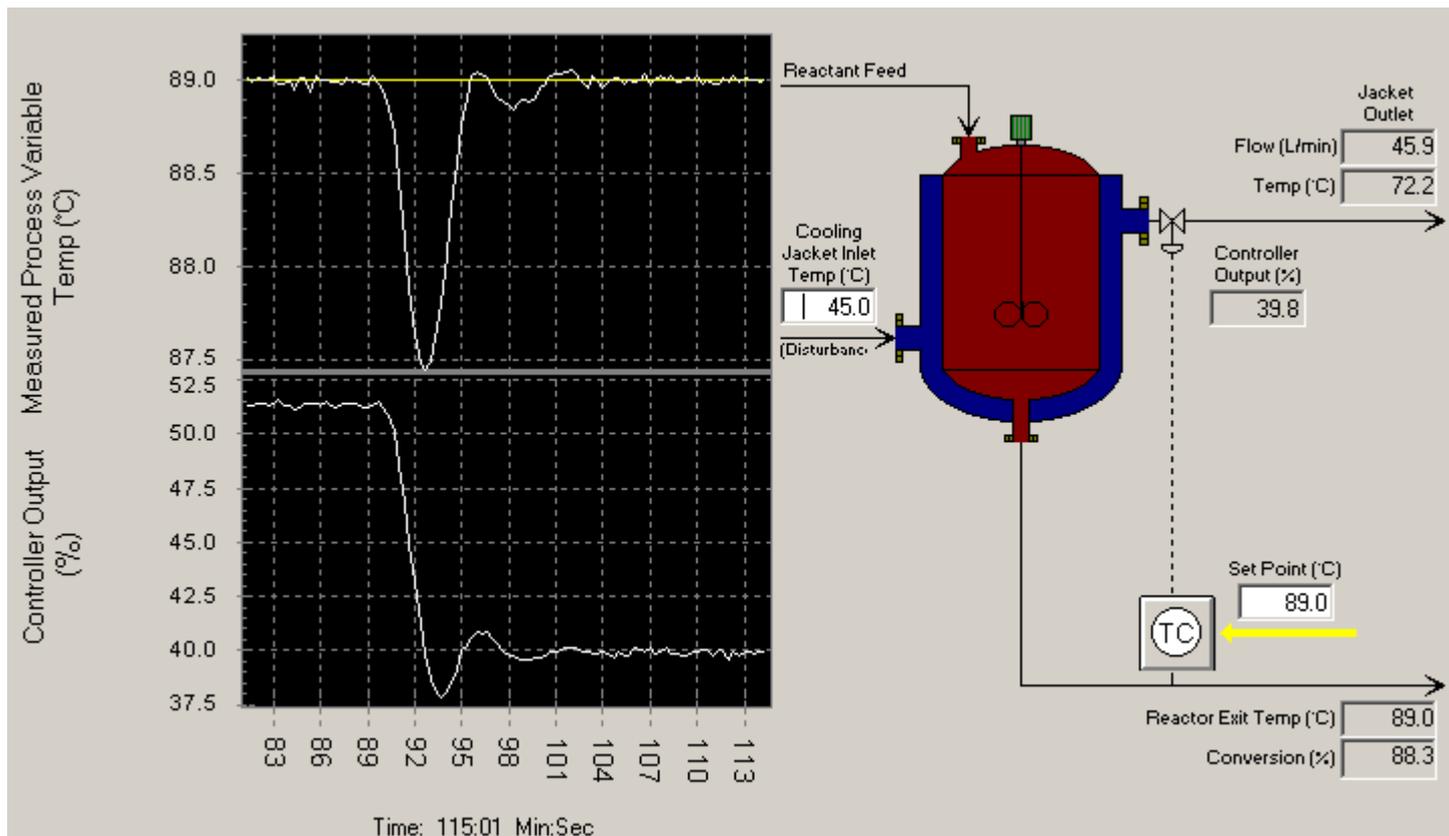


Discusión y evaluación de respuestas tipo del modelo con expertos en el proceso



Respuestas cualitativas

Respuestas cualitativas del modelo a entradas tipo





Respuestas cualitativas

- ✓ Test de Turing: ¿Se puede discriminar una respuesta del modelo y de la realidad si se presentan en el mismo formato?
- ✓ Trazas de la evolución de variables tipo flujo, etc. a través de los distintos componentes



Comprobación con datos experimentales

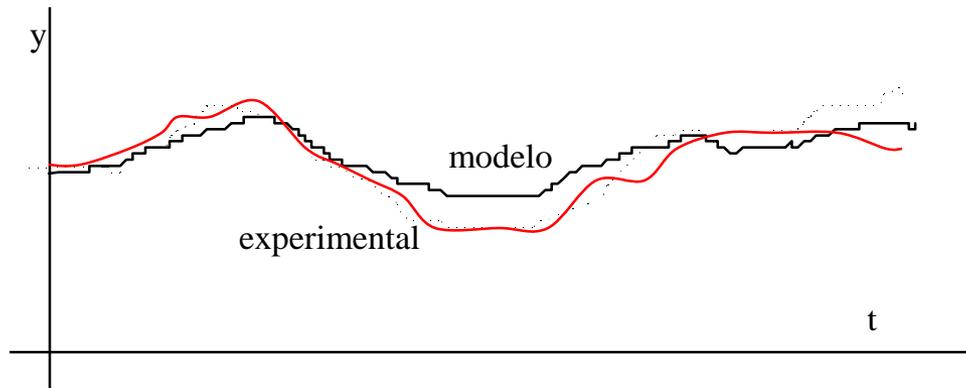
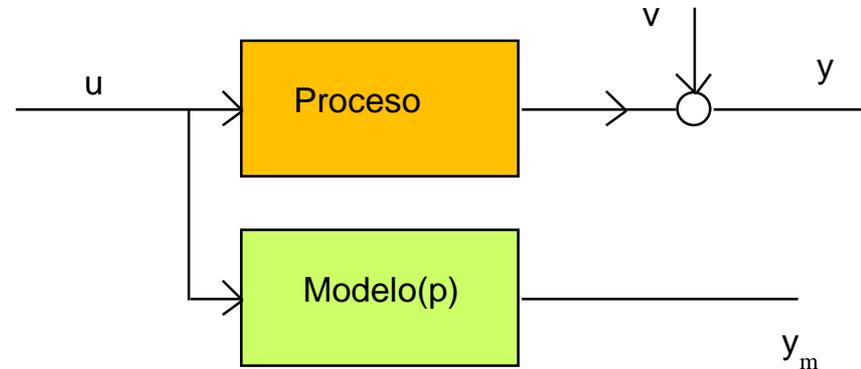


