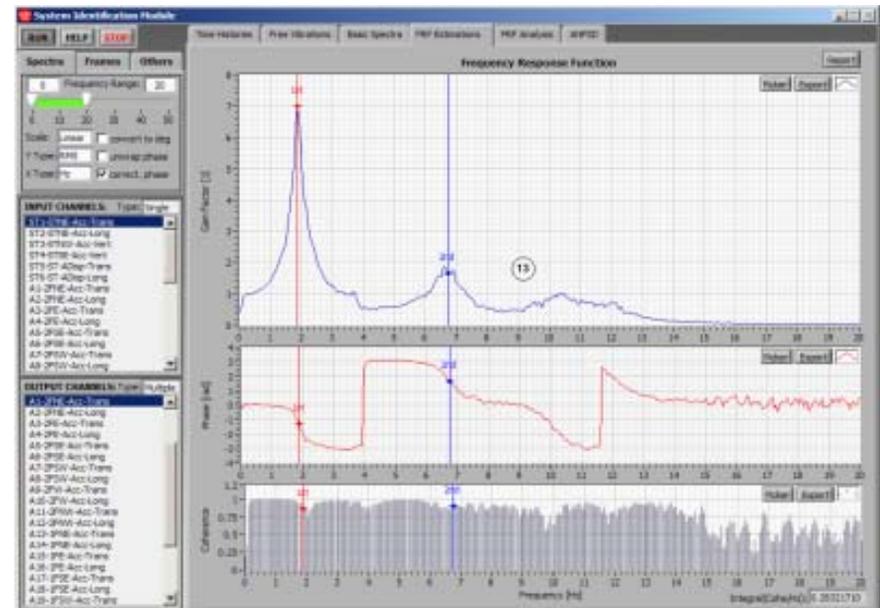
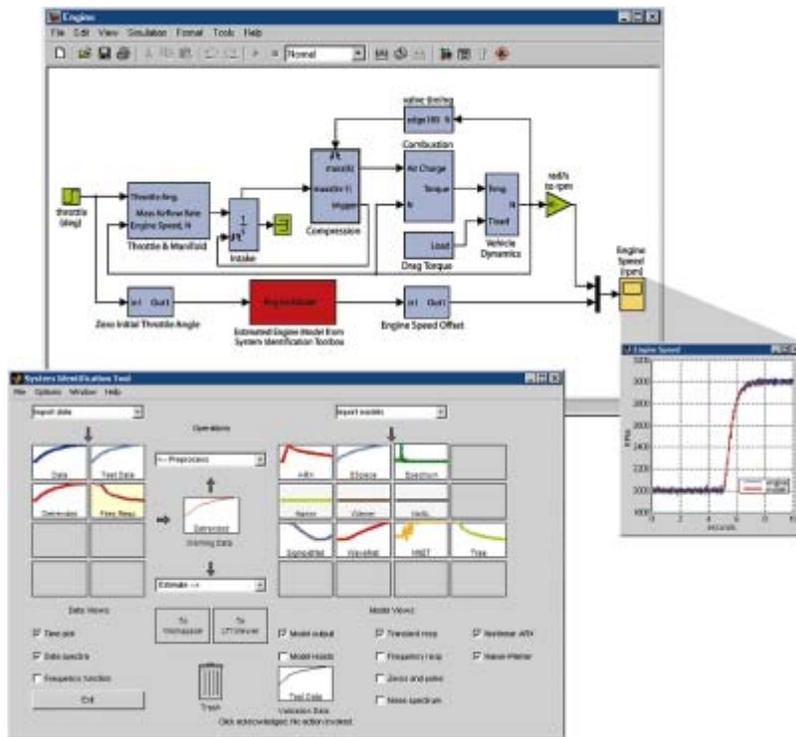


Identificazione di Sistema



Introduzione

- Un modello matematico di un'apparecchiatura o di un processo è di importanza assoluta in campo ingegneristico.
- Un modello numerico di un processo trova applicazione nella chimica, nell'elettronica, nella meccanica, nelle scienze economiche, ...
- Un modello ci permette di trovare risposte senza necessariamente misurare i processi reali o effettuare **esperimenti**.
- Un'altra caratteristica fondamentale consiste nella possibilità di **predire** la risposta futura di un sistema → controllo multivariabile basato su modello; ottimizzazione di processo.



Alcune domande



- Come si costruisce un modello matematico a fini predittivi?
- Come utilizzare i dati sperimentali (se disponibili) per la sintesi del modello?
- Come verificare l'attendibilità/affidabilità del modello?

- I modelli matematici possono essere divisi in due classi:
 - Modelli fondamentali;
 - Modelli empirici.

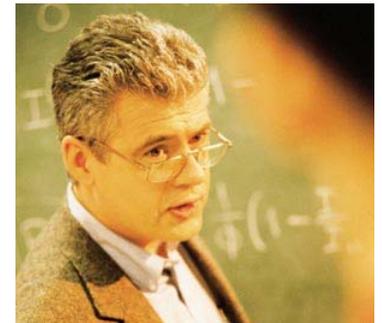
- **MODELLI FONDAMENTALI**: derivati da una conoscenza approfondita della fisica del processo. Leggi di conservazione: massa, energia, quantità di moto. Modelli basati su principi primi. Modelli deterministici.

- **MODELLI EMPIRICI**: il processo non viene visto nell'ottica delle leggi fisiche, bensì lo si conosce solo tramite osservazioni quantitative, esperimenti e misurazioni.



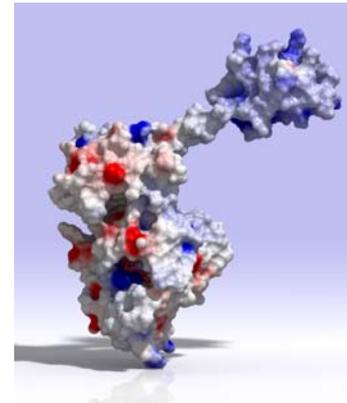
Modelli black-box

- Un modello di tipo black-box è completamente indifferente alla fisica del processo da identificare.
- In linea di principio è possibile costruire un modello di tipo black-box senza avere alcuna conoscenza specifica del processo da modellare.
- Evidentemente una **conoscenza approfondita** del processo permette di incrementare la bontà del modello prodotto.
- In genere occorre **mediare** tra bontà della predizione prodotta dal modello e sua complessità.



Quando è richiesta l'identificazione

- Se esiste **ignoranza** del processo da modellare;
- Se la **complessità** del processo da modellare è elevata;
- Se si ha necessità di ottenere risposte in **tempi ridotti**:
 - lato realizzazione del modello;
 - lato simulazione del modello.

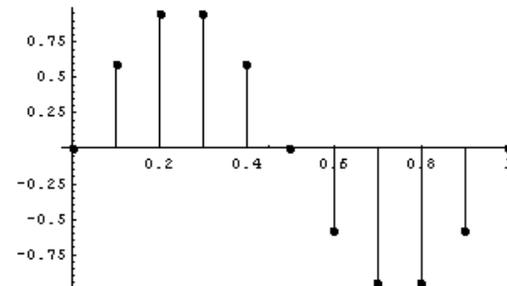
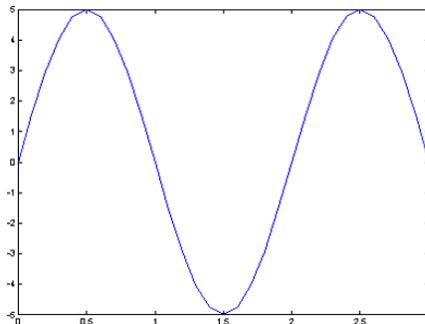


- Se è necessario un modello che non produca errori matematici (**floating point exceptions**).



Segnali e sistemi

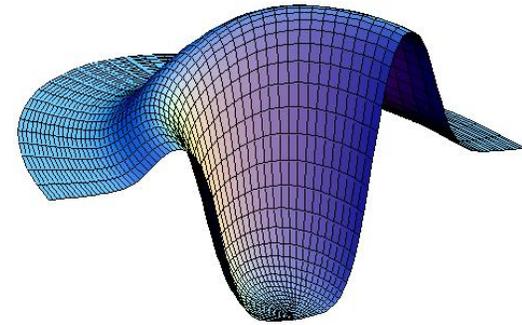
- È utile approfondire i concetti di **segnale** e **sistema** secondo la nomenclatura proposta da Bosh e Klaw, 1994.
- **SEGNALE**
 - È qualcosa che porta con sé dell'**informazione**.
 - Segnali **deterministici**: completamente determinati ad esempio da una espressione matematica;
 - Segnali **stocastici**: l'esatto valore futuro del segnale non può essere predetto. Il segnale viene descritto tramite un approccio statistico ad esempio in termini di media e deviazione standard.
 - Un segnale può essere **continuo** o **discreto**.



