Teoria dei Sistemi e Controllo ottimo (TSC)

Docente: Giulia Michieletto

Stima dello Stato & Controllo Adattativo

Corso di Laurea Magistrale in Ingegneria Meccatronica

A.A. 2023-2024

Introduzione alla
Stima dello Stato

- problema di stima dello stato
- principali nozioni di teoria della probabilità
- introduzione ai processi stocastici

problema di stima dello stato di un sistema dinamico

⚠ $\mathbf{z}_k = \mathbf{z}(k), \quad \mathbf{z}^k = \{\mathbf{z}_0 \dots \mathbf{z}_k\}$

sistema LTI a tempo discreto

$$\mathbf{x}(t+1) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t)$$
$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t)$$
$$\mathbf{x}(0) = \mathbf{x}_0$$

$\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^p$

$\mathbf{w} \in \mathbb{R}^n$: errore di processo

$\mathbf{v} \in \mathbb{R}^p$: errore di misura

stima dello stato: inferire informazioni sullo stato $\mathbf{x}_t$ dalle le osservazioni $\mathbf{y}^{t'}$

- *predizione* se $t > t'$
- *filtraggio* se $t = t'$
- *regolarizzazione* se $t < t'$

la determinazione di $\mathbf{x}_t$ sarebbe banale se non fosse per la presenza degli errori
$\rightarrow$ sfruttare le equazioni di stato **&** le equazioni di misura

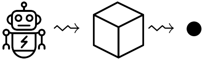
Approccio alla Ricerca

Problema

Formulazione matematica → Analisi/Sintesi → Validazione

Implementazione

equazioni dinamiche



- **come modellare e trattare l'errore di processo e l'errore di misura?**

  nozioni di teoria della probabilità applicate ai processi stocastici

Introduzione alla
Stima dello Stato

- problema di stima dello stato
- principali nozioni di teoria della probabilità
- introduzione ai processi stocastici

probabilità (d'evento)

$E_i,\ i, \in \{1 \dots n\}$ — *i*-esimo evento

$Pr(E_i) \in [0,1], \sum_{i=1}^{n} Pr(E_i) = 1$ — probabilità dell'*i*-esimo evento

$Pr(E_i E_j),\ i,j \in \{1 \dots n\}, i \neq j$ — *probabilità congiunta*

$Pr(E_i E_j) = Pr(E_i)Pr(E_j)$ — probabilità congiunta di eventi indipendenti

$Pr(E_i + E_j),\ i,j \in \{1 \dots n\}, i \neq j$ — *probabilità composta*

$Pr(E_i + E_j) = Pr(E_i) + Pr(E_j)$ — probabilità composta di eventi mut. esclusivi

$Pr(E_i | E_j),\ i,j \in \{1 \dots n\}, i \neq j$ — *probabilità condizionata*

$Pr(E_i | E_j) = Pr(E_i E_j)/Pr(E_j)$ — probabilità condizionata di eventi non indipendenti

variabili e vettori aleatori

variabile aleatoria (casuale o stocastica): variabile che puo' assumere valori diversi in corrispondenza di altrettanti eventi che costituiscono una partizione dello spazio delle probabilità

dati E spazio misurabile & $(\Omega, \mathcal{F}, \nu)$ spazio di probabilità

- Ω spazio campionario o insieme degli eventi
- $\mathcal{F}$ σ-algebra su Ω
- ν misura di probabilità

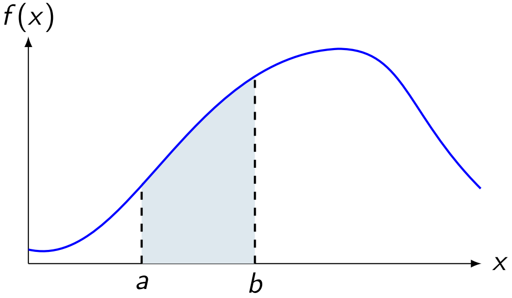
variabile aleatoria X è una funzione misurabile $X\colon \Omega \to E$

- variabili aleatorie continue
- variabili aleatorie discrete

variabili aleatorie continue

funzione di densità di probabilità

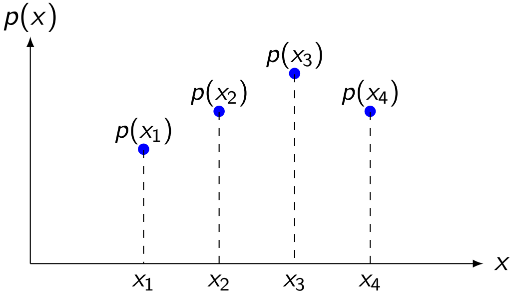
$f(x) = \frac{dF(x)}{dx}, \; F(x) = Pr(X \leq x)$



variabili aleatorie discrete

funzione di probabilità

$p(x) = Pr(X = x)$



proprietà stocastiche

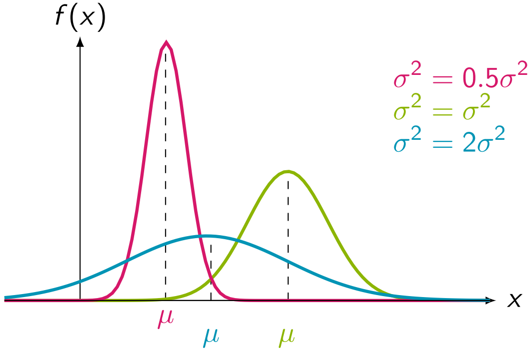
media (valore atteso): numero che formalizza l'idea euristica di valore medio di un fenomeno aleatorio

- variabili aleatorie continue: $\mu_X = \mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx$
- variabili aleatorie discrete: $\mu_X = \mathbb{E}[X] = \sum_{i=1}^{\infty} x_i \, p(x_i)$

varianza: funzione che fornisce una misura della variabilità dei valori assunti dalla variabile stessa

$$\sigma_X^2 = \mathbb{E}\left[\left(X - \mathbb{E}[X]\right)^2\right]$$

esempio classico



distribuzione gaussiana $\mathcal{N}(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

probabilità congiunta: distribuzione di probabilità associata al vettore $\begin{bmatrix} X & Y \end{bmatrix}$ con X e Y variabili aleatorie definite sullo stesso spazio di probabilità

- variabili aleatorie continue: $Pr(\{X \leq x\} \cup \{Y \leq y\}) = F_{XY}(x, y) \;\rightarrow\; f_{XY}(x, y)$
- variabili aleatorie discrete: $Pr(\{X = x\} \cup \{Y = y\}) = p_{XY}(x, y)$

probabilità condizionata di Y dato X: date due variabili aleatorie X e Y, probabilità di Y quando è conosciuto il valore assunto da X

- variabili aleatorie continue: $f_Y(y|X = x) = f_{XY}(x, y)/f_X(x)$
- variabili aleatorie discrete: $p_Y(y|X = x) = p_{XY}(x, y)/p_Y(y)$

vettore aleatorio (variabile aleatoria multivariata): vettore di variabili aleatorie

media (valore atteso): vettore dei valori attesi delle variabili aleatorie componenti il vettore aleatorio

$$\mathbf{m}_X = \mathbb{E}[X] = \begin{bmatrix} \mathbb{E}[X_1] & \dots & \mathbb{E}[X_n] \end{bmatrix}^\top$$

matrice di varianza (o di covarianza): matrice che fornisce una misura di variabilità congiunta per ogni coppia di variabili aletorie componenti il vettore aleatorio

$$\mathbf{\Sigma}_X = \begin{bmatrix} \mathbb{E}\left[(X_1 - \mathbb{E}[X_1])(X_1 - \mathbb{E}[X_1])^\top\right] & \dots & \mathbb{E}\left[(X_1 - \mathbb{E}[X_1])(X_n - \mathbb{E}[X_n])^\top\right] \\ \vdots & & \vdots \\ \mathbb{E}\left[(X_n - \mathbb{E}[X_n])(X_1 - \mathbb{E}[X_1])^\top\right] & \dots & \mathbb{E}\left[(X_n - \mathbb{E}[X_n])(X_n - \mathbb{E}[X_n])^\top\right] \end{bmatrix}$$

Introduzione alla
Stima dello Stato

- problema di stima dello stato
- principali nozioni di teoria della probabilità
- introduzione ai processi stocastici

definizione di processo stocastico

processo stocastico $\{X_t, t \in T\}$: collezione di variabili aleatorie contraddistinte dal parametro $t \in T$ ($T \subseteq \mathbb{R}$ o $T \subseteq \mathbb{N}$, *tempo* continuo o discreto)

$\rightarrow$ fissata una certa distribuzione di probabilià $\mathcal{X}$ per X_t,
il processo stazionario di riduce a un segnale $x(t) \sim \mathcal{X}$
(*realizzazione del processo stocastico*)