Comparación de las respuestas del modelo y datos experimentales distintos a los usados en la parametrización

$$\frac{1}{N} \sum_{t=1}^N (y(t) - y_m(t))^2$$

condiciones
iniciales

perturbaciones



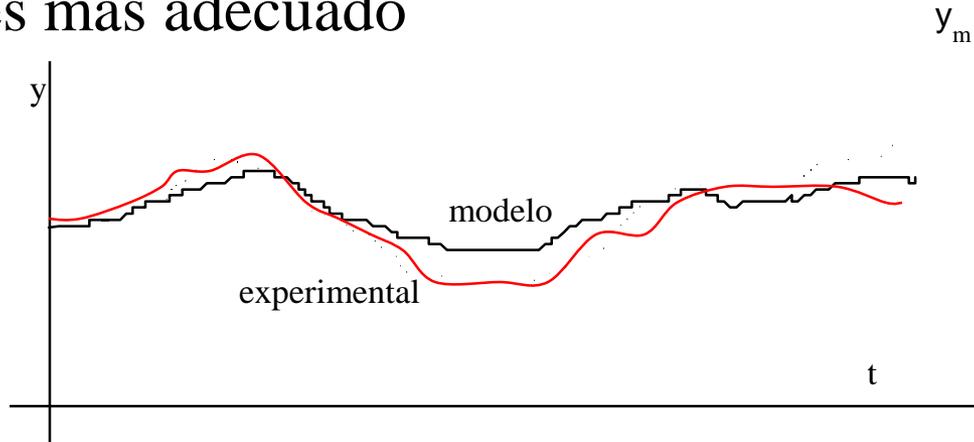


Indices de error

Medida numérica de las discrepancias entre la respuesta del modelo y los datos experimentales

Pueden usarse para comparar varios modelos de un proceso y decidir sobre cual es mas adecuado

$$\frac{1}{N} \sum_{t=1}^N (y(t) - y_m(t))^2$$





Indices de error

Muchos índices ponderan la suma de los errores y el número de parámetros d en el modelo

Final Prediction Error FPE

$$\frac{1}{N} \left(\frac{1+d/N}{1-d/N} \right) \sum_{t=1}^N (y(t) - y_m(t))^2$$

Estima de la varianza del error de predicción con los datos nuevos

Akaike Information Criterion AIC

$$\left(1 + \frac{2d}{N} \right) \sum_{t=1}^N (y(t) - y_m(t))^2$$
$$N \log \left(\frac{\sum_{t=1}^N (y(t) - y_m(t))^2}{N} \right) + 2d$$



Indices de error

Bayesian Information
Criterion BIC

$$N \log \left(\frac{\sum_{t=1}^N (y(t) - y_m(t))^2}{N} \right) + d \log(N)$$

Es consistente: La probabilidad de seleccionar un modelo erroneo tiende a cero con N

Rissanen's Minimal
Description length

$$\left(1 + \frac{2d}{N} \log N\right) \sum_{t=1}^N (y(t) - y_m(t))^2$$

Penaliza mas la complejidad del modelo



Indices de error

Test F para ver si el modelo j (con mas parámetros) es significativamente mejor que el modelo i, con un nivel α de confianza

$$\frac{\left(\left[\sum_{t=1}^N (y(t) - y_m(t))^2 \right]_{\text{mod } i} - \left[\sum_{t=1}^N (y(t) - y_m(t))^2 \right]_{\text{mod } j} \right) / (d_j - d_i)}{\left[\sum_{t=1}^N (y(t) - y_m(t))^2 \right]_{\text{mod } j} / (N - d_j)}$$

Debe compararse con la distribución $F(d_j - d_i, N - d_j)$ con el nivel α de confianza

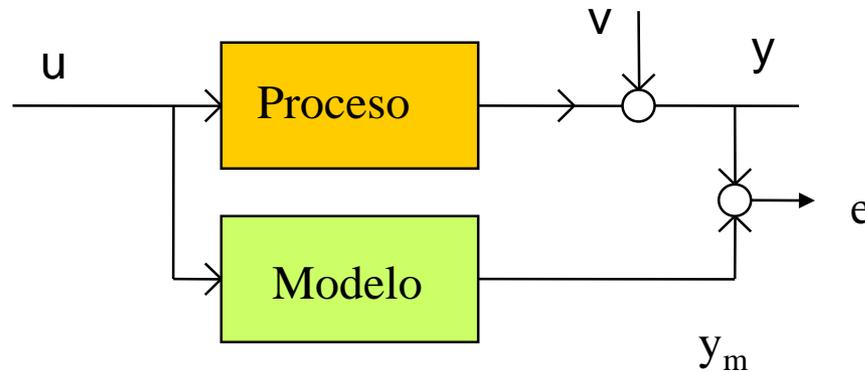


Errores

- ✓ Los errores de la estimación provienen de:
 - Errores de medida de los datos. Idealmente se suelen suponer independientes, de media cero y varianza constante. Para caracterizarlos pueden usarse datos de las salidas con entradas constantes.
 - Errores del modelo o la estima de parámetros. Si el modelo fuera perfecto, los residuos deberían tener las mismas características estadísticas que los errores de medida.