Segnali e sistemi

- **SISTEMA**

- È un insieme di **relazioni tra variabili e segnali**.
- Il sistema è una idealizzazione del processo reale.



- **MODELLI STATE-SPACE**

- Lo stato di un sistema dinamico è tale da contenere l'intera storia passata del processo;
- È quindi possibile predire il comportamento futuro del sistema senza la necessità di conoscere la storia passata del processo.
- Matematicamente ciò corrisponde a richiedere che il sistema contempli un numero di **variabili di stato**, x_i , tale da poter essere descritto da un insieme di n equazioni differenziali del primo ordine.
- n è definito "ordine del sistema".

Segnali e sistemi

- **MODELLI STATE-SPACE**

- Oltre alle variabili di stato ci sono quelle di **input**, u_i , che sono le cosiddette “driving force” del sistema.
- La soluzione del sistema di n equazioni differenziali fornisce un quadro completo relativo al comportamento futuro del sistema.
- Conoscere i valori x_i può essere di utilità trascurabile dal momento che le variabili di stato non necessariamente sono osservabili.
- Le variabili di interesse, y_i , sono quelle di **output** e sono funzione delle variabili x_i e u_i .
- La formulazione **deterministica continua** di un sistema **state-space** è:

$$\dot{\mathbf{x}}(t) = \mathbf{G}(\mathbf{x}(t), \mathbf{u}(t))$$

$$\mathbf{y}(t) = \mathbf{H}(\mathbf{x}(t), \mathbf{u}(t))$$



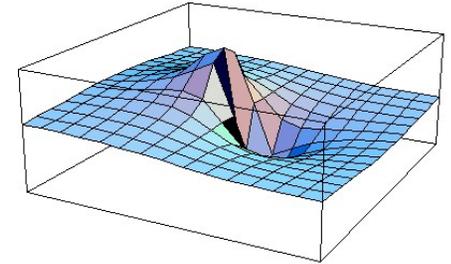
Segnali e sistemi

- **MODELLI STATE-SPACE**

- La formulazione **deterministica discreta** di un sistema **state-space** è:

$$\mathbf{x}(t+1) = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t))$$

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t))$$



- **MODELLI BLACK-BOX**

- Come detto in precedenza i modelli black-box comunicano con l'osservatore unicamente tramite segnali. All'osservatore non è quindi dato di conoscere le funzioni **g** e **h** che governano il sistema.
- Dato che l'unica informazione a disposizione di un utente, relativa ad un sistema black-box, sono i segnali di input e di output allora le variabili di stato **x** non sono note.
- Un modello black-box non utilizza perciò le variabili di stato.

Modelli black-box

- È quindi ragionevole assegnare ad un modello black-box una struttura:

$$\text{output del sistema} = \mathbf{f}(\text{input del sistema}) \quad \Rightarrow \quad \mathbf{y} = \mathbf{f}(\mathbf{u})$$

- Se il modello è dinamico l'output del sistema dipenderà dalla storia passata degli input e degli output:

$$\text{output attuale} = \mathbf{f}(\text{output passati, input passati}) \quad \Rightarrow \quad \mathbf{y}_{now} = \mathbf{f}(\mathbf{y}_{old}, \mathbf{u}_{old})$$

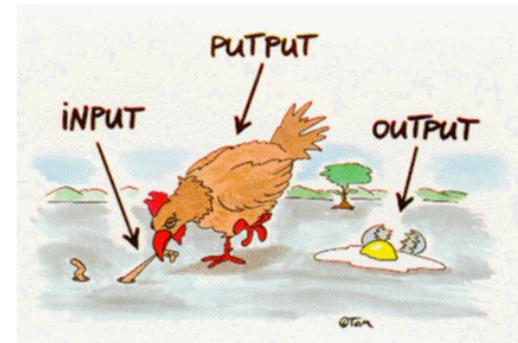
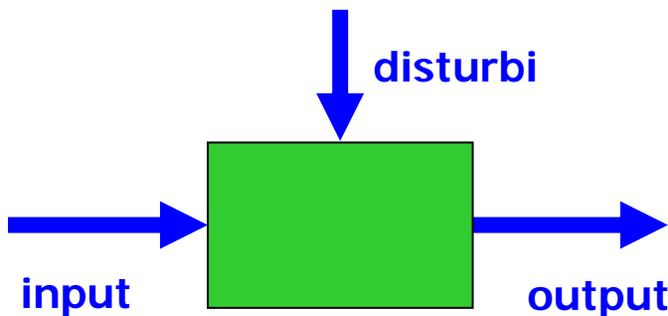
- Affinché la procedura di identificazione del legame funzionale, \mathbf{f} , abbia successo è opportuno introdurre dei **parametri adattivi**, \mathbf{p} .

$$\text{output attuale} = \mathbf{f}(\text{output passati, input passati, parametri}) \quad \Rightarrow \quad \mathbf{y}_{now} = \mathbf{f}(\mathbf{y}_{old}, \mathbf{u}_{old}, \mathbf{p})$$

Modelli black-box

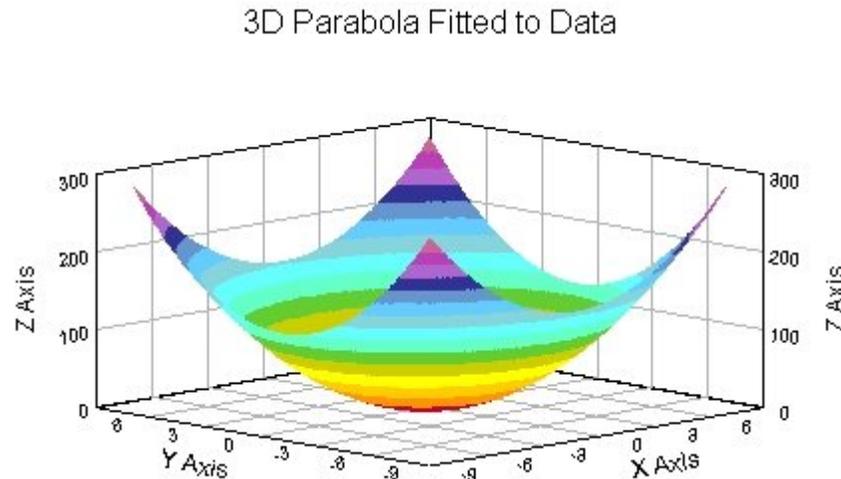
- I parametri \mathbf{p} possono essere utilizzati come variabili indipendenti nella procedura di identificazione (gradi di libertà) con l'obiettivo che il modello black-box descriva nel miglior modo possibile i dati di input-output provenienti dal processo reale.
- Un possibile miglioramento del modello black-box, \mathbf{f} , può avvenire considerando l'errore, \mathbf{e} , che si compie calcolando la risposta del sistema rispetto alla misura del processo.

$$\text{output attuale} = \mathbf{f} \left(\begin{array}{l} \text{output passati, input passati,} \\ \text{errori passati, parametri} \end{array} \right) \quad \Rightarrow \quad \mathbf{y}_{now} = \mathbf{f}(\mathbf{y}_{old}, \mathbf{u}_{old}, \mathbf{e}_{old}, \mathbf{p})$$



Identificazione di sistema

- Sono tre gli step richiesti alla identificazione di sistema:
 1. Selezione dell'insieme di input, \mathbf{u} , output, \mathbf{y} , errori, \mathbf{e} , in termine di numero di incognite ed ampiezza dell'intervallo temporale;
 2. Selezione del regressore, \mathbf{f} ;
 3. Regressione del modello rispetto alle osservazioni tramite utilizzo dei parametri come variabili indipendenti (gradi di libertà).



Regressori

- Un modello non deve necessariamente considerare tutti gli input ed output di processo osservabili.
- Nel caso di un sistema da utilizzare a **fini controllistici** può ad esempio essere sufficiente considerare le variabili **controllate**, quelle **manipolate** ed eventualmente i **disturbi misurabili**.
- In generale si hanno:
 - r variabili di output, \mathbf{y} ;
 - m variabili di input, \mathbf{u} .
- Il vettore degli errori ha invece struttura: $\mathbf{e} = \mathbf{y}_{real} - \mathbf{y}$
- Le $r + m$ variabili del modello vengono quindi campionate (misurate e memorizzate) ogni tempo t_s .
- Tali variabili hanno il compito di descrivere la storia del sistema (tenendo conto dell'effetto di affievolimento del segnale con il tempo).



