

# Sistemi e Modelli

Identificazione, a.a. 2011-2012

Prof. Luca Schenato

Dipartimento di Ingegneria dell'Informazione  
Università degli Studi di Padova

6 Dicembre 2011

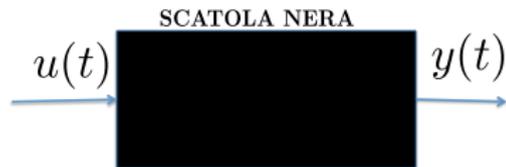
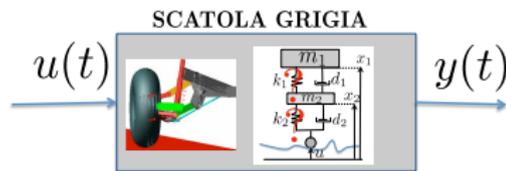
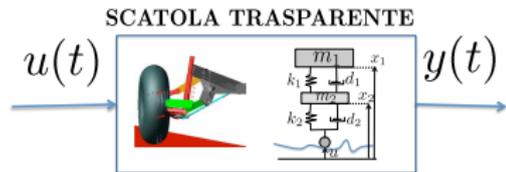


DEPARTMENT OF  
INFORMATION  
ENGINEERING  
UNIVERSITY OF PADOVA



# Come si costruisce un modello matematico?

- **Scatola trasparente:** si usano leggi costruttive del fenomeno (fisiche, biologiche, chimiche, etc) per derivare il modello matematico
- **Scatola grigia:** si usano leggi costruttive del fenomeno e misure sperimentali  $\{u(t), y(t)\}$  per stimare alcuni parametri non noti (massa, costante di degradazione, etc..)
- **Scatola nera (Modelli di dati):** si usano solamente delle misure sperimentali di  $\{u(t), y(t)\}$



# Identificazione (Modelli di dati, Data-based models)

## Definizione

Propone **algoritmi**, spesso basati sulla **statistica**, per la costruzione automatica di **modelli matematici** di sistemi statici e dinamici a partire da **dati sperimentali**, generalmente un set di uscite  $\{y_i\}_{i=1}^N$  ed ingressi  $\{u_i\}_{i=1}^N$ .

- **Diversi obiettivi:**
  - predizione
  - simulazione
  - stima di parametri fisici non direttamente osservabili (costanti elastiche, coefficienti di flusso, ecc.)
  - controllo in retroazione

# Approccio

- 1 **Scelta classe dei Modelli:** scelta della classe dei possibili modelli: lineari/non-lineari, parametrici/non-parametrici, deterministici/stocastici, tempo discreto/continuo, statico/dinamico

$$y_i = \theta + v_i, \quad v_i \sim \mathcal{N}(0, \sigma^2), \quad i = 1, \dots, N \text{ i.i.d.}$$

- 2 **Scelta Stimatore:** criterio/funzionale di costo per la scelta del modello "ottimo" all'interno della classe

$$\hat{\theta} = \operatorname{argmin}_{\hat{\theta}} E[(\hat{\theta} - \theta)^2]$$

- 3 **Validazione Stimatore:** verifica a-posteriori se stimatore ottenuto è consistente con classe del modello.

residui  $e_i = y_i - \hat{\theta}_i$  sono rumore bianco (i.i.d, media nulla)?

- 4 **Scelta tra stimatori:** compromesso fra complessità e prestazione (rasoio di Occam)

$$y_i = \theta u_i + v_i \quad \text{oppure} \quad y_i = \theta_1 u_i + \theta_2 u_i^2 + \theta_3 e^{u_i} + v_i ?$$

# Considerazioni generali

- Gli esperimenti sono **irripetibili**: impossibile ottenere esattamente lo stesso set di dati  $\{y_i\}_{i=1}^N$ ,  $\{u_i\}_{i=1}^N$ , ma **presentano una certa continuità/ripetitività**: il modello che genera i dati ha “memoria” (vedi concetto di **stazionarietà**)
- **Modelli deterministici** ingresso/uscita sono spesso **inadeguati** e sensibili alle incertezze: rumori/disturbi di processo ( $u$ ), rumore/disturbi di misura ( $y$ ), modello vero non necessariamente appartenente alla classe scelta.
- **Approccio statistico**: i dati futuri continueranno ad essere generati dallo stesso **meccanismo stocastico** che ha prodotti i dati osservati.  
Vantaggi:
  - modello probabilistico tiene conto di incertezza sulla scelta della classe di modelli scelta
  - permette di identificare solo i fenomeni rilevanti mentre il resto si manifesta come “rumore”, anche nel caso di sistemi perfettamente deterministici

# Regole d'oro

- Tutti i modelli sono sbagliati, ma alcuni sono utili
- Verificare a posteriori bontà dello stimatore (es: i residui sono bianchi ?)
- Usare per la validazione (se possibile) set di dati differenti da quelli usati per ottenere lo stimatore
- Uso ragionato dell' informazione a priori sul sistema (es: a volte meglio un modello lineare rispetto a non-lineare)
- modello deve sempre poter essere falsificato da nuovi dati
- rasoio di Occam: utilizzare il modello più semplice che raggiunge lo scopo desiderato o la stessa prestazione

Non esiste una ricetta universale per l'identificazione. L'identificazione fornisce alcuni strumenti che devono poi essere usati in maniera consapevole e critica dall'operatore.