Test estadísticos



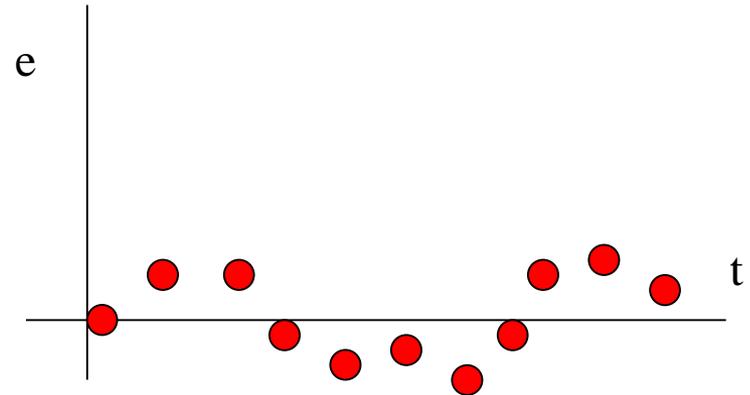
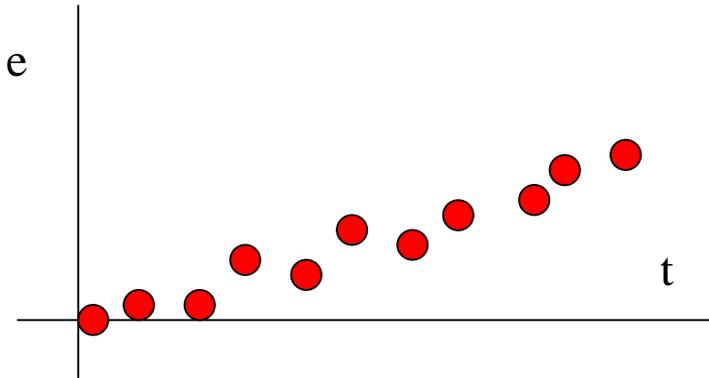
Si el modelo es correcto los residuos no deben presentar una estructura “sistemática”, sino ser el resultado de las perturbaciones aleatorias que actúan sobre el proceso

$$\text{Residuos } e(t) = y(t) - y_m(t)$$



Test estadísticos

✓ Inspección visual de los residuos



- nº de cambios de signo de los residuos (ncs)
- ¿Se ajustan a un modelo AR?

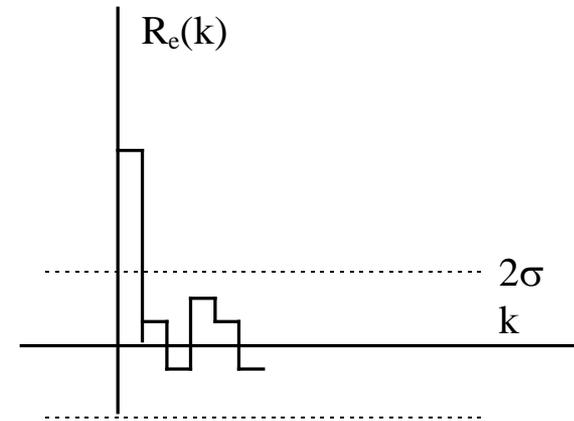
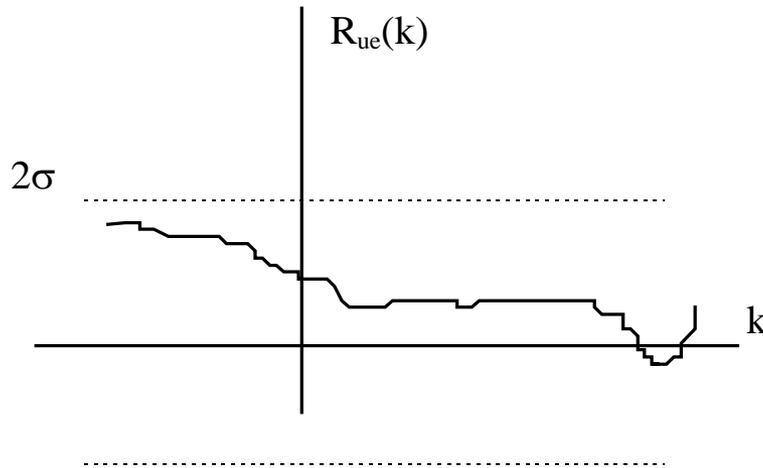
Test:
$$\frac{\text{ncs} - \frac{N}{2}}{\sqrt{N/2}}$$

Puede compararse con una distribución $N(0,1)$ para ver si hay un sesgo en los signos



Test estadísticos

No correlación entre la entrada y los residuos: Los errores deben ser independientes de un experimento particular



Lazo abierto, lazo cerrado



Varianza de los parámetros

Las regiones de confianza de los parámetros dan una medida de la calidad de la calibración y, por tanto, de la bondad el modelo.

$$C_p = \begin{bmatrix} c_{11} & c_{12} & \dots \\ c_{21} & c_{22} & \dots \\ \vdots & \vdots & \dots \end{bmatrix}$$