- processi stocastici a **stati continui**/discreti
- processi stocastici a **tempo continuo**/discreto

caratterizzazione di un processo stocastico

Una variabile aleatoria (continua) X è caratterizzata dalla sua funzione di densità di probabilità $f_X(\cdot)$

Un processo stocastico (a tempo e stati continui) $\{X_t, t \in T\}$ è completamente caratterizzato quando è assegnata la densità di probabilità congiunta $f(x(t_1), \ldots, x(t_n))$ di un generico numero finito $n \geq 1$ (*ordine del processo stocastico*) di valori $x(t_i) = X_{t_i}, i \in \{1 \ldots n\}$ per ogni possibilie scelta di $t_1 \ldots t_n$

- **funzione di media** $m_X(t) = \mathbb{E}[X_t]$
- **funzione di varianza** $\Sigma_X(t) = \mathbb{E}\left[(X_t - \mathbb{E}[X_t])^2\right]$
- **funzione di auto-covarianza** $R_X(t, t') = \mathbb{E}\left[(X_t - \mathbb{E}[X_t])\left(X_{t'} - \mathbb{E}[X_{t'}]^\top\right)\right]$

processo indipendente e identicamente distribuito (i.i.d.)
$f(x(t_1), \ldots, x(t_n)) = \prod_{i=1}^{n} f(x(t_i))$

processo stazionario (in senso forte)
$f(x(t_1), \ldots, x(t_n)) = f(x(t_1 + k), \ldots, x(t_n + k))$

processo stazionario (in senso debole)
$m_X(t) = m_X$ (primo ordine) o $R_X(t, t') = R_X(t' - t)$ (secondo ordine)

processo bianco
$R_X(t, t') = 0 \; \forall t \neq t'$

processo Gaussiano stazionario bianco e a media nulla:
processo stocastico $\{X_t, t \in T \subseteq \mathbb{R}\}$ tale che $x(t) \sim \mathcal{N}(m_X(t), \Sigma_X(t))$ e

- $m_X(t) = \mathbb{E}[x(t)] = 0 \; \forall t$
- $\Sigma_X(t) = \mathbb{E}[x(t)x(t)^\top] = \Sigma_X \; \forall t$
- $R_X(t, t') = \mathbb{E}[x(t)x(t')^\top] = 0 \; \forall t \neq t'$

processo Gaussiano: modello di errore di processo & misura

▶ praticamente perchè ha senso modellare gli errori tramite processo Gaussiano?

teorema del limite centrale:
data una variabile casuale X distribuita secondo una funzione f_X, se f_X ha un secondo momento finito, allora data un'altra variabile casuale Y definito come la media di molte istanze di X, ovvero $Y \equiv \frac{1}{N}\sum_{i=1}^{N} X_i$ la distribuzione f_Y è gaussiana

molti processi fisici sono la somma di processi più piccoli

Stima ai Minimi Quadrati

formulazione del problema

- stima ai minimi quadrati lineari
  - soluzione deterministica
  - soluzione ricorsiva
  - soluzione pesata
- introduzione del forgetting factor (FF)
- interpretazione geometrica e probabilistica

stima dello stato iniziale

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t) \\ y(t) &= \mathbf{h}^\top \mathbf{x}(t) + v(t) \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{aligned} \qquad \begin{aligned} &\mathbf{x} \in \mathbb{R}^n,\ \mathbf{u} \in \mathbb{R}^m \\ &\mathbf{y} \in \mathbb{R}^p \quad (p = 1) \end{aligned}$$

sistema osservabile

- $\ell \geq n$ **dati di uscita** (misure) & $\mathbf{w}(t) = \mathbf{0},\ v(t) = 0\ \forall t \in [0, \ell - 1]$

$$y_t = \mathbf{h}^\top \mathbf{F}^t \mathbf{x}_0 + \sum_{k=0}^{t-1} \mathbf{h}^\top \mathbf{F}^{t-1-k} \mathbf{G}\mathbf{u}_k \quad \text{con} \quad y_t = y(t),\ \mathbf{u}_k = \mathbf{u}(k)$$

$$\mathbf{y}^\ell = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{\ell-1} \end{bmatrix} = \begin{bmatrix} \mathbf{h}^\top \\ \mathbf{h}^\top \mathbf{F} \\ \vdots \\ \mathbf{h}^\top \mathbf{F}^{\ell-1} \end{bmatrix} \mathbf{x}_0 + \begin{bmatrix} \mathbf{0} & \dots & \dots & \mathbf{0} \\ \mathbf{h}^\top \mathbf{G} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & & & \vdots \\ \mathbf{h}^\top \mathbf{F}^{\ell-2}\mathbf{G} & \dots & \mathbf{h}^\top \mathbf{F}\mathbf{G} & \mathbf{h}^\top \mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{\ell-1} \end{bmatrix} = \mathcal{O}_\ell \mathbf{x}_0 + \mathbf{Z}_\ell \mathbf{u}^\ell$$

se $\mathbf{z}^\ell = (\mathbf{y}^\ell - \mathbf{Z}_\ell \mathbf{u}^\ell) \in Im(\mathcal{O}_\ell)$
allora il sistema $\mathbf{y}^\ell = \mathcal{O}_\ell \mathbf{x}_0 + \mathbf{Z}_\ell \mathbf{u}^\ell$ ammette un'unica soluzione $\mathbf{x}_0 \in \mathbb{R}^n$

in generale, gli errori di processo e misura sono **non nulli**
e quindi $(\mathbf{y}^\ell - \mathbf{Z}_\ell \mathbf{u}^\ell) \notin Im(\mathcal{O}_\ell)$

lo stato iniziale si può stimare con il **metodo dei minimi quadrati**

stimatore ai minimi quadrati

modello dei dati

$$y_i = s_i(\mathbf{x}) + v_i, \; i \in \{1 \dots \ell\} \quad \text{con} \quad s_i(\cdot) : \mathbb{R}^n \to \mathbb{R}$$

determinare $\mathbf{x}$ in modo da approssimare i dati osservati
ottimizzando un certo indice di ‘somiglianza’

stima ai minimi quadrati (o LS - least squares)

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathcal{X}} \sum_{i=1}^{\ell} (y_i - s_i(\mathbf{x}))^2$$

$s_i(\mathbf{x})$ (possibili) funzioni complicate di $\mathbf{x}$ $\Longrightarrow$ calcolo di $\hat{\mathbf{x}}$ non banale

- ipotesi di linearità

modello di regressione lineare

$$y_i = s_{i,1}x_1 + \cdots + s_{i,n}x_n + v_i, \; i \in \{1 \ldots \ell\}$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_\ell \end{bmatrix} = \begin{bmatrix} s_{1,1} & \cdots & s_{1,n} \\ \vdots & & \vdots \\ s_{\ell,1} & \cdots & s_{\ell,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_\ell \end{bmatrix} \quad \Leftrightarrow \quad \mathbf{y} = \mathbf{S}\mathbf{x} + \mathbf{v}$$

$\mathbf{S} \in \mathbb{R}^{\ell \times n}$
matrice di regressione

stimatore ai minimi quadrati lineare

stima ai minimi quadrati lineare (o LLS - linear least squares)

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}} \|\mathbf{y} - \mathbf{S}\mathbf{x}\|^2 = \arg\min_{\mathbf{x}\in\mathcal{X}} \mathbf{v}^\top \mathbf{v}$$

Stima ai Minimi Quadrati

formulazione del problema

- stima ai minimi quadrati lineari
  - soluzione deterministica
  - soluzione ricorsiva
  - soluzione pesata
- introduzione del forgetting factor (FF)
- interpretazione geometrica e probabilistica

stima ai minimi quadrati lineare

stima ai minimi quadrati lineare (o LLS - linear least squares)

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}} \|\mathbf{y} - \mathbf{S}\mathbf{x}\|^2 = \arg\min_{\mathbf{x}\in\mathcal{X}} \mathbf{v}^\top \mathbf{v}$$

caso multivariato ($\mathbf{x} \in \mathbb{R}^n$)

$J(\mathbf{x}) = \|\mathbf{y} - \mathbf{S}\mathbf{x}\|^2$

$$\begin{aligned}
\nabla_{\mathbf{x}} J(\mathbf{x}) &= \nabla_{\mathbf{x}} (\mathbf{y} - \mathbf{S}\mathbf{x})^\top (\mathbf{y} - \mathbf{S}\mathbf{x}) \\
&= \nabla_{\mathbf{x}} \left(\mathbf{y}^\top \mathbf{y} - \mathbf{y}^\top \mathbf{S}\mathbf{x} - \mathbf{x}^\top \mathbf{S}^\top \mathbf{y} + \mathbf{x}^\top \mathbf{S}^\top \mathbf{S}\mathbf{x} \right) \\
&= \nabla_{\mathbf{x}} \left(\mathbf{y}^\top \mathbf{y} - 2\mathbf{y}^\top \mathbf{S}\mathbf{x} + \mathbf{x}^\top \mathbf{S}^\top \mathbf{S}\mathbf{x} \right) \\
&= -2\mathbf{S}^\top \mathbf{y} + 2\mathbf{S}^\top \mathbf{S}\mathbf{x}
\end{aligned}$$