# Programma seconda parte del corso

- **Introduzione a identificazione (Cap.6 Cobelli-Carson):**
  - motivazioni, scopo, obiettivi, applicazioni, approccio, etc
  - ingredienti: classe del modello, funzionale di prestazione, calcolo stimatore "ottimo" all'interno della classe, validazione modello,
  - modello: modello piu' semplice che soddisfa lo scopo
  - regole d'oro
- **Richiami di analisi multivariabile (ANALISI 2):**
  - funzioni multivariabili, derivate parziali, gradiente, Hessiano, interpretazione geometrica
  - norma vettore, prodotto interno, traccia, forme quadratiche,
  - matrici semidefinite positive, diagonalizzazione, ellissoide associato, minimo di funz. quad. definita positiva e per semidefinita positiva, radice di una matrice semidefinita positiva
  - somma di quadratiche, minimo corrispondente, prodotto matrici come somma vettori

# Programma seconda parte del corso

## ● Richiami di teoria della probabilità (ANALISI DEI DATI)

- Richiamo variabili aleatorie: definizione, densità di probabilità, funzione cumulativa, variabili multidimensionali ( $y \in R^m, \theta \in R^n$ , numero campioni  $N$ ).
- densità parametriche e non parametriche
- v.a. gaussiana: densità per il caso scalare, media e varianza, funzione cumulativa, caso multivariabile, interpretazione geometrica con ellisse di incertezza, intervallo di confidenza.
- v.a.  $\chi(n)$ : densità, media e varianza, legame con v.a. gaussiane (somma di quadrati)
- probabilità condizionata; regola di Bayes; distinzione fra variabile aleatoria, realizzazione, misura, funzione di variabili aleatorie, parametri, densità di una variabile aleatoria,
- teorema limite centrale, legge dei grandi numeri, disuguaglianza di Chebyshev, media campionaria, numero di campioni ed intervallo di confidenza

## Programma seconda parte del corso

- Richiami di statistica (ANALISI DEI DATI ?)
  - campionamento, definizione di statistica, sufficiente, completa
  - stimatore: definizione, errore, errore quadratico medio, deviazione di campione, bias, varianza, deviazione standard, approccio Fisheriano e Bayesiano
  - stimatori unbiased (UE), stimatori a minima varianza (MVE), stimatori unbiased a minima varianza (MVUE) uniformi. stimatori lineari unbiased (BLUE).
  - Discussione sull'utilità in certi contesti di stimatori biased (vedi deconvoluzione)
  - proprietà asintotiche: consistenza, normalità asintotica, efficienza (asintotica)
  - matrice di informazione, disuguaglianza cramer-rao,  $UE + (\text{var} = \text{cramer-rao}) = \text{MVUE}$  ()
  - condizioni MVUE: Teo.mi di Lehmann-Scheffe' e Rao-Blackwell (?)
  - stimatori massima verosimiglianza (MV), massimo a posteriori (MAP), Bayesiani (B)
  - proprietà stimatori MV: invarianza, proprietà asintotiche, relazione con MAP

# Programma seconda parte del corso

- **Minimi quadrati (Cap.8 Cobelli-Carson)**
  - definizione del problema dei minimi quadrati: definizione di modello e funzionale di prestazione/costo
  - minimi quadrati pesati
  - problema di modellizzazione: modello non-lineare che diventa lineare tramite cambio di variabile, modello lineare nei parametri ma non-lineare negli ingressi (radial basis functions), etc..
  - minimi quadrati non-lineari e algoritmo di Gauss-Newton
- **Sistemi lineari statici con rumore gaussiano (Cap.8 Cobelli-Carson)**
  - stima MV della media e della varianza
  - equivalenza tra stima MV nei modelli lineari e problema MQ: teorema Gauss-Markov
  - proprietà stimatore MV nei modelli lineari: correttezza, cramer-rao, distribuzione, interpretazione geometrica
  - proprietà stimatore MV della varianza (?): distribuzione  $\chi$
  - stima Bayesiana, equivalenza con MAP per densità simmetriche e unimodali

# Programma seconda parte del corso

- **Validazione e scelta del modello (Cap.8 Cobelli-Carson)**
  - trade-off bias-variance, overfitting, underestimation, rasoio di Occam
  - bianchezza dei residui
  - correlogramma
- **Identificabilità parametri in modelli dinamici (Cap.7 Cobelli-Carson)**
  - motivazioni, applicazioni
  - problema di tipo algebrico
  - esempi modelli compartimentali ed altri
- **Deconvoluzione (Cap.9 Cobelli-Carson)**
  - definizione del problema dei minimi quadrati: definizione di modello e funzionale di "fit"
  - minimi quadrati pesati
  - problema di modellizzazione: modello non-lineare che diventa lineare tramite cambio di variabile, modello lineare nei parametri ma non-lineare negli ingressi (radial basis functions), etc..
  - minimi quadrati non-lineari e algoritmo di Newton-Raphson