Los términos cruzados de la matriz de covarianza dan una idea de la independencia de los parámetros, pues deben estar incorrelacionados



Sensibilidad

¿ Cuanto varia J ante un cambio unidad en p ?

$$\min_p J = \sum_{j=1}^N [y(t) - y_m(p, u, t)]^2$$

$$\dot{x}(t) = f(x(t), u(t)) \quad y_m(t) = g(x(t), u(t))$$

$$\underline{p} \leq p \leq \bar{p}$$

Puede obtenerse a partir del problema dual de optimización en ciertos casos



Validación

Capacidad de predicción del modelo

Distorsiones de los parámetros

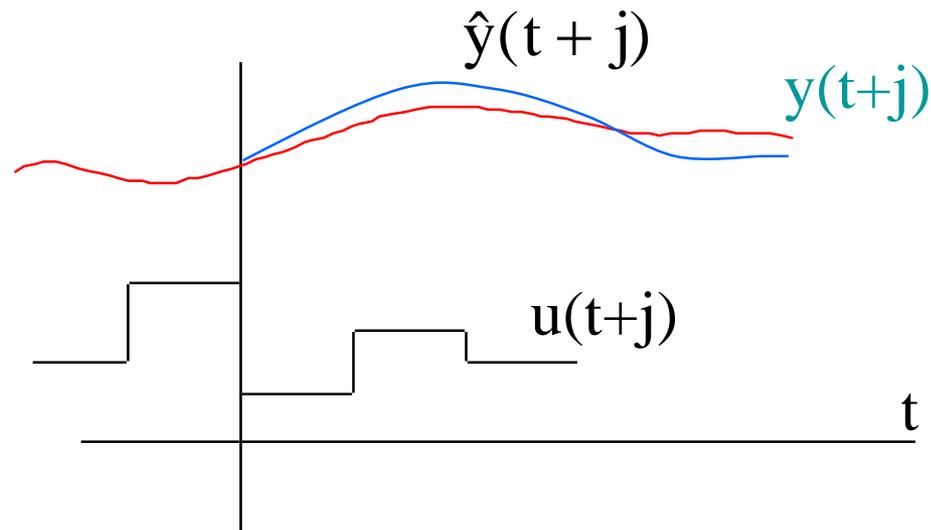
Coherencia de las distintas representaciones

Varianza de las estimaciones

$$\text{var}\{G(e^{j\omega t}, \hat{\theta})\} \cong \frac{d}{N} \frac{\Phi_v(\omega)}{\Phi_u(\omega)}$$



Capacidad de predicción



En ciertos casos, se utilizan fórmulas de predicción de controladores predictivos tales como el DMC, GPC, etc.

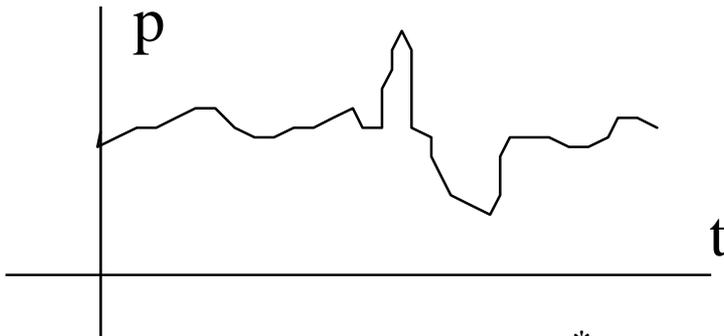


Distorsión de parámetros Butterfield 1986



Supongamos que ahora los parámetros p pueden variar a lo largo del tiempo,

¿ Como habría que modificar el valor de los parámetros del modelo para que sus respuestas coincidieran exactamente con las del proceso?



El valor de la distorsión da una medida de la credibilidad del modelo

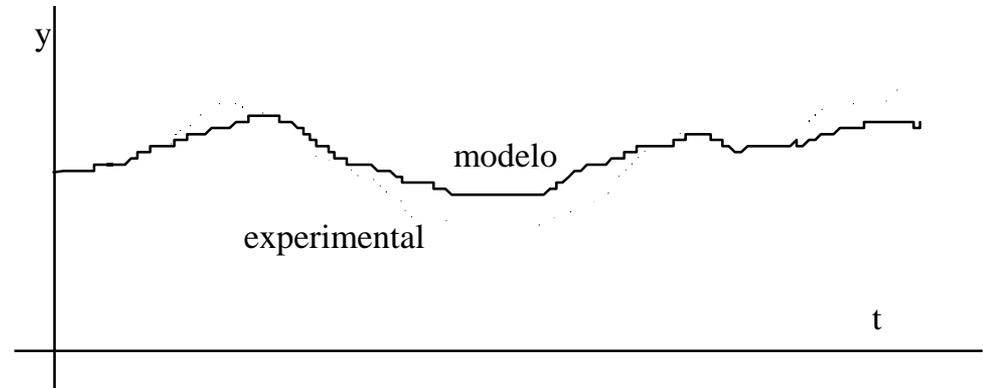
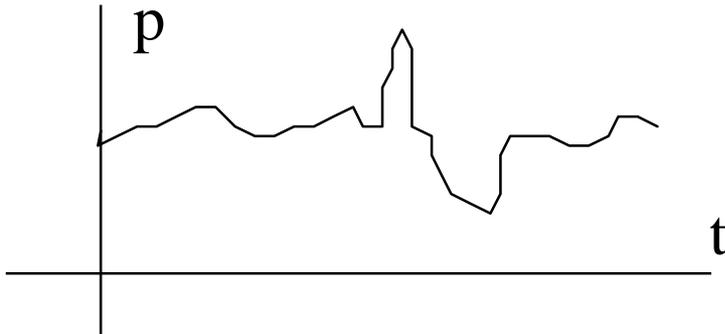
$$\min_{\Delta \mathbf{p}} y_m(t, \mathbf{p}^* + \Delta \mathbf{p}(t)) = y(t) \quad t = 1, \dots, n$$