$\implies \mathbf{S}^\top \mathbf{S}\mathbf{x} = \mathbf{S}^\top \mathbf{y}$ sistema di equazioni normali

le equazioni normali ammettono soluzioni se e solo se le cononne di $\mathbf{S}$ sono linearmente indipendenti (se e solo se $rank(\mathbf{S}) = n \leq \ell$)

ipotesi: $\mathbf{S}^\top\mathbf{S}$ invertibile

stimatore ai minimi quadrati lineare

$$\hat{\mathbf{x}} = \left(\mathbf{S}^\top\mathbf{S}\right)^{-1}\mathbf{S}^\top\mathbf{y}$$

caso scalare ($x \in \mathbb{R}$)

modello dei dati

$y_i = s_i x + v_i,\ i \in \{1 \dots \ell\}$ con $s_i = 1$ (senza perdita di generalità)

$$\begin{bmatrix} y_1 \\ \vdots \\ y_\ell \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} x + \begin{bmatrix} v_1 \\ \vdots \\ v_\ell \end{bmatrix} \quad \Leftrightarrow \quad \mathbf{y} = \mathbf{S}x + \mathbf{v}, \quad \mathbf{S} = \mathbf{1}_\ell$$

$$\Longrightarrow \hat{x} = \left(\mathbf{S}^\top \mathbf{S}\right)^{-1} \mathbf{S}^\top \mathbf{y}$$

$$= \left(\mathbf{1}_\ell^\top \mathbf{1}_\ell\right)^{-1} \mathbf{1}_\ell^\top y = \frac{1}{\ell} \mathbf{1}_\ell^\top y = \frac{1}{\ell} \sum_{i=1}^{\ell} y_i \qquad \text{media delle misure}$$

$$J(x) = \|\mathbf{y} - \mathbf{1}_\ell x\|^2 = \sum_{i=1}^{\ell} (y_i - x)^2$$

$$\begin{aligned}\frac{d}{dx} J(x) &= \frac{d}{dx}\left(\sum_{i=1}^{\ell}(y_i - x)^2\right) = \sum_{i=1}^{\ell} \frac{d}{dx}\left(y_i^2 - 2y_i x + x^2\right) \\ &= \sum_{i=1}^{\ell}(-2y_i + 2x) = -2\sum_{i=1}^{\ell} y_i + 2\ell x\end{aligned}$$

$$\implies \hat{x} = \frac{1}{\ell}\sum_{i=1}^{\ell} y_i \quad \text{media delle misure}$$

stima LS dello stato iniziale

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t) \\ y(t) &= \mathbf{h}^\top \mathbf{x}(t) + v(t) \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{aligned} \qquad \begin{aligned} &\mathbf{x} \in \mathbb{R}^n,\ \mathbf{u} \in \mathbb{R}^m \\ &y \in \mathbb{R} \end{aligned}$$

sistema
osservabile

- $\ell \geq n$ **dati di uscita** (misure)

$$\hat{\mathbf{x}}_0 = \arg\min_{\mathbf{x} \in \mathcal{X}} \|\mathbf{z}^\ell - \mathcal{O}_\ell \mathbf{x}\|^2 \qquad \mathbf{z}^\ell = \mathbf{y}^\ell - \mathbf{Z}_\ell \mathbf{u}^\ell$$

ipotesi: sistema osservabile $\rightarrow$ $\mathcal{O}_\ell^\top \mathcal{O}_\ell$ invertibile

$$\boxed{\hat{\mathbf{x}}_0 = \left(\mathcal{O}_\ell^\top \mathcal{O}_\ell\right)^{-1} \mathcal{O}_\ell^\top \mathbf{z}^\ell}$$

Stima ai Minimi Quadrati

formulazione del problema

- stima ai minimi quadrati lineari
  - soluzione deterministica
  - soluzione ricorsiva
  - soluzione pesata
- introduzione del forgetting factor (FF)
- interpretazione geometrica e probabilistica

caso scalare ($x \in \mathbb{R}$)

modello dei dati
$y_i = x_r + v_i, \quad i \in \{1 \dots \ell\}$

$$\Longrightarrow \qquad \hat{x} = \frac{1}{\ell}\sum_{i=1}^{\ell} y_i$$

- **dato $y_{\ell+1}$ come aggiornare la stima $\hat{x}$?**
  scrittura ricorsiva

sia $\hat{x}(k)$ la stima ottenuta a partire dalle misure $\{y_i\}_{i=1}^{k}$

$$\boxed{\begin{aligned}\hat{x}(\ell+1) &= \frac{1}{\ell+1}\sum_{i=1}^{\ell+1} y_i = \frac{1}{\ell+1} y_{\ell+1} + \frac{1}{\ell+1}\sum_{i=1}^{\ell} y_i \\ &= \frac{1}{\ell+1} y_{\ell+1} + \frac{\ell}{\ell+1}\frac{1}{\ell}\sum_{i=1}^{\ell} y_i = \frac{1}{\ell+1} y_{\ell+1} + \frac{\ell}{\ell+1}\hat{x}(\ell)\end{aligned}}$$

osservazioni

$$\hat{x}(\ell) = \frac{1}{\ell+1} y_{\ell+1} + \frac{\ell}{\ell+1}\hat{x}(\ell) = \frac{1}{\ell+1} y_{\ell+1} + \left(1 - \frac{1}{\ell+1}\right)\hat{x}(\ell)$$

1. $\hat{x}(\ell+1) = \frac{1}{\ell+1} y_{\ell+1} + \frac{\ell}{\ell+1}\hat{x}(\ell)$

$= (1-\lambda_{\ell+1})\hat{x}_{y_{\ell+1}} + \lambda_{\ell+1}\hat{x}(\ell)$

$\hat{x}_{y_{\ell+1}} = y_{\ell+1} = \arg\min_{x\in\mathcal{X}} (y_{\ell+1} - x)^2$
$\lambda_{\ell+1} = \frac{\ell}{\ell+1} \in (0,1)$

2. $\hat{x}(\ell+1) = \frac{1}{\ell+1} y_{\ell+1} + \left(1 - \frac{1}{\ell+1}\right)\hat{x}(\ell)$

$= \hat{x}(\ell) + \frac{1}{\ell+1}\left(y_{\ell+1} - \hat{x}(\ell)\right)$

$(y_{\ell+1} - \hat{x}(\ell))$: errore di stima sul nuovo dato
$\frac{1}{\ell+1}$: guadagno di correzione

caso multivariato ($\mathbf{x} \in \mathbb{R}^n$)

modello dei dati
$\mathbf{y} = \mathbf{S}\mathbf{x} + \mathbf{v}, \quad \mathbf{S} \in \mathbb{R}^{\ell \times n}$ $\quad \Longrightarrow \quad$ $\hat{\mathbf{x}} = \left(\mathbf{S}^\top \mathbf{S}\right)^{-1} \mathbf{S}^\top \mathbf{y}$

- **dato $y_{\ell+1}$ come aggiornare la stima $\hat{\mathbf{x}}$?**

scrittura ricorsiva

siano

$$\mathbf{y}^{\ell+1} = \begin{bmatrix} y_1 \\ \vdots \\ y_\ell \\ y_{\ell+1} \end{bmatrix} = \begin{bmatrix} \mathbf{y}^\ell \\ y_{\ell+1} \end{bmatrix} \qquad \mathbf{S}(\ell+1) = \begin{bmatrix} s_{1,1} & \cdots & s_{1,n} \\ \vdots & & \vdots \\ s_{\ell,1} & \cdots & s_{\ell,n} \\ s_{\ell+1,1} & \cdots & s_{\ell+1,n} \end{bmatrix} = \begin{bmatrix} \mathbf{S}(\ell) \\ \mathbf{s}_{\ell+1}^\top \end{bmatrix}$$

$$\mathbf{y}^{\ell+1} \in \mathbb{R}^{\ell+1}, \; \mathbf{y}^\ell \in \mathbb{R}^\ell, \; y_{\ell+1} \in \mathbb{R} \qquad \mathbf{S}(\ell+1) \in \mathbb{R}^{(\ell+1)\times n}, \; \mathbf{S}(\ell) \in \mathbb{R}^{\ell \times n}, \; \mathbf{s}_{\ell+1} \in \mathbb{R}^n$$

sia $\hat{\mathbf{x}}(k)$ la stima ottenuta dalle misure $\{y_i\}_{i=1}^k$ e $\mathbf{S}(k)$ la relativa matrice di regressione

$$\boxed{\hat{\mathbf{x}}(\ell+1) = \left(\mathbf{S}(\ell+1)^\top \mathbf{S}(\ell+1)\right)^{-1} \mathbf{S}(\ell+1)^\top \mathbf{y}_{\ell+1}}$$