Regressori

- Il sistema da identificare assume quindi la forma:

$$\mathbf{y}(t) = \mathbf{f}[y_1(t-1), \dots, y_1(t-n_{y_1}), \dots, y_r(t-1), \dots, y_r(t-n_{y_r}), \\ u_1(t-1), \dots, u_1(t-n_{u_1}), \dots, u_m(t-1), \dots, u_m(t-n_{u_m}), \\ e_1(t-1), \dots, e_1(t-n_{e_1}), \dots, e_r(t-1), \dots, e_r(t-n_{e_r})]$$

- Si introduce il vettore $\boldsymbol{\varphi}$ le cui componenti sono definite **regressori**:

$$\boldsymbol{\varphi}(t) = [y_1(t-1), \dots, y_1(t-n_{y_1}), \dots, y_r(t-1), \dots, y_r(t-n_{y_r}), \\ u_1(t-1), \dots, u_1(t-n_{u_1}), \dots, u_m(t-1), \dots, u_m(t-n_{u_m}), \\ e_1(t-1), \dots, e_1(t-n_{e_1}), \dots, e_r(t-1), \dots, e_r(t-n_{e_r})]^T$$

- Se d è il numero totale di variabili del sistema, allora la lunghezza ℓ del vettore $\boldsymbol{\varphi}$ è definita come l'**ordine totale** del modello:

$$\ell = \sum_{i=1}^d n_i$$

Regressori

- Considerando il fatto che gli input possono avere un effetto ritardato sugli output è possibile introdurre dei **tempi di ritardo**, n_{ki} , nel modello del sistema in numero pari agli m input:

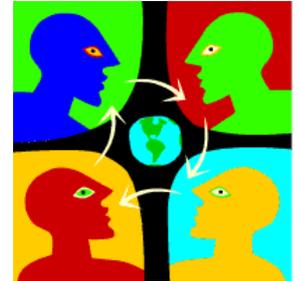
$$\begin{aligned} \boldsymbol{\varphi}(t) = & [y_1(t-1), \dots, y_1(t-n_{y_1}), \dots, y_r(t-1), \dots, y_r(t-n_{y_r}), \\ & u_1(t-n_{k_1}-1), \dots, u_1(t-n_{k_1}-n_{u_1}), \dots, u_m(t-n_{k_m}-1), \dots, u_m(t-n_{k_m}-n_{u_m}), \\ & e_1(t-1), \dots, e_1(t-n_{e_1}), \dots, e_r(t-1), \dots, e_r(t-n_{e_r})]^T \end{aligned}$$



La funzione f

- La funzione f , tramite i parametri \mathbf{p} , **mappa** il vettore dei regressori nelle variabili di output \mathbf{y} :

$$\mathbf{y}(t) = \mathbf{f}[\boldsymbol{\varphi}(t), \mathbf{p}]$$



- È possibile distinguere tra mappature (funzioni) **lineari** e **non lineari**.
- Il modello più semplice per la funzione f è:

$$\mathbf{y}(t) = \mathbf{p} \times \boldsymbol{\varphi}(t)$$

N.B.: il vettore \mathbf{p} è un vettore riga mentre $\boldsymbol{\varphi}$ è un vettore colonna.

- Analogamente se \mathbf{p} è una matrice si ottiene il vettore \mathbf{y} .
- Se ora ipotizziamo che il vettore di output \mathbf{y} sia pari alla somma di due termini relativi ad un **contributo deterministico** (non disturbato) $\boldsymbol{\rho}$ e ad un **contributo di disturbo** $\boldsymbol{\omega}$ è possibile scrivere:

$$\mathbf{y}(t) = \boldsymbol{\rho}(t) + \boldsymbol{\omega}(t)$$

La funzione f

- Nella rappresentazione $\mathbf{y}(t) = \boldsymbol{\rho}(t) + \boldsymbol{\omega}(t)$ il termine $\boldsymbol{\omega}$ rappresenta il contributo alla variabile di output \mathbf{y} non modellabile deterministicamente.
- $\boldsymbol{\omega}$ produce degli effetti stocastici sul sistema, è cioè il rumore e/o la deviazione del sistema dalla ideale linearità.
- La porzione deterministica del modello può essere espressa come:

$$\boldsymbol{\rho}(t) = \mathbf{G}(q, \mathbf{p}) \mathbf{u}(t)$$

- Dove \mathbf{G} è una matrice **funzione di trasferimento razionale** nell'**operatore di traslazione** q :

$$\mathbf{G}(q, \mathbf{p}) = \frac{\mathbf{B}(q)}{\mathbf{A}(q)} = \frac{\mathbf{b}_1 q^{-n_k} + \mathbf{b}_2 q^{-n_k-1} + \dots + \mathbf{b}_{n_u} q^{-n_k-n_u}}{1 + \mathbf{a}_1 q^{-1} + \dots + \mathbf{a}_{n_\rho} q^{-n_\rho}}$$

- Scritto in modo esplicito:

$$\boldsymbol{\rho}(t) + \mathbf{a}_1 \boldsymbol{\rho}(t-1) + \dots + \mathbf{a}_{n_\rho} \boldsymbol{\rho}(t - n_\rho) = \mathbf{b}_1 \mathbf{u}(t - n_k) + \dots + \mathbf{b}_{n_u} \mathbf{u}(t - n_k - n_u)$$

Tempo di ritardo



La funzione f

- Allo stesso modo è possibile modellare la porzione di disturbo:

$$\omega(t) = \mathbf{H}(q, \mathbf{p}) \mathbf{e}(t)$$

- Dove \mathbf{H} è una matrice **funzione di trasferimento razionale** nell'**operatore di traslazione** q :

$$\mathbf{H}(q, \mathbf{p}) = \frac{\mathbf{D}(q)}{\mathbf{C}(q)} = \frac{\mathbf{d}_1 q^{-1} + \mathbf{d}_2 q^{-2} + \dots + \mathbf{d}_{n_e} q^{-n_e}}{1 + \mathbf{c}_1 q^{-1} + \dots + \mathbf{c}_{n_\omega} q^{-n_\omega}}$$

- Scritto in modo esplicito:

$$\omega(t) + \mathbf{c}_1 \omega(t-1) + \dots + \mathbf{c}_{n_\omega} \omega(t - n_\omega) = \mathbf{d}_1 \mathbf{e}(t-1) + \dots + \mathbf{d}_{n_e} \mathbf{e}(t - n_e)$$

- La formulazione generale del problema diviene quindi:

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



**Formulazione di
BOX-JENKINS**

La funzione f

- Si noti che la formulazione di Box-Jenkins dipende da 5 parametri strutturali: \mathbf{n}_u \mathbf{n}_e \mathbf{n}_ω \mathbf{n}_ρ \mathbf{n}_k
- e da quattro parametri adattivi: \mathbf{a} \mathbf{b} \mathbf{c} \mathbf{d}
- Una semplificazione del modello di Box-Jenkins si ottiene imponendo che:

$$\mathbf{A}(q) = \mathbf{C}(q)$$

$$\mathbf{A}(q)\mathbf{y}(t) = \mathbf{B}(q)\mathbf{u}(t) + \mathbf{D}(q)\mathbf{e}(t) \quad \text{Modello ARMAX}$$

- Si ha allora:
- L'acronimo **ARMAX** deriva da:
 - **A**UTO**R**EGRESSIVE $\mathbf{A}(q)\mathbf{y}(t)$
 - **M**OVING **A**AVERAGE $\mathbf{D}(q)\mathbf{e}(t)$
 - **E**XOGENEOUS INPUT $\mathbf{B}(q)\mathbf{u}(t)$ ← È un "extra input" che in Economia è definito come "Exogeneous Input"

Regressori

- Il vettore dei regressori nel caso di modelli **ARMAX** è:

$$\boldsymbol{\varphi}_{ARMAX}(t) = [\mathbf{y}(t-1), \dots, \mathbf{y}(t-\mathbf{n}_y); \mathbf{u}(t-\mathbf{n}_k-1), \dots, \mathbf{u}(t-\mathbf{n}_k-\mathbf{n}_u); \mathbf{e}(t-1), \dots, \mathbf{e}(t-\mathbf{n}_e)]$$

- Se si esclude il termine di media mobile si ottiene il modello **ARX**:

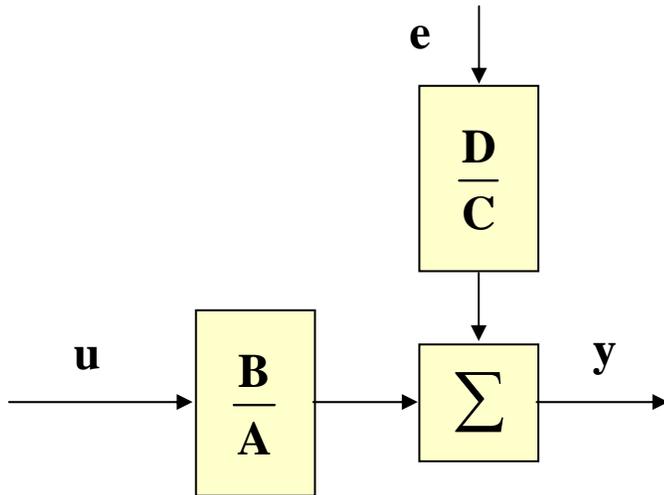
$$\boldsymbol{\varphi}_{ARX}(t) = [\mathbf{y}(t-1), \dots, \mathbf{y}(t-\mathbf{n}_y); \mathbf{u}(t-\mathbf{n}_k-1), \dots, \mathbf{u}(t-\mathbf{n}_k-\mathbf{n}_u)]$$

- Se si parte da un modello ARX e si pone $n_y = 0$ si ottiene un modello **FIR** (**F**inite **I**mpulse **R**esponse):

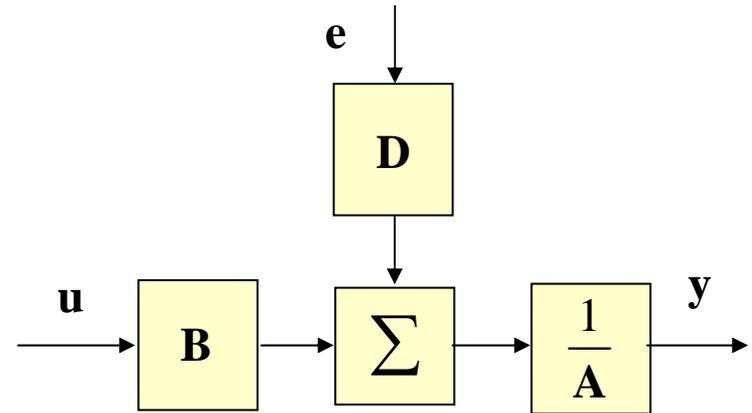
$$\boldsymbol{\varphi}_{FIR}(t) = [\mathbf{u}(t-\mathbf{n}_k-1), \dots, \mathbf{u}(t-\mathbf{n}_k-\mathbf{n}_u)]$$



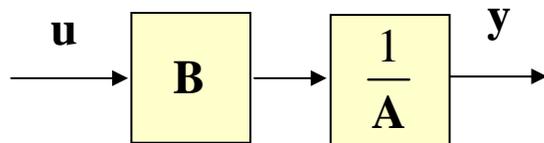
Modelli identificativi



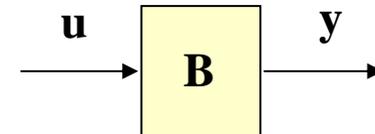
BOX-JENKINS



ARMAX



ARX



FIR

Modelli ARX

- I modelli matematici finora descritti possono avere struttura scalare, vettoriale e mista:
 - **SISO**: Single Input – Single Output
 - **MISO**: Multiple Input – Single Output
 - **MIMO**: Multiple Input – Multiple Output
- **CARATTERISTICHE**
 - Il modello ARX è **lineare** sia nei **regressori** che nei **parametri**.
 - Come tale **non** è in grado di descrivere **più stati stazionari**;
 - Per definizione **non** può descrivere comportamenti **non lineari**;
 - La sua identificazione è relativamente **semplice**;
 - I **tempi di calcolo** per una predizione sono estremamente **ridotti**.
- **ESEMPIO ARX SISO**

$$y(t) + a_1 y(t-1) + a_2 y(t-2) + \dots + a_{n_y} y(t-n_y) = b_1 u(t-1) + b_2 u(t-2) + \dots + b_{n_u} u(t-n_u)$$



Modelli ARMAX

• CARATTERISTICHE

- Un modello ARMAX è **lineare** nelle **variabili** di input, output, errori.
- Il dettaglio di predizione ARMAX è migliore di quello di un modello ARX grazie alla presenza dei termini di errore, **e**;
- Il termine di errore è in grado, per certi versi, di tener conto di:
 - non linearità del processo;
 - disturbi non misurati;
 - rumore delle misure.

- Non è in grado di descrivere **più stati stazionari**;
- Gli output del modello sono calcolati tramite prodotto scalare del vettore dei regressori per quello dei parametri;

$$y(t) = \mathbf{p} \times \boldsymbol{\varphi}(t, \mathbf{p})$$

- Il modello ARMAX è **non lineare** nei **parametri** di regressione.



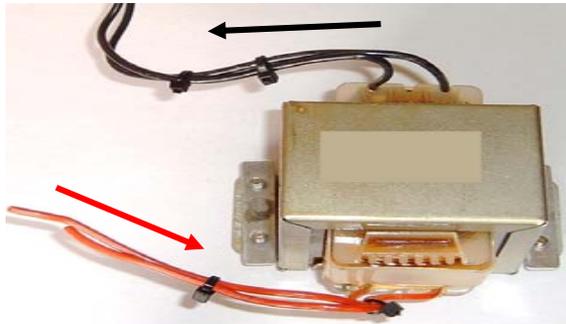
Modelli ARMAX

• CARATTERISTICHE

- La determinazione dei parametri, \mathbf{p} , richiede una procedura di **minimizzazione non lineare** → tempi elevati di CPU.
- Tempi di calcolo per una predizione maggiori rispetto al modello ARX.
- La presenza del termine di errore $\mathbf{e} = \mathbf{y}_{real} - \mathbf{y}$ richiede la misurazione in campo delle variabili di output. Come tale il modello ARMAX non è indicato per effettuare previsioni di comportamento con n passi in avanti.

• ESEMPIO ARMAX SISO

$$y(t) + a_1 y(t-1) + a_2 y(t-2) + \dots + a_{n_y} y(t-n_y) = b_1 u(t-1) + b_2 u(t-2) + \dots + b_{n_u} u(t-n_u) + d_1 e(t-1) + d_2 e(t-2) + \dots + d_{n_e} e(t-n_e)$$



Modelli non lineari



- Riprendendo il concetto di funzione \mathbf{f} intesa come **mappa** del vettore dei regressori $\boldsymbol{\varphi}$ nelle variabili di output \mathbf{y} tramite i parametri \mathbf{p} :

$$\mathbf{y}(t) = \mathbf{f}[\boldsymbol{\varphi}(t), \mathbf{p}]$$

è possibile fare uso di una funzione \mathbf{f} non lineare di qualsivoglia forma.

- Di solito però si adotta una espansione di una funzione base $f_k(\boldsymbol{\varphi}(t))$ tale che:

$$\mathbf{f}[\boldsymbol{\varphi}(t), \mathbf{p}] = \sum_k p_k f_k(\boldsymbol{\varphi}(t))$$

- Una delle forme più usate è l'espansione polinomiale degli elementi del regressore:

$$\mathbf{y}(t) = \sum_{k=1}^N \alpha_k \varphi_k + \sum_{k=1}^N \sum_{l \geq k}^N \beta_k \varphi_k \varphi_l + \sum_{k=1}^N \sum_{l \geq k}^N \sum_{m \geq l}^N \gamma_k \varphi_k \varphi_l \varphi_m + \dots$$

N.B.: tale espansione è **lineare** nei **parametri** ma **non lineare** negli N **regressori**.