Distorsión de parámetros

$$\frac{d \mathbf{x}_p(t)}{dt} = f(\mathbf{x}_p(t), u(t), \mathbf{p}^* + \Delta \mathbf{p}(t), t)$$

$$y_p(t) = y(t) = g(\mathbf{x}_p, u(t), t) \quad \text{Problema de índice}$$





Otros criterios de identificación

- ✓ Patrones de respuesta a entradas especiales
- ✓ Respuesta en frecuencia
- ✓ Funciones de correlación y espectros
- ✓ Minimización de una función de error
- ✓ Minimización de errores de predicción
- ✓ **Máxima verosimilitud**
- ✓ **Identificación global**
- ✓ **Minimización de distancia paramétrica**
- ✓ Identificación de subespacios



Máxima verosimilitud (ML)

- ✓ La respuesta del modelo $y_m(t)$ a los datos experimentales de entrada $u(\cdot)$ depende del valor de los parámetros del modelo θ
- ✓ Sea $p(y(t)|t-1, \theta)$ la probabilidad de que $y_m(t) = y(t)$ con el modelo de parámetros θ y datos hasta el instante $t-1$
- ✓ Criterio: elegir los parámetros θ que maximizan la probabilidad de obtener los datos $y(1), y(2), \dots, y(N)$ medidos



Máxima verosimilitud

$p(y(t)|t-1, \theta)$ probabilidad de que $y_m(t) = y(t)$
con el modelo de parámetros θ y datos hasta $t-1$

La probabilidad de obtener los valores $y(1), y(2), \dots, y(N)$ es:

$$\prod_{t=1}^N p_t(y(t):t-1, \theta)$$

Si la distribución es gaussiana:

$$\prod_{t=1}^N p_t(y(t):t-1, \theta) = \prod_{t=1}^N \frac{\exp\left[\frac{-(y(t) - \hat{y}_m(t:t-1, \theta))^2}{2\sigma^2}\right]}{\sigma\sqrt{2\pi}}$$



Máxima verosimilitud

$$\prod_{t=1}^N p_t(y(t):t-1, \theta) = \prod_{t=1}^N \frac{\exp\left[\frac{-(y(t) - \hat{y}_m(t:t-1, \theta))^2}{2\sigma^2}\right]}{\sigma\sqrt{2\pi}}$$

Normalmente en lugar de maximizar esta función se minimiza su logaritmo cambiado de signo

$$\min_{\theta} [-\log L(\mathbf{y}, \theta)] = \min_{\theta} \sum_{t=1}^N \frac{(y(t) - \hat{y}_m(t:t-1, \theta))^2}{2\sigma^2} + N \log \sigma\sqrt{2\pi}$$

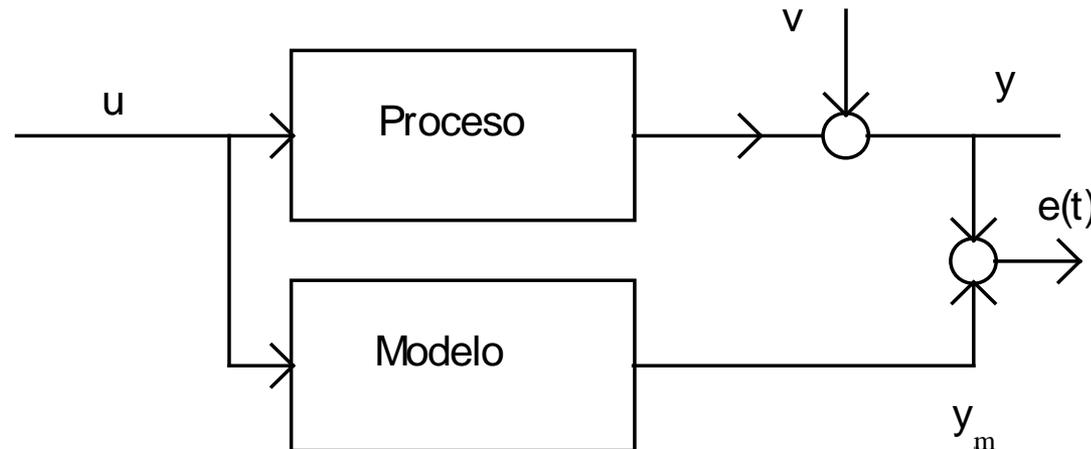
El problema se expresa como una minimización numérica de suma de errores de predicción $y(t) - y_m(t | t-1, \theta)$



Identificación Global

Criterio de minimización del error :Dado un conjunto de datos experimentales $u(t)$, $y(t)$, buscar los parámetros θ que minimizan la función de coste V :