- $\mathbf{S}(\ell+1)^\top \mathbf{S}(\ell+1) = \begin{bmatrix} \mathbf{S}(\ell)^\top & \mathbf{s}_{\ell+1} \end{bmatrix} \begin{bmatrix} \mathbf{S}(\ell) \\ \mathbf{s}_{\ell+1}^\top \end{bmatrix} = \mathbf{S}(\ell)^\top \mathbf{S}(\ell) + s_{\ell+1} \mathbf{s}_{\ell+1}^\top$
- $\mathbf{S}(\ell+1)^\top \mathbf{y}_{\ell+1} = \begin{bmatrix} \mathbf{S}(\ell)^\top & \mathbf{s}_{\ell+1} \end{bmatrix} \begin{bmatrix} \mathbf{y}^\ell \\ y_{\ell+1} \end{bmatrix} = \mathbf{S}(\ell)^\top \mathbf{y}_\ell + \mathbf{s}_{\ell+1} y_{\ell+1}$

$$\boxed{\hat{\mathbf{x}}(\ell+1) = \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell) + \mathbf{s}_{\ell+1} \mathbf{s}_{\ell+1}^\top\right)^{-1} \left(\mathbf{S}(\ell)^\top \mathbf{y}^\ell + \mathbf{s}_{\ell+1} y_{\ell+1}\right)}$$

$$\boxed{\hat{\mathbf{x}}(\ell+1) = \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell) + \mathbf{s}_{\ell+1}\mathbf{s}_{\ell+1}^\top\right)^{-1} \left(\mathbf{S}(\ell)^\top \mathbf{y}^\ell + \mathbf{s}_{\ell+1} y_{\ell+1}\right)}$$

lemma d'inversione di matrice: $(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{C}^{-1}+\mathbf{DA}^{-1}\mathbf{B}\right)^{-1}\mathbf{DA}^{-1}$

$$\left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell) + \mathbf{s}_{\ell+1}\mathbf{s}_{\ell+1}^\top\right)^{-1} = \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1} +$$

$$\underbrace{-\left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1}\mathbf{s}_{\ell+1}\left(1+\mathbf{s}_{\ell+1}^\top\left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1}\mathbf{s}_{\ell+1}\right)^{-1}}_{\mathbf{k}(\ell+1)\in\mathbb{R}^n}\mathbf{s}_{\ell+1}^\top\left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1}$$

$$= \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1} - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1}$$

$$\boxed{\begin{aligned}\hat{\mathbf{x}}(\ell+1) = {} & \left(\mathbf{S}(\ell)^\top\mathbf{S}(\ell)\right)^{-1}\left(\mathbf{S}(\ell)^\top\mathbf{y}^\ell + \mathbf{s}_{\ell+1}y_{\ell+1}\right) + \\ & - \left(\mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\left(\mathbf{S}(\ell)^\top\mathbf{S}(\ell)\right)^{-1}\right)\left(\mathbf{S}(\ell)^\top\mathbf{y}^\ell + \mathbf{s}_{\ell+1}y_{\ell+1}\right)\end{aligned}}$$

- $\left(\mathbf{S}(\ell)^\top\mathbf{S}(\ell)\right)^{-1}\left(\mathbf{S}(\ell)^\top\mathbf{y}^\ell + \mathbf{s}_{\ell+1}y_{\ell+1}\right) = \hat{\mathbf{x}}(\ell) + \left(\mathbf{S}(\ell)^\top\mathbf{S}(\ell)\right)^{-1}\mathbf{s}_{\ell+1}y_{\ell+1}$

- $$\begin{aligned} & -\left(\mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\left(\mathbf{S}(\ell)^\top\mathbf{S}(\ell)\right)^{-1}\right)\left(\mathbf{S}(\ell)^\top\mathbf{y}^\ell + \mathbf{s}_{\ell+1}y_{\ell+1}\right) = \\ & = -\mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\hat{\mathbf{x}}(\ell) - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\left(\mathbf{S}(\ell)^\top\mathbf{S}(\ell)\right)^{-1}\mathbf{s}_{\ell+1}y_{\ell+1}\end{aligned}$$

$$\boxed{\begin{aligned}\hat{\mathbf{x}}(\ell+1) = {} & \hat{\mathbf{x}}(\ell) + \left(\mathbf{S}(\ell)^\top\mathbf{S}(\ell)\right)^{-1}\mathbf{s}_{\ell+1}y_{\ell+1} + \\ & -\mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\hat{\mathbf{x}}(\ell) - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\left(\mathbf{S}(\ell)^\top\mathbf{S}(\ell)\right)^{-1}\mathbf{s}_{\ell+1}y_{\ell+1}\end{aligned}}$$

$$\boxed{\begin{aligned}\hat{\mathbf{x}}(\ell+1) = \hat{\mathbf{x}}(\ell) + \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1} \mathbf{s}_{\ell+1} y_{\ell+1} + \\ -\mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top \hat{\mathbf{x}}(\ell) - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1} \mathbf{s}_{\ell+1} y_{\ell+1}\end{aligned}}$$

- $\left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1} \mathbf{s}_{\ell+1} y_{\ell+1} - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1} \mathbf{s}_{\ell+1} y_{\ell+1} = \ldots = \mathbf{k}(\ell+1) y_{\ell+1}$

$$\boxed{\hat{\mathbf{x}}(\ell+1) = \hat{\mathbf{x}}(\ell) - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top \hat{\mathbf{x}}(\ell) + \mathbf{k}(\ell+1) y_{\ell+1}}$$

$$\boxed{\hat{\mathbf{x}}(\ell+1) = \hat{\mathbf{x}}(\ell) - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^{\top}\hat{\mathbf{x}}(\ell) + \mathbf{k}(\ell+1)y_{\ell+1}}$$

$$\hat{\mathbf{x}}(\ell+1) = \hat{\mathbf{x}}(\ell) + \mathbf{k}(\ell+1)\left(y_{\ell+1} - \mathbf{s}_{\ell+1}^{\top}\hat{\mathbf{x}}(\ell)\right)$$

- **dato $y_{\ell+1}$ come calcolare $\mathbf{k}(\ell+1)$?**

aggiornamento del guadagno

$$\begin{aligned}\mathbf{k}(\ell+1) &= \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1} \mathbf{s}_{\ell+1} \left(1 + \mathbf{s}_{\ell+1}^\top \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1} \mathbf{s}_{\ell+1}\right)^{-1} \\ &= \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1} \left(1 + \mathbf{s}_{\ell+1}^\top \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\right)^{-1}\end{aligned}$$

$$\bar{\mathbf{S}}(\ell) = \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1}$$

$$\begin{aligned}\bar{\mathbf{S}}(\ell+1) &= \left(\mathbf{S}(\ell+1)^\top \mathbf{S}(\ell+1)\right)^{-1} = \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell) + \mathbf{s}_{\ell+1}\mathbf{s}_{\ell+1}^\top\right)^{-1} \\ &= \left(\bar{\mathbf{S}}(\ell)^{-1} + \mathbf{s}_{\ell+1}\mathbf{s}_{\ell+1}^\top\right)^{-1} = \bar{\mathbf{S}}(\ell) - \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\left(1 + \mathbf{s}_{\ell+1}^\top \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\right)^{-1} s_{\ell+1}^\top \bar{\mathbf{S}}(\ell) \\ &= \bar{\mathbf{S}}(\ell) - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top \bar{\mathbf{S}}(\ell)\end{aligned}$$

algoritmo di stima ricorsiva

$$\hat{\mathbf{x}}(\ell+1) = \hat{\mathbf{x}}(\ell) + \mathbf{k}(\ell+1)\left(y_{\ell+1} - \mathbf{s}_{\ell+1}^\top \hat{\mathbf{x}}(\ell)\right)$$
$$\mathbf{k}(\ell+1) = \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\left(1 + \mathbf{s}_{\ell+1}^\top \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\right)^{-1}$$
$$\bar{\mathbf{S}}(\ell+1) = \left(\mathbf{I}_n - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\right)\bar{\mathbf{S}}(\ell)$$

condizioni iniziali: $\hat{\mathbf{x}}(0)$ arbitrario, $\bar{\mathbf{S}}(0)$ invertibile