Modelli non lineari

- Un'altra alternativa è quella di utilizzare delle funzioni f_k ottenute tramite parametrizzazione di una singola **funzione madre di base**, κ .
- Di solito κ è a sua volta parametrizzata tramite due parametri di natura diversa:
 - β_k che fa riferimento alla scala o a qualche proprietà direzionale;
 - γ_k che fa riferimento ad una posizione o ad una traslazione.



$$f_k(\varphi(t)) = \kappa(\varphi(t), \beta_k, \gamma_k) = \kappa(\beta_k(\varphi - \gamma_k))$$

- Se ad esempio $\kappa = \cos(x)$, l'espansione diviene una **serie di Fourier** ove β_k sono le frequenze e γ_k le fasi:

$$\mathbf{f}[\varphi(t), \mathbf{p}] = \sum_k p_k f_k(\varphi(t)) = \sum_k \alpha_k \kappa(\varphi(t), \beta_k, \gamma_k) = \sum_k \alpha_k \cos(\beta_k(\varphi(t) - \gamma_k))$$

- Un'altra funzione madre di base comunemente usata è la **sigmoide**:

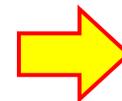
$$\mathbf{f}[\varphi(t), \mathbf{p}] = \sum_k p_k f_k(\varphi(t)) = \sum_k \alpha_k \kappa(\varphi(t), \beta_k, \gamma_k) = \sum_k \alpha_k \frac{1}{1 + e^{(\beta_k(\varphi(t) - \gamma_k))}}$$



Modelli NARX

- In analogia con gli omologhi modelli lineari è possibile introdurre i modelli non lineari: **NARMAX**, **NARX**, **NFIR**.
- Per quanto riguarda i modelli **NARX** si può pensare di costruirli a partire dai vettori di input ed output tramite due mappature consecutive:
 - Dapprima si costruisce una **mappatura polinomiale non lineare**, g , del vettore dei regressori:

$$\eta(t) = g(\varphi(t))$$



- In seguito il nuovo vettore dei regressori η viene **mappato linearmente** nel vettore degli output y :

$$y(t) = \mathbf{p} \times \eta(t)$$

$$\eta_1 = \varphi_1$$

$$\vdots$$

$$\eta_N = \varphi_N$$

$$\eta_{N+1} = \varphi_1 \varphi_1$$

$$\eta_{N+2} = \varphi_1 \varphi_2$$

$$\vdots$$

$$\eta_{N+N} = \varphi_1 \varphi_N$$

$$\eta_{2N+1} = \varphi_2 \varphi_2$$

$$\vdots$$

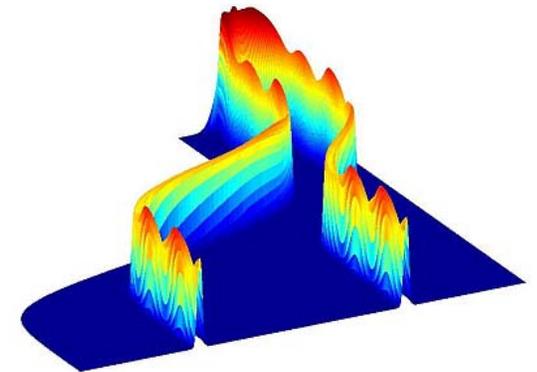
$$\eta_{2N+N-1} = \varphi_2 \varphi_N$$

$$\vdots$$


Modelli NARX

- **CARATTERISTICHE**

- dato che la formulazione è **lineare nei parametri** allora è sufficiente un algoritmo basato sui minimi quadrati per determinare il vettore \mathbf{p} ;
- conseguentemente lo **sforzo calcolistico** dell'identificazione è assolutamente analogo a quello compiuto nel caso di modelli ARX;
- grazie alla non linearità un modello NARX è in grado di descrivere **molteplicità di stati stazionari**;
- **non** richiede **misure in campo** del vettore \mathbf{y} per effettuare delle previsioni in quanto non contempla i termini di errore, \mathbf{e} ;
- un modello NARX, se non opportunamente strutturato, può produrre **output non limitati** con evidente degrado delle capacità predittive del modello.



Modelli NARX

- **ESEMPIO NARX SISO**

- Si consideri come esempio un NARX SISO basato su di un'espansione quadratica polinomiale di un vettore di regressori con quattro elementi (due di input e due di output): $y(t-1), y(t-2), u(t-1), u(t-2)$

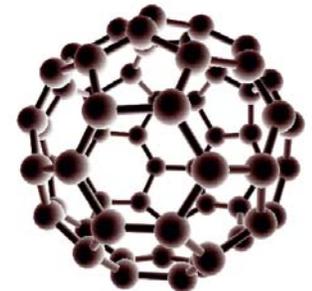
Il sistema NARX risulta essere:

$$y(t) = a_1 y(t-1) + a_2 y(t-2) + a_3 u(t-1) + a_4 u(t-2) + a_5 y(t-1) y(t-1) + a_6 y(t-1) y(t-2) + a_7 y(t-1) u(t-1) + a_8 y(t-1) u(t-2) + a_9 y(t-2) y(t-2) + a_{10} y(t-2) u(t-1) + a_{11} y(t-2) u(t-2) + a_{12} u(t-1) u(t-1) + a_{13} u(t-1) u(t-2) + a_{14} u(t-2) u(t-2)$$

- Il modello è basato su **14** parametri nonostante le ridotte dimensioni.
- Esistono problemi di **sovraparametrizzazione**. Ad esempio una espansione polinomiale cubica con un vettore di regressori composto da 6 elementi condurrebbe ad un modello con **83** parametri.
- Occorre individuare delle tecniche di calcolo in grado di annullare i coefficienti non significativi → algoritmi Stepwise.

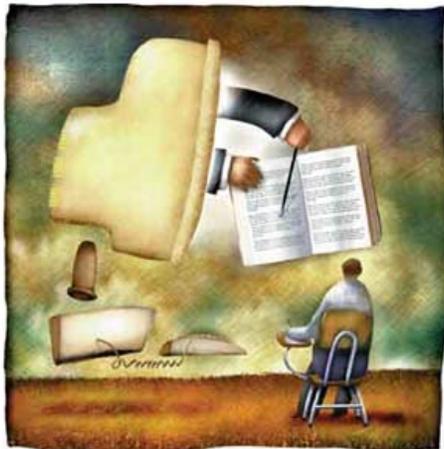
La procedura di identificazione

- **Determinazione dei limiti del sistema e delle variabili necessarie**: in base alle esigenze che hanno condotto alla realizzazione del modello si definisce l'**esatto numero di variabili** di input e di output. Si individua anche il **range di variabilità** delle stesse al fine di creare un opportuno dominio di apprendimento per la fase successiva di identificazione. Le variabili vengono scelte in base alla conoscenza fisica/empirica del processo, indici di correlazione, trial & error.
- **Progettazione della sperimentazione**: una volta individuate le variabili occorre definire la **frequenza di campionamento**. Inoltre è opportuno che tutte le variabili di input siano disturbate. Occorre anche considerare se sia possibile disturbare le variabili che in modo non diretto entrano a far parte del sistema da identificare.
- **Selezione della struttura del modello**: occorre definire la **lunghezza** del vettore dei regressori, l'**ordine** del modello rispetto ad ogni variabile, la **linearità** o **non linearità** del modello rispetto i regressori ed i parametri.



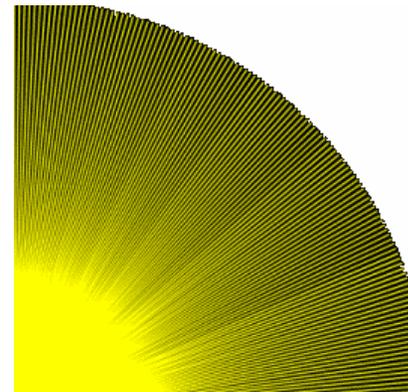
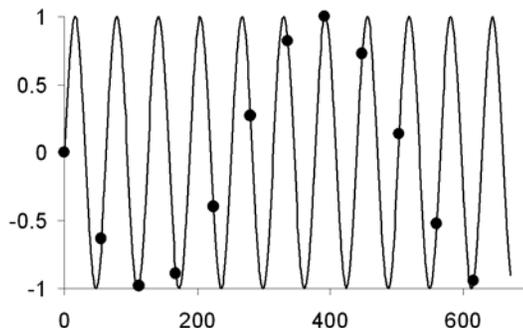
La procedura di identificazione

- **Determinazione dei parametri**: occorre scegliere l'**algoritmo numerico** necessario alla determinazione dei parametri del modello. Occorre distinguere tra modelli **deterministici** (minimizzazione dell'errore commesso) e modelli **stocastici** (metodo della massima verosimiglianza).
- **Simulazione e convalida**: una volta identificato il modello occorre **testare** la sua capacità predittiva e quindi la sua bontà utilizzando un insieme di dati che non siano già stati usati. A tal fine, la procedura di convalida del modello si basa su di un insieme di dati di convalida (**cross-validation set**) opportunamente scelto a priori e mantenuto distinto dall'insieme di addestramento (**learning set**).

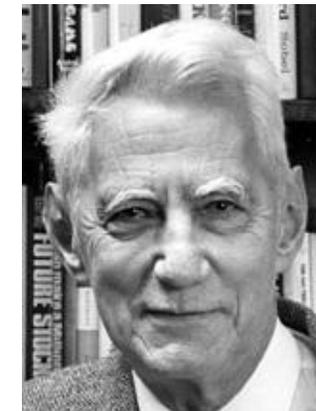


Tempo di campionamento

- Per produrre i due insiemi di dati per l'apprendimento e la verifica occorre definire l'intervallo temporale con il quale effettuare le misure.
- In generale il tempo di campionamento è preso pari ad una frazione, 5%-20% del tempo caratteristico del sistema.
- **Teorema di Shannon**: se si campiona un segnale a frequenza almeno doppia rispetto alla massima frequenza del segnale stesso, è possibile ricostruire, senza perdita di informazione il segnale originale.
- Se il teorema non viene rispettato si può incorrere in problemi di **aliasing**. Non è cioè possibile ricostruire il segnale originario partendo dai dati campionati.