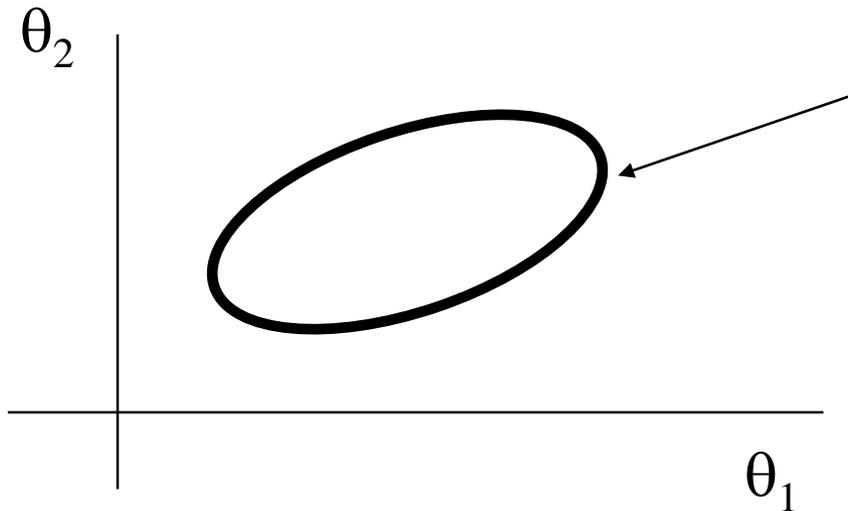
$$V = \frac{1}{N} \sum_{t=1}^N e(t)^2 = \frac{1}{N} \sum_{t=1}^N [(y(t) - y_m(t, \theta))]^2$$





Identificación global

$$V = \frac{1}{N} \sum_{t=1}^N [(y(t) - M(\theta)u(t))]^2 \quad f(\theta) = V = \text{cte.}$$



Isodistancia:
Lugar geométrico de los puntos del espacio paramétrico con igual valor de V

Todos los modelos sobre una isodistancia tienen la misma validez respecto al criterio de suma de cuadrados de errores



Identificación global

Pratique de l'identification

Jacques Richalet

Edit. Hermes

Matching the uncertainty of the model given by global identification techniques to the robustness of a model based predictive controller