⚠ $\bar{\mathbf{S}}(\ell)$ invertibile per $\ell \geq n$ $\quad \Longrightarrow \bar{\mathbf{S}}(\ell) = \left(\bar{\mathbf{S}}_0^{-1} + \left(\mathbf{S}(\ell)^\top \mathbf{S}(\ell)\right)^{-1}\right), \ \bar{\mathbf{S}}_0 = \frac{1}{\epsilon}\mathbf{I}_n$

formulazione del problema

- stima ai minimi quadrati lineari
  - soluzione deterministica
  - soluzione ricorsiva
  - soluzione pesata
- introduzione del forgetting factor (FF)
- interpretazione geometrica e probabilistica

introduzione dei pesi

- **come sfruttare la conoscenza sulla bontà delle misure?**
  dare più importanza ai dati che sono noti essere più precisi
  con l'introduzione di coefficienti di peso

stima ai minimi quadrati lin. standard (o LLS - lin. least squares)	**stima ai minimi quadrati lin. pesata** (o WLLS - weighted lin. least squares)
$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}} \Vert\mathbf{y} - \mathbf{S}\mathbf{x}\Vert^2$	$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}} \Vert\mathbf{y} - \mathbf{S}\mathbf{x}\Vert_{\mathbf{W}}^2$
$\Vert\mathbf{y} - \mathbf{S}\mathbf{x}\Vert^2 = (\mathbf{y} - \mathbf{S}\mathbf{x})^\top (\mathbf{y} - \mathbf{S}\mathbf{x})$	$\Vert\mathbf{y} - \mathbf{S}\mathbf{x}\Vert_{\mathbf{W}}^2 = (\mathbf{y} - \mathbf{S}\mathbf{x})^\top \mathbf{W} (\mathbf{y} - \mathbf{S}\mathbf{x})$

$\mathbf{W} \in \mathbb{R}^{\ell \times \ell}$ matrice dei pesi

- in genere: $\mathbf{W} = \mathit{diag}\left(\begin{bmatrix} w_1 \dots w_\ell \end{bmatrix}\right)$, $w_i \geq 0$
- $\mathbf{W}$ semi definita positiva
- scelta naturale $w_i = \frac{1}{\sigma_{v_i}^2}$ con $\sigma_{v_i}^2 \geq 0$ nota

  $\longrightarrow \quad \boxed{\mathbf{W} = \mathbf{\Sigma}_v^{-1}}$

soluzione deterministica pesata

caso multidimensionale

$$J(\mathbf{x}) = \|\mathbf{y} - \mathbf{S}\mathbf{x}\|_{\mathbf{W}}^2$$

$$\begin{aligned}
\nabla_{\mathbf{x}} J(\mathbf{x}) &= \nabla_{\mathbf{x}} (\mathbf{y} - \mathbf{S}\mathbf{x})^\top \mathbf{W} (\mathbf{y} - \mathbf{S}\mathbf{x}) \\
&= \nabla_{\mathbf{x}} \left(\mathbf{y}^\top \mathbf{W}\mathbf{y} - \mathbf{y}^\top \mathbf{W}\mathbf{S}\mathbf{x} - \mathbf{x}^\top \mathbf{S}^\top \mathbf{W}\mathbf{y} + \mathbf{x}^\top \mathbf{S}^\top \mathbf{W}\mathbf{x}\mathbf{S} \right) \\
&= \nabla_{\mathbf{x}} \left(\mathbf{y}^\top \mathbf{W}\mathbf{y} - 2\mathbf{x}^\top \mathbf{S}^\top \mathbf{W}\mathbf{y} + \mathbf{x}^\top \mathbf{S}^\top \mathbf{S}\mathbf{x} \right) \\
&= -2\mathbf{S}^\top \mathbf{W}\mathbf{y} + 2\mathbf{S}^\top \mathbf{W}\mathbf{S}\mathbf{x}
\end{aligned}$$

$\implies \mathbf{S}^\top \mathbf{W}\mathbf{S}\mathbf{x} = \mathbf{S}^\top \mathbf{W}\mathbf{y}$ sistema di equazioni normali *ipotesi*: $\mathbf{S}^\top \mathbf{W} \mathbf{S}$ invertibile

stimatore ai minimi quadrati lineare pesato

$$\hat{\mathbf{x}} = \left(\mathbf{S}^\top \mathbf{W} \mathbf{S} \right)^{-1} \mathbf{S}^\top \mathbf{W} \mathbf{y}$$

soluzione ricorsiva pesata

$$
\begin{aligned}
\hat{\mathbf{x}}(\ell+1) &= \hat{\mathbf{x}}(\ell) + \mathbf{k}(\ell+1)\left(y_{\ell+1} - \mathbf{s}_{\ell+1}^\top \hat{\mathbf{x}}(\ell)\right)\\
\mathbf{k}(\ell+1) &= \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\left(w_{\ell+1} + \mathbf{s}_{\ell+1}^\top \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\right)^{-1}\\
\bar{\mathbf{S}}(\ell+1) &= \left(\mathbf{I}_n - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\right)\bar{\mathbf{S}}(\ell)
\end{aligned}
$$

$$
\bar{\mathbf{S}}(\ell) = \left(\mathbf{S}(\ell)^\top \mathbf{W}\mathbf{S}(\ell)\right)^{-1}
$$

$$
\mathbf{k}(\ell+1) = \left(\mathbf{S}(\ell)^\top \mathbf{W}\mathbf{S}(\ell)\right)^{-1}\mathbf{s}_{\ell+1}\left(w_{\ell+1} + \mathbf{s}_{\ell+1}^\top \left(\mathbf{S}(\ell)^\top \mathbf{W}\mathbf{S}(\ell)\right)^{-1}\mathbf{s}_{\ell+1}\right)^{-1}
$$

Stima ai Minimi Quadrati

formulazione del problema

- stima ai minimi quadrati lineari
  - soluzione deterministica
  - soluzione ricorsiva
  - soluzione pesata
- introduzione del forgetting factor (FF)
- interpretazione geometrica e probabilistica

introduzione del forgetting factor (FF)

$$\begin{aligned}
\hat{\mathbf{x}}(\ell+1) &= \hat{\mathbf{x}}(\ell) + \mathbf{k}(\ell+1)\left(y_{\ell+1} - \mathbf{s}_{\ell+1}^\top \hat{\mathbf{x}}(\ell)\right) \\
\mathbf{k}(\ell+1) &= \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\left(w_{\ell+1} + \mathbf{s}_{\ell+1}^\top \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\right)^{-1} \\
\bar{\mathbf{S}}(\ell+1) &= \left(\mathbf{I}_n - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\right)\bar{\mathbf{S}}(\ell)
\end{aligned}$$

scegliere i pesi in modo da rendere l'algoritmo più reattivo
$\Longrightarrow$ pesare meno la misure più vecchie

modello dei dati

$$y_i = \mathbf{s}_i^\top \mathbf{x} + v_i, \quad \mathbf{s}_i = \begin{bmatrix} s_{i,1} & \dots & s_{i,n} \end{bmatrix}^\top \in \mathbb{R}^n, \; i \in \{1 \dots \ell\}$$

stima ai minimi quadrati con **forgetting factor** (fattore di oblio)

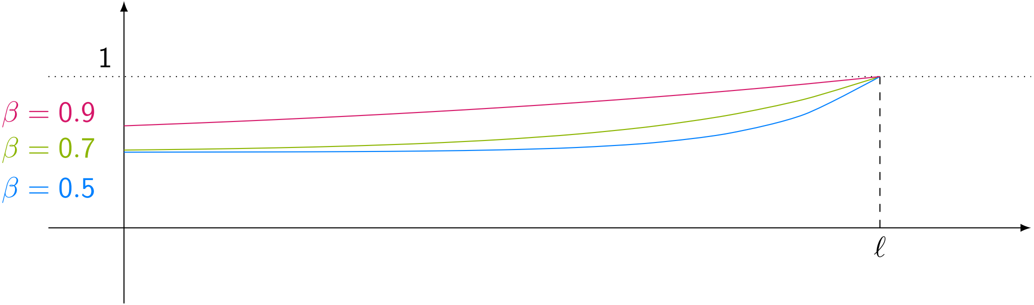
$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}} \sum_{i=1}^{\ell} \beta^{\ell-i}(y_i - \mathbf{s}_i^\top \mathbf{x})^2 = \arg\min_{\mathbf{x}\in\mathcal{X}} \sum_{\mathbf{i}=\mathbf{1}}^{\ell} \beta^{\ell-\mathbf{i}}\mathbf{v}_\mathbf{i}^2$$