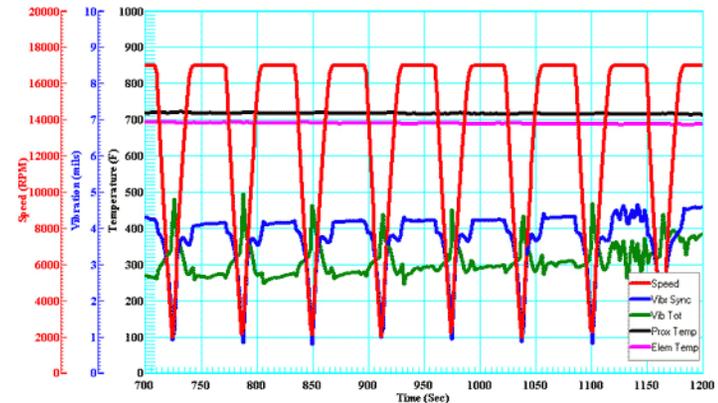


Claude Shannon
1916-2001



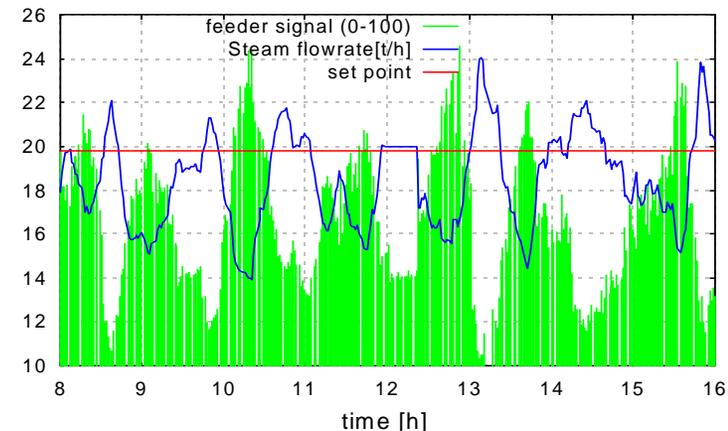
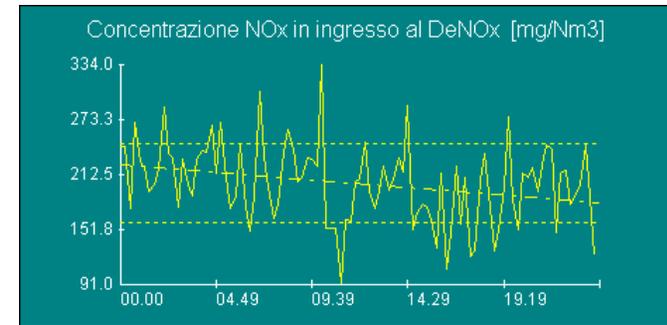
Tempo di campionamento

- Avendo definito t_s il tempo di campionamento del segnale, se questi è **troppo elevato** non è possibile descrivere e quindi identificare l'effettiva dinamica del processo.
- Se t_s è **troppo piccolo**:
 - occorre manipolare un numero eccessivo di dati mentre non si migliora l'informazione in essi contenuta;
 - si rischia di aumentare il campionamento del rumore;
 - dati troppo simili tra loro creano problemi numerici all'algoritmo di identificazione;
 - il tempo di elaborazione aumenta sia in fase di identificazione che in fase di predizione;
 - il costo di acquisizione può aumentare.



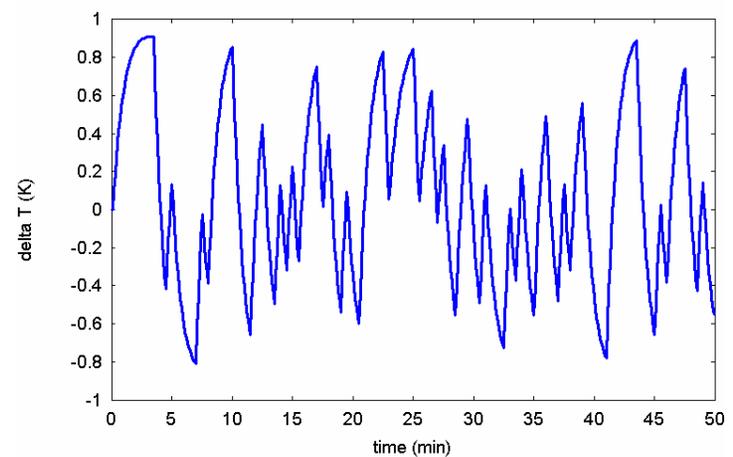
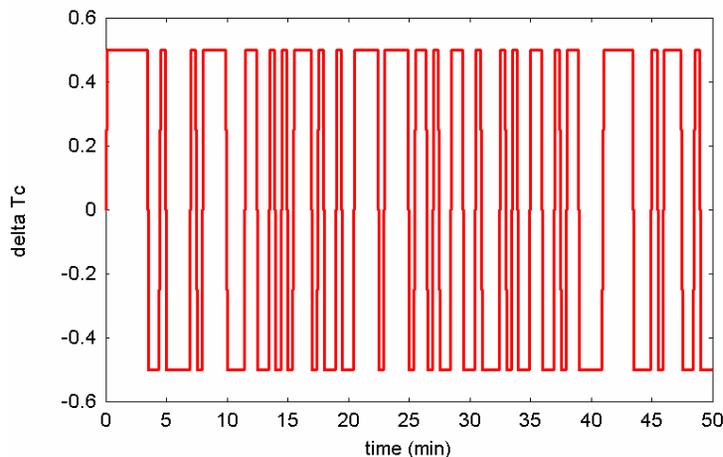
Pretrattamento dei dati

- All'atto dell'acquisizione dei dati dal campo è possibile applicare degli opportuni operatori matematici in grado di smorzare le eccessive oscillazioni (ad esempio **media mobile**).
- È possibile applicare dei **filtri** taglia alto, taglia basso per cancellare variazioni improvvise oltre i normali intervalli di operatività.
- È possibile eliminare i cosiddetti **outlier** tramite opportune tecniche di analisi statistica.
- **DETREND**: ai dati viene tolto il valore medio. In tal modo le variabili campionate esprimono lo scostamento dai valori di stazionario o comunque dalle condizioni operative medie. È così possibile utilizzare il modello (a costo di risultati mediocri) anche per altri stati stazionari.



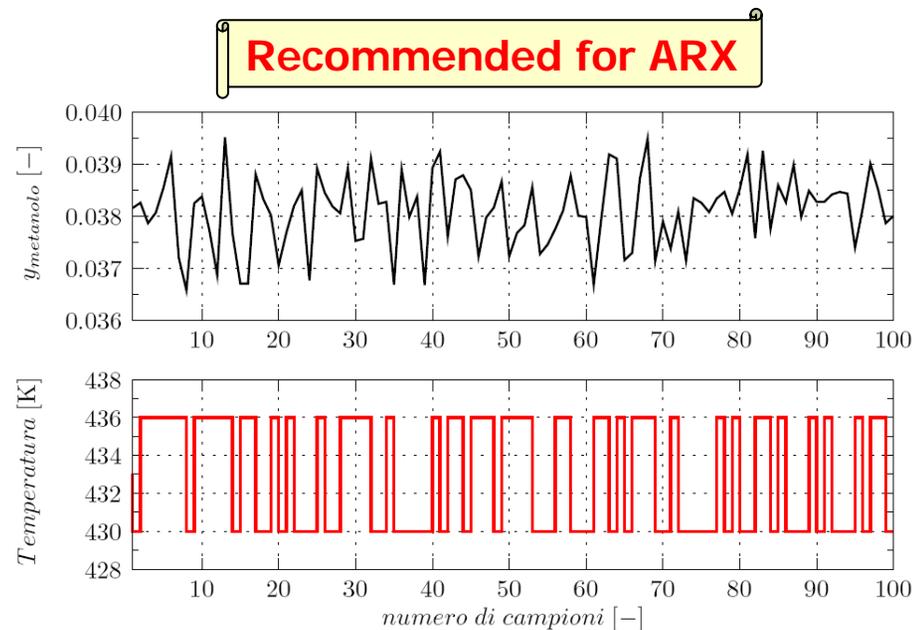
Generazione di sequenze di disturbo

- La raccolta dei dati di input-output per le procedure di identificazione e di convalida avviene disturbando le variabili di input del processo secondo tecniche distinte, volte a coprire adeguatamente il campo operativo del processo studiato.
- **PRBS**: **P**seudo **R**andom **B**inary **S**equence. Si scelgono due estremi di banda, x_{MIN} , x_{MAX} , per la variabilità della grandezza x da disturbare e si varia in modo random il suo valore tra uno dei due estremi con una sequenza binaria ($0 = x_{MIN}$, $1 = x_{MAX}$). Si misura in corrispondenza il vettore di output.