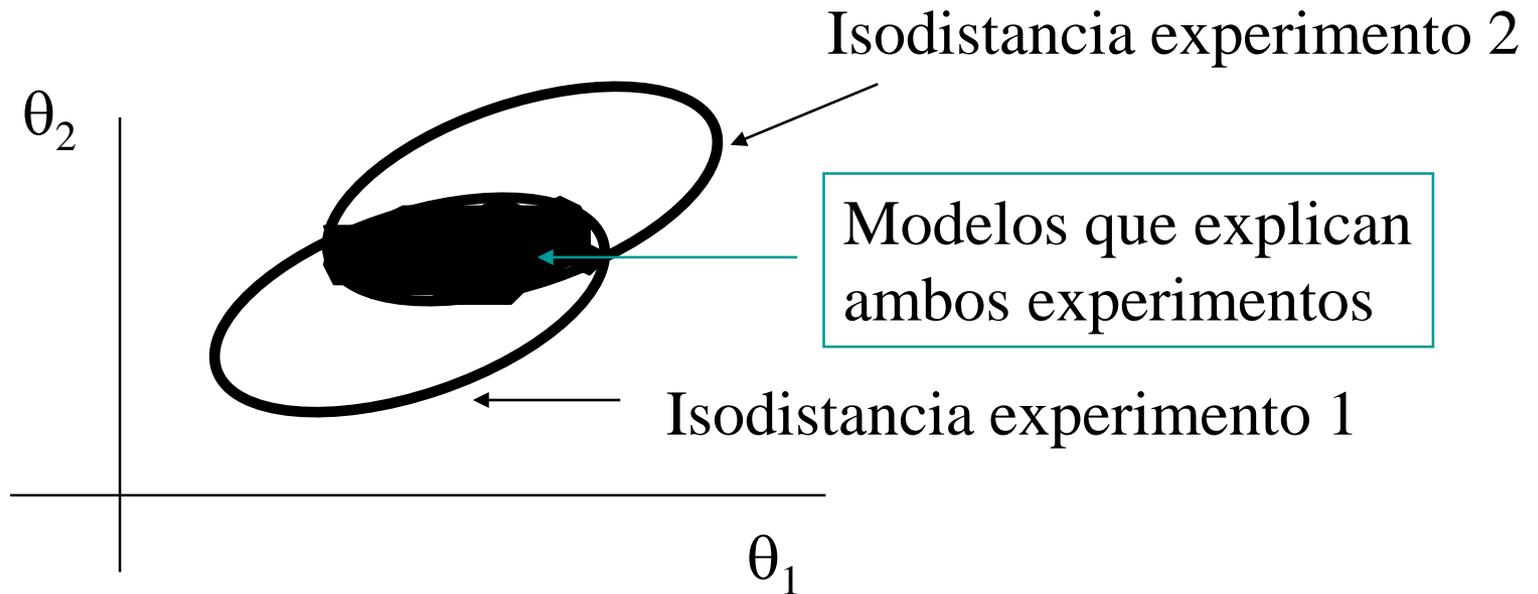
M. Sanzo, J. Richalet, C. de Prada

Advances in Model Based Predictive Control

Oxford University Press, 1994



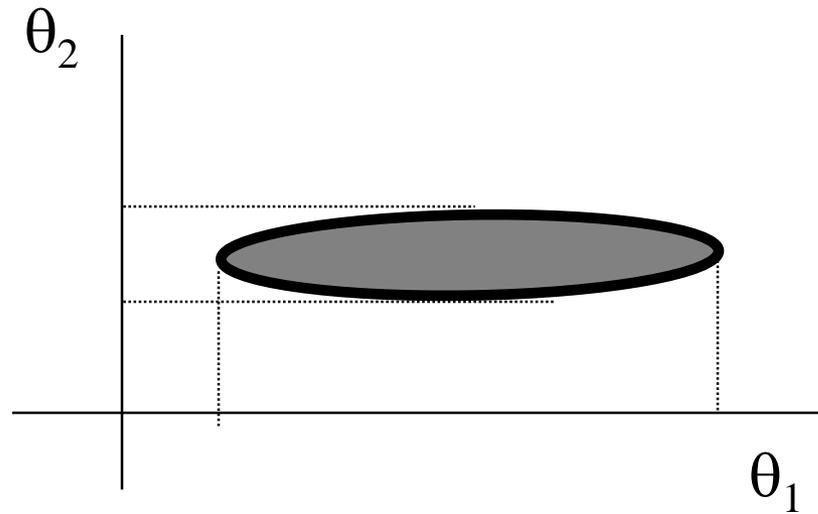
Identificación global



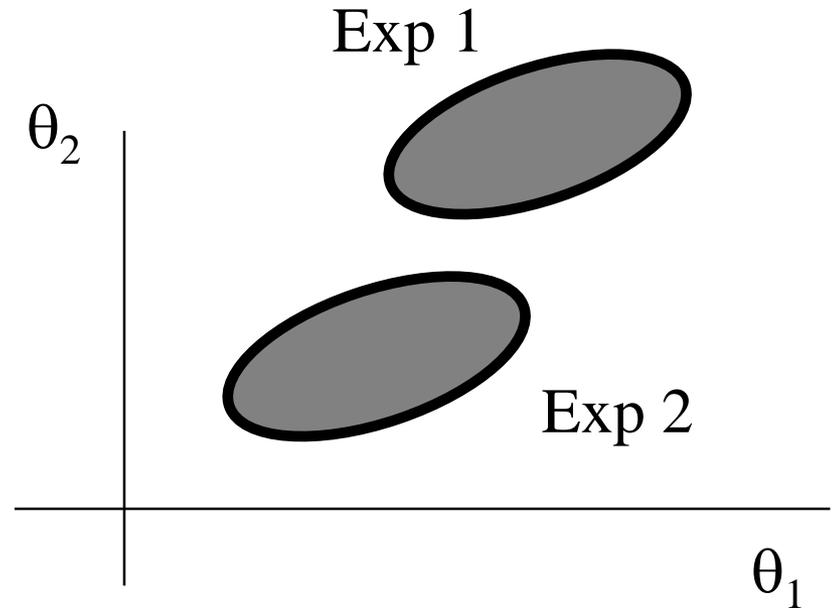
Las isodistancias dependen de los datos experimentales y del nivel de exactitud V elegido



Isodistancias



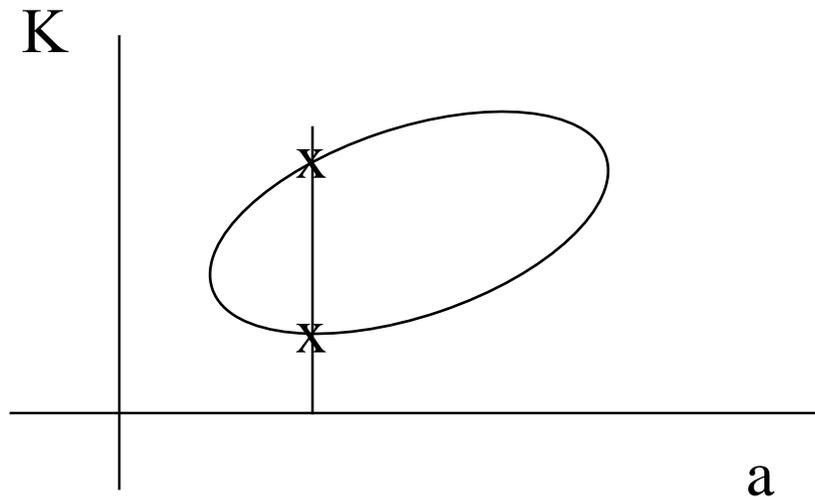
Informan sobre la calidad de los parámetros, y el rediseño de experimentos



Nivel de calidad no compatible con los experimentos



Cálculo de Isodistancias



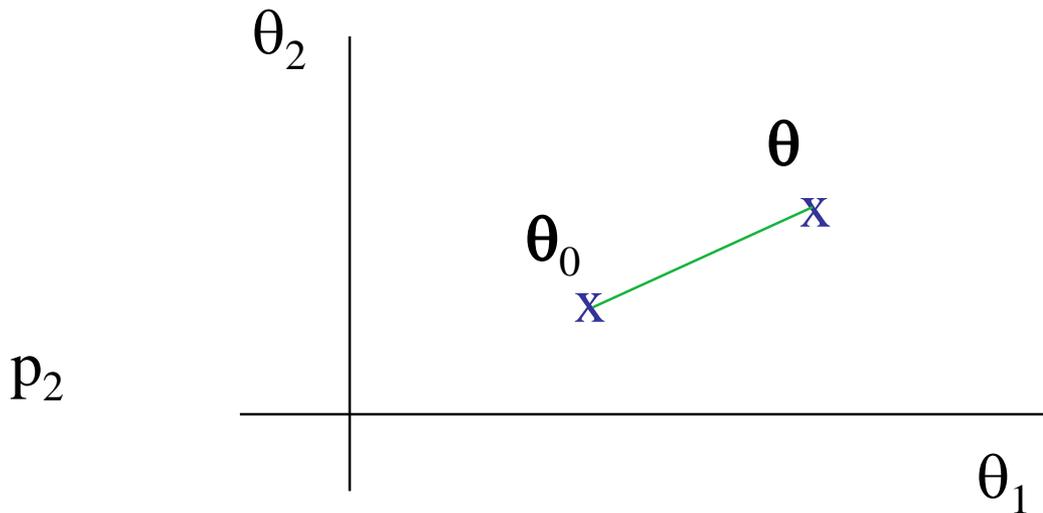
Ejemplo con un modelo sencillo de primer orden