$$\beta \in (0,1], \text{ tipicamente } \beta \in [0.95, 0.99]$$

- $\beta = 1$: stima ai minimi quadrati standard
- $\beta < 1$: la misura corrente ($i = \ell$) ha peso unitario ($\beta^0 = 1$)
  la misura k passi indietro ($i = \ell - k$) ha peso minore ($\beta^k < 1$)

esempio: andamento del FF

$\ell = 10$



soluzione ricorsiva con FF: caso scalare

modello dei dati

$$y_i = x + v_i, \; i \in \{1 \dots \ell\}$$

$$J(x) = \sum_{i=1}^{\ell} \beta^{\ell-i}(y_i - x)^2$$

$$\begin{aligned} \frac{d}{dx} J(x) &= \sum_{i=1}^{\ell} \beta^{\ell-i} \; (-2y_i + 2x) = 2\sum_{i=1}^{\ell} \beta^{\ell-i} x - 2\sum_{i=1}^{\ell} \beta^{\ell-i} y_i \\ &= 2\sum_{k=0}^{\ell-1} \beta^k x - 2\sum_{i=1}^{\ell} \beta^{\ell-i} y_i = 2\frac{1-\beta^\ell}{1-\beta} x - 2\sum_{i=1}^{\ell} \beta^{\ell-i} y_i \end{aligned}$$

$$\Longrightarrow \; \hat{x} = \frac{1-\beta}{1-\beta^\ell} \sum_{i=1}^{\ell} \beta^{\ell-i} y_i \quad \text{con} \quad \sum_{i=1}^{\ell} \frac{1-\beta}{1-\beta^\ell} \beta^{\ell-i} = 1$$

$$
\begin{aligned}
\hat{x}(\ell+1) &= \tfrac{1-\beta}{1-\beta^{\ell+1}} \textstyle\sum_{i=1}^{\ell+1} \beta^{\ell+1-i} y_i = \tfrac{1-\beta}{1-\beta^{\ell+1}} \left(y_{\ell+1} + \textstyle\sum_{i=1}^{\ell} \beta^{\ell+1-i} y_i \right) \\
&= \tfrac{1-\beta}{1-\beta^{\ell+1}} y_{\ell+1} + \tfrac{1-\beta}{1-\beta^{\ell+1}} \tfrac{1-\beta^{\ell}}{1-\beta^{\ell}} \beta \beta^{-1} \textstyle\sum_{i=1}^{\ell} \beta^{\ell+1-i} y_i \\
&= \tfrac{1-\beta}{1-\beta^{\ell+1}} y_{\ell+1} + \beta \tfrac{1-\beta^{\ell}}{1-\beta^{\ell+1}} \tfrac{1-\beta}{1-\beta^{\ell}} \textstyle\sum_{i=1}^{\ell} \beta^{\ell-i} y_i \\
&= \tfrac{1-\beta}{1-\beta^{\ell+1}} y_{\ell+1} + \beta \tfrac{1-\beta^{\ell}}{1-\beta^{\ell+1}} \hat{x}(\ell)
\end{aligned}
$$

combinazione convessa

$$\hat{x}(\ell+1) = \lambda \hat{x}(\ell) + (1-\lambda) y_{\ell+1}$$

$$\Longrightarrow \ \hat{x}(\ell+1) = \hat{x}(\ell) + \frac{1-\beta}{1-\beta^{\ell+1}} \left(y_{\ell+1} - \hat{x}(\ell) \right)$$

soluzione ricorsiva con FF: caso multivariato

$$\begin{aligned}
\hat{\mathbf{x}}(\ell+1) &= \hat{\mathbf{x}}(\ell) + \mathbf{k}(\ell+1)\left(y_{\ell+1} - \mathbf{s}_{\ell+1}^\top \hat{\mathbf{x}}(\ell)\right) \\
\mathbf{k}(\ell+1) &= \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\left(\beta + \mathbf{s}_{\ell+1}^\top \bar{\mathbf{S}}(\ell)\mathbf{s}_{\ell+1}\right)^{-1} \\
\bar{\mathbf{S}}(\ell+1) &= \frac{1}{\beta}\left(\mathbf{I}_n - \mathbf{k}(\ell+1)\mathbf{s}_{\ell+1}^\top\right)\bar{\mathbf{S}}(\ell)
\end{aligned}$$

osservazioni

- $\beta = 1$ (stima ai minimi quadrati standard)
  $\lim_{\ell\to\infty} \|\bar{\mathbf{S}}(\ell)\| = 0 \;\Rightarrow\; \lim_{\ell\to\infty} \mathbf{k}(\ell+1) = 0 \;\Rightarrow\; \lim_{\ell\to\infty} \hat{\mathbf{x}}(\ell+1) = \hat{\mathbf{x}}(\ell)$

  $\beta < 1$ (stima ai minimi quadrati con FF)
  $\bar{\mathbf{S}}(\ell+1) \propto \frac{1}{\beta} > 1 \;\Rightarrow\; \lim_{\ell\to\infty} \mathbf{k}(\ell+1) \neq 0 \;\Rightarrow\; \lim_{\ell\to\infty} \hat{\mathbf{x}}(\ell+1) \propto (\hat{\mathbf{x}}(\ell), y_{\ell+1})$

- $\beta \ll 1$: riduzione della precisione di stima nel caso di parametri costanti

  FF variabile

  $$\beta(\ell) = \begin{cases} 0.95 & |e(\ell)| \geq \bar{e} \\ 1 - \alpha\,(1 - \beta(n-1)) & |e(\ell)| < \bar{e} \end{cases}$$

  $e(\ell) = y_{\ell+1} - \mathbf{s}_{\ell+1}^\top \hat{\mathbf{x}}(\ell)$
  $\bar{e}$ valore di soglia
  α param. di progetto

Stima ai Minimi Quadrati

formulazione del problema

- stima ai minimi quadrati lineari
  - soluzione deterministica
  - soluzione ricorsiva
  - soluzione pesata
- introduzione del forgetting factor (FF)
- interpretazione geometrica e probabilistica

interpretazione geometrica

⚠ $\mathbf{s}_i = \begin{bmatrix} s_{i,1} & \dots & s_{i,n} \end{bmatrix}^\top \in \mathbb{R}^n,\ i \in \{1 \dots \ell\}$

$\varsigma_j = \begin{bmatrix} s_{1,j} & \dots & s_{\ell,j} \end{bmatrix}^\top \in \mathbb{R}^\ell,\ j \in \{1 \dots n\}$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_\ell \end{bmatrix} = \begin{bmatrix} s_{1,1} \\ \vdots \\ s_{\ell,1} \end{bmatrix} x_1 + \cdots + \begin{bmatrix} s_{1,n} \\ \vdots \\ s_{\ell,n} \end{bmatrix} x_n + \begin{bmatrix} v_1 \\ \vdots \\ v_\ell \end{bmatrix} \quad \Leftrightarrow \quad \mathbf{y} - \varsigma_1 x_1 - \cdots - \varsigma_n x_n = \mathbf{v}$$

approssimare $\mathbf{y} \in \mathbb{R}^\ell$ tramite il vettore $\mathbf{y}^* \in \mathbb{R}^\ell$, combinazione lineare degli n vettori $\varsigma_1 \dots \varsigma_n \in \mathbb{R}^\ell$, *minimizzando* la norma del vettore $\mathbf{v} \in \mathbb{R}^\ell$

$\Rightarrow$ $\mathbf{y}^*$ **proiezione ortogonale** di $\mathbf{y}$ su $\langle \varsigma_1 \dots \varsigma_n \rangle$

- $\mathbf{y}^*$ **proiezione ortogonale** di $\mathbf{y}$ su $\langle \varsigma_1 \dots \varsigma_n \rangle \quad \Longrightarrow \quad (\mathbf{y} - \mathbf{y}^*) \perp \langle \varsigma_1 \dots \varsigma_n \rangle$
- $\mathbf{y}^*$ **combinazione lineare** di $\varsigma_1 \dots \varsigma_n \quad \Longrightarrow \quad \mathbf{y}^* = \varsigma_1 \hat{\mathbf{x}}_1 + \dots + \varsigma_n \hat{\mathbf{x}}_n$