Sequenze PRBS

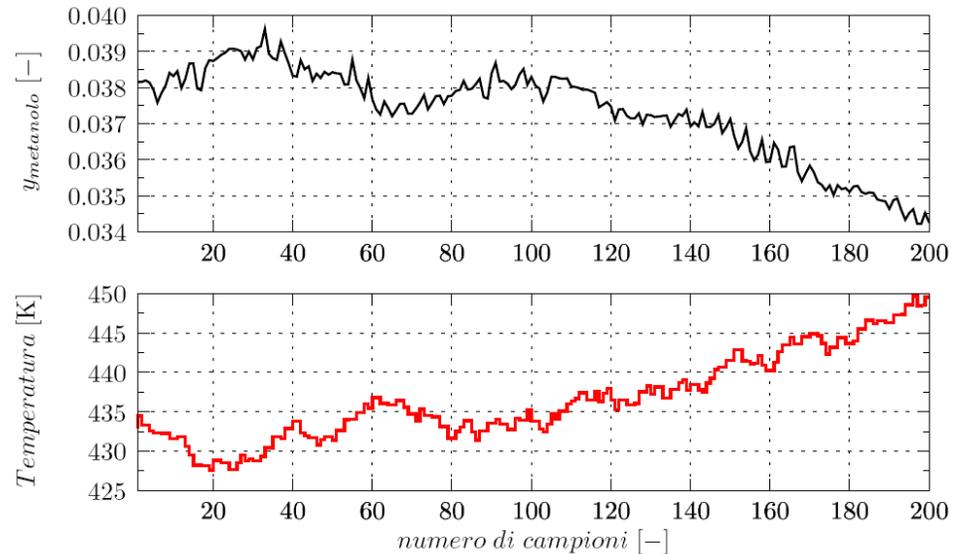
1. Le variabili di input possono assumere **solo due valori**, uguali in ampiezza ma di segno opposto, $\pm\Delta u$, rispetto alle condizioni stazionarie (prese come riferimento 0 a seguito della procedura di "detrend").
2. Il passaggio da una condizione **positiva** a quella **negativa**, e viceversa, avviene in modo **random** al fine di dare alla sequenza la caratteristica di rumore bianco (media nulla).
3. Il disturbo sulle variabili di input viene effettuato ogni n tempi di campionamento, t_s , con una probabilità 0.5 di cambiare segno.
4. Di solito l'intervallo $n \cdot t_s$ è pari al 20% del tempo richiesto al sistema per esaurire il transitorio.
5. L'ampiezza Δu del disturbo deve essere sufficientemente elevata da eliminare i disturbi di misurazione dovuti al rumore del sistema.



Sequenze PRS

- La caratteristica principale che distingue un sistema lineare da uno non lineare è la dipendenza proporzionale del comportamento a seguito dell'ampiezza del disturbo.
- Se l'intensità di un ingresso raddoppia, anche l'ampiezza della risposta raddoppia in un sistema lineare. Non altrettanto in un sistema non lineare.
- In una sequenza **PRS**, **P**seudo **R**andom **S**equence, il valore della variabile di input viene calcolato sommando al valore attuale una quantità random avente distribuzione uniforme nell'intervallo $[-\Delta u, +\Delta u]$.
- Le variabili di input subiscono un disturbo random che è una **frazione costante** in modulo dell'ampiezza massima di variabilità ($\sim 15-20\%$).

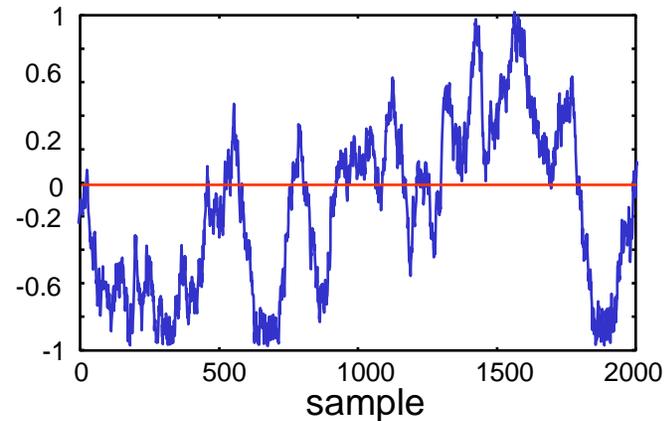
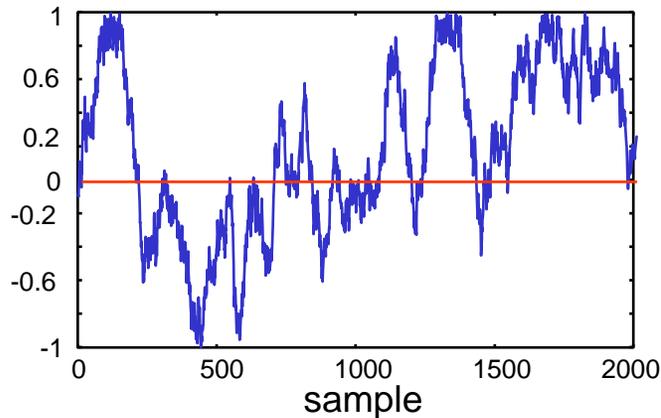
Recommended for ARMAX



Sequenze completamente random

- Un'ulteriore generalizzazione della sequenza PRS è l'assegnazione di incrementi di ampiezza random alle variabili di input all'interno di un prefissato range di variabilità $[-\Delta u, +\Delta u]$.
- Unici accorgimenti sono:
 - limitare il massimo incremento di variazione per singolo disturbo;
 - mantenere le variabili di input all'interno del range di corretta operatività (lower and upper bounds).

Recommended for NARX and ANN

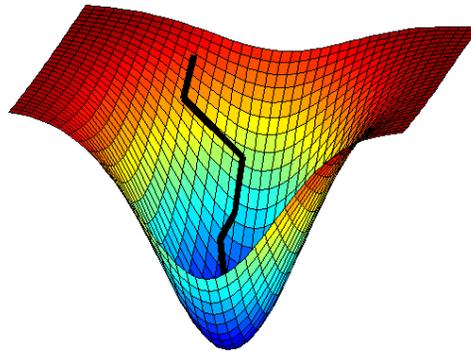
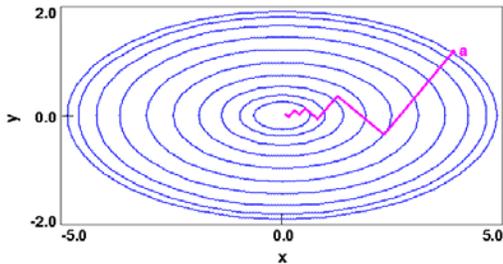


Stima dei parametri

- I parametri \mathbf{p} del modello vengono stimati tramite una procedura di regressione.
- Nel caso di segnali deterministici il problema numerico da risolvere è:

$$\min_{\mathbf{p}} \left\{ \sum_{i=1}^{n_y} \sum_{t=1}^{n_s} \left[y_i^{real}(t) - f_i(\boldsymbol{\varphi}(t), \mathbf{p}) \right]^2 \right\}$$

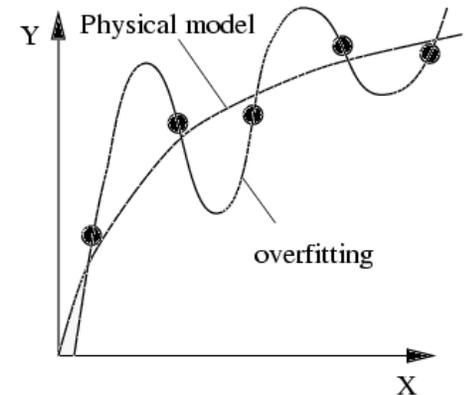
con n_y il numero di variabili di output ed n_s il numero di campionamenti.