Para cada valor de a se resuelve una ecuación de segundo grado para obtener el correspondiente valor de K

$$V = \frac{1}{N} \sum_{t=1}^N \left[(y(t) - \frac{Kq^{-1}}{1 + aq^{-1}} u(t)) \right]^2$$



Minimización de la distancia paramétrica



Criterio

$$\min_{\theta} \|\theta_0 - \theta\|$$

θ_0 Valor “verdadero” de los parámetros



Minimización de la distancia paramétrica

$$y(t) = \varphi(t)' \theta_0 + v(t)$$

$$y_m(t) = \varphi(t)' \theta(t-1)$$

$$V(t) = \|\theta_0 - \theta(t)\|^2$$

$$\theta(t) = \theta(t-1) + \Delta\theta(t)$$

$$V(t) = \|\theta_0 - \theta(t-1) - \Delta\theta(t)\|^2$$

Descripción del proceso
Modelo

Distancia paramétrica

Algoritmo de estimación

Evolución de $V(t)$

¿Como escoger $\Delta\theta$ para hacer que $V(t)$ sea decreciente?



Minimización de la distancia paramétrica



$$V(t) = \|\theta_0 - \theta(t-1) - \Delta\theta(t)\|^2 = \|\theta_0 - \theta(t-1)\|^2 + \|\Delta\theta(t)\|^2 - 2\Delta\theta(t)'[\theta_0 - \theta(t-1)]$$

$$V(t) = V(t-1) + \|\Delta\theta(t)\|^2 - 2\Delta\theta(t)'[\theta_0 - \theta(t-1)]$$

$$y(t) = \varphi(t)'\theta_0 + v(t)$$

$$y_m(t) = \varphi(t)'\theta(t-1)$$

$$y(t) - y_m(t) = \varphi(t)'[\theta_0 - \theta(t-1)] + v(t)$$

$$\Delta\theta(t) = -\rho\varphi(t)$$

Para eliminar el término desconocido $\theta_0 - \theta$:

$$V(t) - V(t-1) = \rho^2 \|\varphi(t)\|^2 + 2\rho\varphi(t)'[\theta_0 - \theta(t-1)] =$$

$$= \rho^2 \|\varphi(t)\|^2 + 2\rho[y(t) - y_m(t) - v(t)]$$

¿Qué valor de ρ hace $V(t) - V(t-1)$ negativo?



Caso sin ruido

Si $v(t)=0$

$$V(t) - V(t-1) = \rho^2 \|\varphi(t)\|^2 + 2\rho[y(t) - y_m(t)]$$

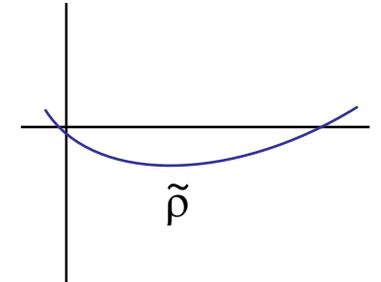
$$\frac{\partial(V(t) - V(t-1))}{\partial\rho} = 0 \quad \text{minimo}$$

$$\tilde{\rho} = -\frac{[y(t) - y_m(t)]}{\|\varphi(t)\|^2} \quad V(t) - V(t-1) \text{ es negativo para } 0 < \rho < 2\tilde{\rho}$$

si se escoge

$$\theta(t) = \theta(t-1) + \mu \frac{[y(t) - y_m(t)]\varphi(t)}{\|\varphi(t)\|^2} \quad \text{con } 0 < \mu < 2$$

$$V(t) = V(t-1) + (\mu^2 - 2\mu) \frac{[y(t) - y_m(t)]^2}{\|\varphi(t)\|^2} \quad V(t) \text{ es decreciente}$$





Algoritmo NLMS

$$\theta(t) = \theta(t-1) + \mu \frac{[y(t) - \varphi(t)' \theta(t-1)] \varphi(t)}{\|\varphi(t)\|^2} \quad \text{con } 0 < \mu < 2$$

$$\text{si } d(t) = \theta(t) - \theta_0 \quad \text{y} \quad y(t) = \varphi(t)' \theta_0 + v(t)$$

$$\begin{aligned} d(t) &= d(t-1) + \mu \frac{[\varphi(t)' \theta_0 + v(t) - \varphi(t)' \theta(t-1)] \varphi(t)}{\|\varphi(t)\|^2} = \\ &= \left[\mathbf{I} - \mu \frac{\varphi(t)' \varphi(t)}{\|\varphi(t)\|^2} \right] d(t-1) + \mu \frac{v(t) \varphi(t)}{\|\varphi(t)\|^2} \end{aligned}$$

Si el ruido $v(t)$ y $\varphi(t)$ están incorrelacionados

$$E\{\|d(t)\|\} = |1 - \mu| E\{\|d(t-1)\|\} \quad \text{para } 0 < \mu < 2 \text{ y la distancia decrece}$$



Algoritmo NLMS

NLMS Normalize Least Mean Square

Mínima media cuadrada normalizada

$$\theta(t) = \theta(t-1) + \mu \frac{[y(t) - \varphi(t)' \theta(t-1)] \varphi(t)}{\sigma + \|\varphi(t)\|^2} \quad \text{con } 0 < \mu < 2$$

- Con ruido, el posible sesgo no depende de μ
- La varianza de la estimación decrece con μ
- La mejor velocidad de convergencia se obtiene con μ proximo a 1
- Compromiso velocidad / precisión
- σ es una constante para evitar que θ  cuando $\varphi(t)$ 