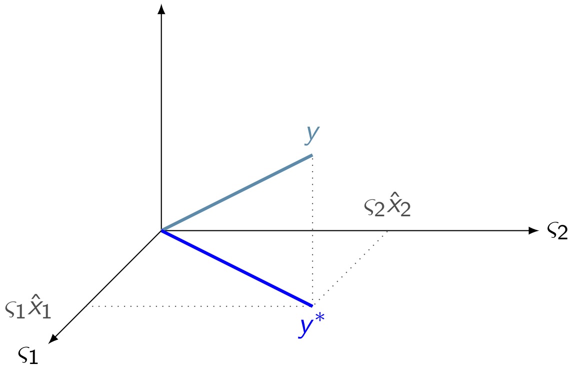
$$\begin{cases} (\mathbf{y} - \mathbf{y}^*)^\top \varsigma_1 = 0 \\ \qquad \vdots \\ (\mathbf{y} - \mathbf{y}^*)^\top \varsigma_n = 0 \end{cases} \iff \begin{cases} \mathbf{y}^\top \varsigma_1 = (\mathbf{y}^*)^\top \varsigma_1 = (\varsigma_1 \hat{\mathbf{x}}_1 + \dots + \varsigma_n \hat{\mathbf{x}}_n)^\top \varsigma_1 \\ \quad \vdots \\ \mathbf{y}^\top \varsigma_n = (\mathbf{y}^*)^\top \varsigma_n = (\varsigma_1 \hat{\mathbf{x}}_1 + \dots + \varsigma_n \hat{\mathbf{x}}_n)^\top \varsigma_n \end{cases}$$

$$\Updownarrow$$

$$\mathbf{S}^\top \mathbf{y} = \mathbf{S}^\top \mathbf{S} \hat{\mathbf{x}} \iff \mathbf{y}^\top \mathbf{S} = (\mathbf{S} \hat{\mathbf{x}})^\top \mathbf{S}$$

esempio

$\ell = 3, \;\; n = 2$



interpretazione probabilistica

- **come valutare la bontà della stima?**

  assunzioni probabilistiche

modello dei dati

$$y_i = \mathbf{s}_i^\top \mathbf{x}_r + v_i, \quad \mathbf{s}_i = \begin{bmatrix} s_{i,1} & \dots & s_{i,n} \end{bmatrix}^\top \in \mathbb{R}^n, \; i \in \{1 \dots \ell\}$$

$\mathbf{x}_r \in \mathbb{R}^n$ valore vero dei parametri

$\{v_i\}_{i=1}^{\ell}$ i.i.d. con media nulla e varianza σ_v^2

modello dei dati

$$\mathbf{y} = \mathbf{S}\mathbf{x}_r + \mathbf{v}$$

$$\implies \quad \hat{\mathbf{x}} = \left(\mathbf{S}^\top \mathbf{S}\right)^{-1} \mathbf{S}^\top \mathbf{y} = \left(\mathbf{S}^\top \mathbf{S}\right)^{-1} \mathbf{S}^\top (\mathbf{S}\mathbf{x}_r + \mathbf{v})$$

$$= \left(\mathbf{S}^\top \mathbf{S}\right)^{-1} \mathbf{S}^\top \mathbf{S}\, \mathbf{x}_r + \left(\mathbf{S}^\top \mathbf{S}\right)^{-1} \mathbf{S}^\top \mathbf{v} = \mathbf{x}_r + \left(\mathbf{S}^\top \mathbf{S}\right)^{-1} \mathbf{S}^\top \mathbf{v}$$

$$\boxed{\hat{\mathbf{x}} = \mathbf{x}_r + \left(\mathbf{S}^\top \mathbf{S}\right)^{-1} \mathbf{S}^\top \mathbf{v}}$$

proprietà stocastiche della stima

$$\hat{\mathbf{x}} = \mathbf{x}_r + \left(\mathbf{S}^\top\mathbf{S}\right)^{-1}\mathbf{S}^\top\mathbf{v} \qquad \mathbf{\Sigma}_v = \mathbb{E}\left[\mathbf{v}\mathbf{v}^\top\right] = \sigma_v^2\mathbf{I}_\ell$$

- $\mathbf{M}_{\hat{\mathbf{x}}} = \mathbb{E}\left[\hat{\mathbf{x}}\right] = \mathbb{E}\left[\mathbf{x}_r + \left(\mathbf{S}^\top\mathbf{S}\right)^{-1}\mathbf{S}^\top\mathbf{v}\right] = \mathbf{x}_r + \left(\mathbf{S}^\top\mathbf{S}\right)^{-1}\mathbf{S}^\top\mathbb{E}\left[\mathbf{v}\right] = \mathbf{x}_r$

  stima corretta/unbiased (mean squared error $MSE\left(\hat{\mathbf{x}}\right) = \mathbb{E}\left[\|\hat{\mathbf{x}} - \mathbf{x}_r\|^2\right] = \mathbf{0}$)

- $\mathbf{\Sigma}_{\hat{\mathbf{x}}} = \mathbb{E}\left[(\hat{\mathbf{x}} - \mathbb{E}\left[\hat{\mathbf{x}}\right])^2\right] = \mathbb{E}\left[\left(\mathbf{x}_r + \left(\mathbf{S}^\top\mathbf{S}\right)^{-1}\mathbf{S}^\top\mathbf{v} - \mathbf{x}_r\right)^2\right] = \mathbb{E}\left[\left(\left(\mathbf{S}^\top\mathbf{S}\right)^{-1}\mathbf{S}^\top\mathbf{v}\right)^2\right]$

  $= \left(\mathbf{S}^\top\mathbf{S}\right)^{-1}\mathbf{S}^\top\mathbb{E}\left[\mathbf{v}\mathbf{v}^\top\right]\mathbf{S}\left(\mathbf{S}^\top\mathbf{S}\right)^{-1} = \left(\mathbf{S}^\top\mathbf{S}\right)^{-1}\mathbf{S}^\top\left(\sigma_v^2\mathbf{I}\right)\mathbf{S}\left(\mathbf{S}^\top\mathbf{S}\right)^{-1}$

  $= \sigma_v^2\left(\mathbf{S}^\top\mathbf{S}\right)^{-1}$

  stima consistente ($\lim_{n\to\infty}\mathbf{\Sigma}_{\hat{x}} = \mathbf{0}$)

conseguenze dell'interpretazione probabilistica

1. *livello di confidenza nella stima*

modello dei dati - caso scalare

$$y_i = x_r + v_i, \quad i \in \{1 \dots \ell\} \qquad \Longrightarrow \qquad \hat{x} = \frac{1}{\ell}\sum_{i=1}^{\ell} y_i$$

$$\{v_i \sim \mathcal{N}(0, \sigma_v^2)\}_{i=1}^{\ell} \quad \text{i.i.d.}$$

- $y_i \sim \mathcal{N}\left(x_r, \sigma_v^2\right)$
- $\hat{x} \sim \mathcal{N}\left(m_{\hat{x}}, \sigma_{\hat{x}}^2\right)$

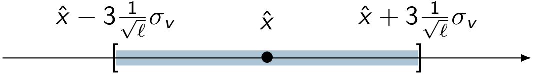
$$m_{\hat{x}} = \mathbb{E}\left[\hat{x}\right] = \mathbb{E}\left[\frac{1}{\ell}\sum_{i=1}^{\ell} y_i\right] = \frac{1}{\ell}\sum_{i=1}^{\ell}\mathbb{E}\left[y_i\right] = \frac{1}{\ell}\ell x_r = x_r$$

$$\sigma_{\hat{x}}^2 = var(\hat{x}) = var\left(\frac{1}{\ell}\sum_{i=1}^{\ell} y_i\right) = \frac{1}{\ell^2}\sum_{i=1}^{\ell} var\left(y_i\right) = \frac{1}{\ell^2}\ell\sigma_v^2 = \frac{1}{\ell}\sigma_v^2$$

$$\hat{x} \sim \mathcal{N}\left(m_{\hat{x}}, \sigma_{\hat{x}}^2\right), \qquad m_{\hat{x}} = x_r, \quad \sigma_{\hat{x}}^2 = \frac{1}{\ell}\sigma_v^2$$

intervallo di confidenza: $\quad \mathcal{P}\left(|x_r - \hat{x}| \leq 3\sigma_{\hat{x}}\right) \geq 99\%$

$$\mathcal{P}\left(|x_r - \hat{x}| \leq 3\frac{1}{\sqrt{\ell}}\sigma_v\right) \geq 99\%$$



x_r (non noto) è nell'intervallo di confidenza con alta probabilità

$\propto \sigma_v$

$$\sigma_v^2 = \mathbb{E}\left[v_i^2\right] = \mathbb{E}\left[(y_i - x_r)^2\right] \approx \mathbb{E}\left[(y_i - \hat{x})^2\right]$$

$$\sigma_v^2 \approx \hat{\sigma}_v^2 = \frac{1}{\ell}\sum_{i=1}^{\ell}\left(y_i - \hat{x}\right)^2$$

2. *equivalenza con la stima a massima verosimiglianza*

stima a massima verosimiglianza (o ML - maximum likelihood)

trovare $\hat{\mathbf{x}}$ che massimizza la probabilità che le misure che si otterrebbero in corrispondenza a $\hat{\mathbf{x}}$ siano uguali a quelle effettivamente ottenute