Convalida del modello

交差検定

- Dopo aver stimato i parametri occorre verificare la bontà del sistema rispetto un insieme di dati differente da quello utilizzato per l'identificazione.
- La procedura di verifica è definita: convalida (**cross-validation**).
- Occorre verificare il cosiddetto **overfitting** del sistema ovvero una eccessiva specificità del modello rispetto ai dati di apprendimento.
- Si testa anche la capacità di estrapolazione.
- La procedura di convalida può essere effettuata:
 - **One step ahead**: utilizzando di passo in passo gli output del processo reale;
 - **Predictive mode**: utilizzando gli output del sistema per procedere con la simulazione passo-passo senza impiegare i dati provenienti dal processo (**pure extended simulation**).



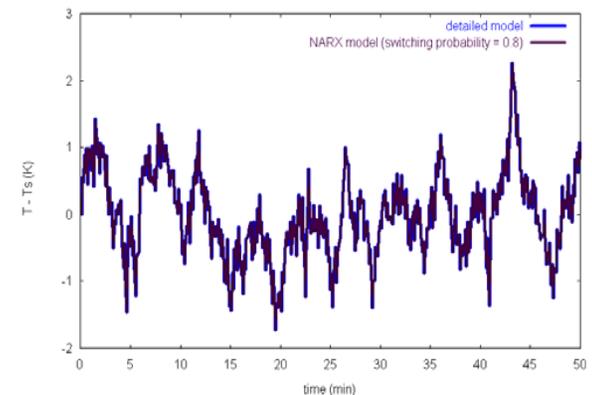
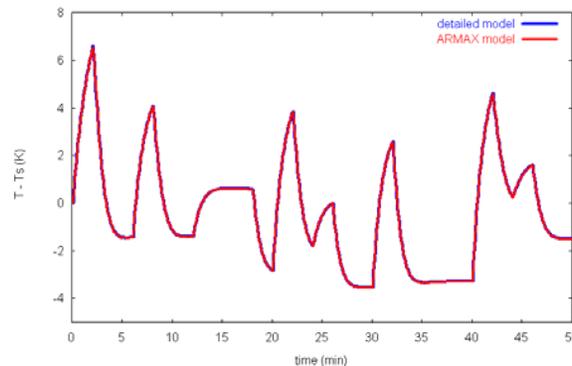
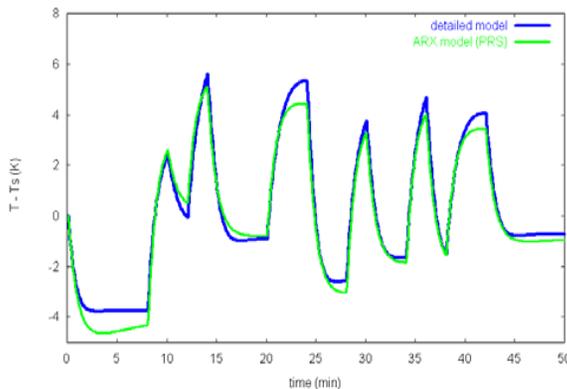
Convalida del modello

- La bontà di un modello è valutata rispetto alla:
 - capacità di riprodurre la **risposta dinamica** del processo reale;
 - stabilità** del modello rispetto a disturbi esterni.
- CROSS VALIDATION INDEX**: permette di stimare la bontà del modello (n_{CVS} = numero di dati per la convalida)



$$CVI = \frac{\sum_{i=1}^{n_{CVS}} \left(y_{real}(i) - y_{system}(i) \right)^2}{\sum_{i=1}^{n_{CVS}} \left(y_{system}(i) - y_{system}^{mean} \right)^2}$$

- Più l'indice *CVI* è prossimo a zero, migliore è la convalida del modello.



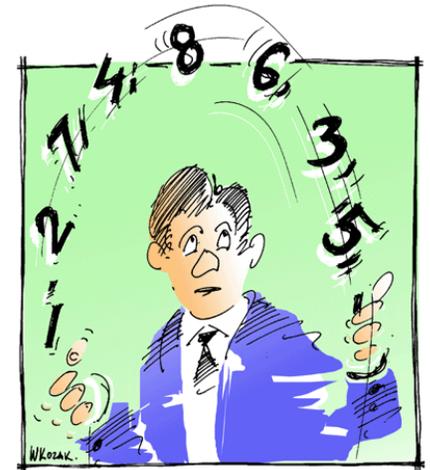
Confronto tra modelli

- Per individuare l'ordine ottimale di un modello (ed evitare il fenomeno dell'overfitting) o effettuare confronti tra modelli aventi struttura diversa (ARX, ARMAX, NARX, ...) è possibile adottare dei **criteri statistici** che considerano e bilanciano attributi quali:
 - **ordine del modello;**
 - **numero di variabili di input e output;**
 - **numero di parametri.**
- I quattro criteri di confronto più noti sono:
 - **AKAIKE INFORMATION CRITERION:**

$$AIC = n_s \log \left(\frac{1}{n_s} \sum_{i=1}^{n_s} (e(i))^2 \right) + 2n_p$$

- **FINAL PREDICTION ERROR CRITERION:**

$$FPE = n_s \log \left(\frac{1}{n_s} \sum_{i=1}^{n_s} (e(i))^2 \right) + n_s \log \left(\frac{n_s + n_p}{n_s - n_p} \right)$$



Confronto tra modelli

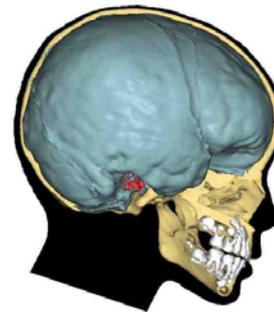
- **BAYESIAN INFORMATION CRITERION:**

$$BIC = n_s \log \left(\frac{1}{n_s} \sum_{i=1}^{n_s} (e(i))^2 \right) + n_p \log(n_s)$$

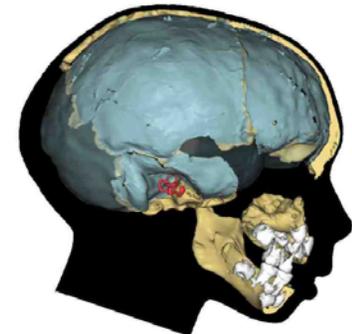
- **LAW OF ITERATED LOGARITHMS CRITERION:**

$$LILC = n_s \log \left(\frac{1}{n_s} \sum_{i=1}^{n_s} (e(i))^2 \right) + 2n_p \log(\log(n_s))$$

n_s numero di campionamenti, n_p numero di parametri ed $e(i) = y_{real}(i) - y_{system}(i)$



Neanderthal



Sapiens

Bibliografia

- Billings S.A., W.S.F. Woon, *"A Prediction – Error and Stepwise – Regression Estimation Algorithm for Non-linear Systems"*, Int. J. Control, Vol 44, no 3, pp 803-822, 1986
- Bosch P.P.J, A.C. van der Klauw, *"Modeling, Identification, and Simulation of Dynamical Systems"*, CRC Press, 1994
- Henson M.A, D.E. Seborg, *"Nonlinear Process Control"*, Prentice Hall, New Jersey, 1997
- Hernández E., Y. Arcun, *"Control of Non-linear Systems using Polynomial ARMA Models"*, AIChE, Vol 39, no 3, pp 446-460, 1993
- Hernando D., A. A. Desrochers, *"Modeling of Non-linear Discrete Time Systems from Input–Output Data "*, Automatica, Vol 24, no 5, pp 629-641, 1988
- Juditsky et. al, *"Nonlinear Black-box Modeling in System Identification: Mathematical Foundations"*, Automatica, Vol 31, no 12, pp 1725-1750, 1995
- Leontaritis I.J., S.A. Billings, *"Input-output Parametric Models for Non-linear Systems. Part I: Deterministic Non-linear Systems"*, Int. J. Control, Vol 41, no 2, pp 303-328, 1985
- Leontaritis I.J., S.A. Billings, *"Input-output Parametric Models for Non-linear Systems. Part II: Stochastic Non-linear Systems"*, Int. J. Control, Vol 41, no 2, pp 329-344, 1985
- **Ljung L., "System Identification: Theory for the User", Prentice Hall, 1998**
- Shannon C.E. , *"A Mathematical Theory of Communication"*, The Bell System Technical Journal, Vol 27, 379–423, 623–656, 1948
- Sjöberg J., et al., *"Nonlinear Black-box Modeling in System Identification: a Unified Overview"*, Automatica, Vol 31, no 12, pp 1691-1724, 1995
- Srinivas G.R., et. Al, *"Non-linear Identification and Control of a High-purity Distillation Column"*, J. Proc. Cont., Vol 5, no 3, pp 149-162, 1995