$\rightarrow$ massimizzare $\mathcal{L}(\mathbf{x}|\mathbf{y}) = f(\mathbf{y}|\mathbf{x})$: funzione di massima verosimiglianza

stima a massima verosimiglianza (o ML - maximum likelihood)

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\mathcal{X}} \mathcal{L}(\mathbf{x}|\mathbf{y})$$

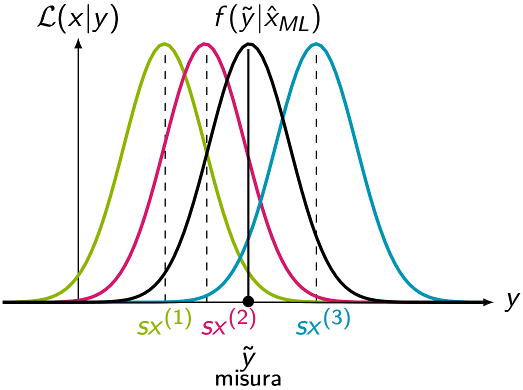
se $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_\ell)$ allora $\mathbf{y} \sim \mathcal{N}(\mathbf{S}\mathbf{x}, \sigma_v^2 \mathbf{I}_\ell)$

se $\mathbf{y} \sim \mathcal{N}(\mathbf{S}\mathbf{x}, \sigma_v^2 \mathbf{I}_\ell)$ allora $\mathcal{L}(\mathbf{x}|\mathbf{y}) = \dfrac{1}{\sqrt{2\pi\sigma_v^2}} exp\left(-\dfrac{1}{2\sigma_v^2}(\mathbf{y} - \mathbf{S}\mathbf{x})^\top (\mathbf{y} - \mathbf{S}\mathbf{x})\right)$

stima ML: $$\hat{\mathbf{x}}_{ML} = \arg\max_{\mathbf{x} \in \mathcal{X}} \left(\frac{1}{\sqrt{2\pi\sigma_v^2}} exp\left(-\frac{1}{2\sigma_v^2}(\mathbf{y} - \mathbf{S}\mathbf{x})^\top (\mathbf{y} - \mathbf{S}\mathbf{x})\right)\right)$$

stima ML: $\hat{\mathbf{x}}_{ML} = \arg\max_{\mathbf{x}\in\mathcal{X}} \left(\frac{1}{\sqrt{2\pi\sigma_v^2}} exp\left(-\frac{1}{2\sigma_v^2} (\mathbf{y} - \mathbf{S}\mathbf{x})^\top (\mathbf{y} - \mathbf{S}\mathbf{x}) \right) \right)$

esempio - caso scalare SISO: $x \in \mathbb{R}$, $y = sx + v \in \mathbb{R}$



se $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \sigma_v^2 \mathbf{I}_\ell)$ allora $\mathbf{y} \sim \mathcal{N}(\mathbf{Sx}, \sigma_v^2 \mathbf{I}_\ell)$

se $\mathbf{y} \sim \mathcal{N}(\mathbf{Sx}, \sigma_v^2 \mathbf{I}_\ell)$ allora $\mathcal{L}(\mathbf{x}|\mathbf{y}) = \dfrac{1}{\sqrt{2\pi\sigma_v^2}} exp\left(-\dfrac{1}{2\sigma_v^2}(\mathbf{y} - \mathbf{Sx})^\top (\mathbf{y} - \mathbf{Sx})\right)$

stima ML: $$\hat{\mathbf{x}}_{ML} = \arg\max_{\mathbf{x}\in\mathcal{X}} \left(\frac{1}{\sqrt{2\pi\sigma_v^2}} exp\left(-\frac{1}{2\sigma_v^2}(\mathbf{y} - \mathbf{Sx})^\top (\mathbf{y} - \mathbf{Sx})\right)\right)$$

stima (WL)LS: $$\hat{\mathbf{x}}_{LS} = \arg\min_{\mathbf{x}\in\mathcal{X}} \left(\frac{1}{\sigma_v^2}(\mathbf{y} - \mathbf{Sx})^\top (\mathbf{y} - \mathbf{Sx})\right)$$

$$\boxed{\hat{\mathbf{x}}_{ML} = \hat{\mathbf{x}}_{LS}}$$

Stima Bayesiana

- approccio fisheriano VS. approccio bayesiano
- stima MAP - Maximum A Posteriori
- stima MMSE - Minimum Mean Square Error

Fisher VS. Bayes

la stima ML è una stima fisheriana

approccio fisheriano

la variabile da stimare è considerata una *grandezza deterministica costante*
$\rightarrow$ esiste il *valore vero* che può essere stimato sfruttando misure/dati sperimentali

approccio bayesiano

la variabile da stimare è considerata una *grandezza aleatoria*
$\rightarrow$ se ne può stimare una realizzazione utilizzando

- informazioni a posteriori: misure/dati sperimentali
- informazioni a priori (indipendenti dai dati)

approccio fisheriano

la variabile da stimare è considerata una *grandezza deterministica costante*
$\rightarrow$ esiste il *valore vero* che può essere stimato sfruttando misure/dati sperimentali

- $\triangle$ algoritmi di ottimizzazione a bassa complessità computazionale
- $\bigtriangledown$ minima varianza d'errore solo sotto ipotesi di
  linearità del modello e gaussianità dell' errore di misura
- $\bigtriangledown$ intervalli di confidenza non realistici
  (causa approssimazione Gaussiana della distribuzione delle stime)

approccio bayesiano

la variabile da stimare è considerata una *grandezza aleatoria*
$\rightarrow$ se ne può stimare una realizzazione utilizzando

- informazioni a posteriori: misure/dati sperimentali
- informazioni a priori (indipendenti dai dati)

informazioni a priori

derivante dalla letteratura, da esperimenti precedenti, da studi passati

informazioni a priori

$f(\mathbf{x})$ densità di probabilità *nota a priori*
(aspettative in termini di distribuzione probabilistica)

$f(\mathbf{x}|\mathbf{y})$ densità di probabilità *da stimare a posteriori*
utilizzando/in funzione delle misure $\mathbf{y}$
(modifica della aspettative iniziali)

stime puntuali & stime degli intervalli di confidenza

teorema di Bayes

$$f(\mathbf{x}|\mathbf{y}) = \frac{f(\mathbf{y}|\mathbf{x})f(\mathbf{x})}{f(\mathbf{y})}$$

Thomas Bayes (1702 – 1761)
matematico e ministro presbiteriano

la densità di probabilità di **x** *a posteriori* di una misura **y** è pari al prodotto della densità di probabilità *a priori* di **x** e della probabilità condizionata della misura **y** dato **x**, diviso per la densità di probabilità della misura **y**

- $f(\mathbf{y}|\mathbf{x})$ likelihood/verosimiglianza
- $f(\mathbf{x})$ prior (bayesiano)
- $f(\mathbf{y})$ distribuzione marginale, $f(\mathbf{y}) = \sum_{\mathbf{x}} f(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{x}} f(\mathbf{y}|\mathbf{x})f(\mathbf{x})$

teorema di Bayes

in alcuni casi la probabilità che accada qualcosa si modifica se, per qualunque motivo, si viene a conoscenza di una informazione aggiuntiva legata all'evento

esempio classico - lancio di dadi

- probabilità che il risultato sia il numero 6: $P(6) = \frac{1}{6}$
- probabilità che il risultato sia il numero 6 sapendo che è uscito un numero pari: $P(6|\text{pari}) = \frac{p(6 \cap \text{pari})}{p(\text{pari})} = \frac{1/6}{1/2} = \frac{1}{3}$

teorema di Bayes

in molte applicazioni un dato è facilmente ottenibile mentre un altro no

esempio meno classico ma più istruttivo - tumori dovuti al fumo

▷ probabilità di avere **un** tumore dato che si fuma?

$$P(T|F) = \frac{p(F|T)p(T)}{p(F)}$$

$p(F\|T)$	essere fumatori dato che si ha **un certo tipo** di tumore
$p(T)$	avere **un certo tipo** di tumore
$p(F)$	essere fumatore

⚠ probabilità di avere **un certo tipo** di tumore dato che si fuma?

teorema di Bayes

fun fact : il teorema di Bayes è accettato nei tribunali

l'errore del procuratore - la fallacia del pubblico ministero

confondere (più o meno consapevolmente) la probabilità che l'imputato sia colpevole, dati gli indizi a suo carico, con la probabilità che si ottengano quegli indizi, se l'imputato è colpevole

$P(C|I) \neq P(I|C)$ C: colvevolezza, I: indizio

se ho avvelenato il gatto del vicino, è molto probabile che poco prima io abbia comprato del veleno, ma se ho comprato del veleno non è detto che sia stato io a ucciderlo

$\otimes$

Stima Bayesiana

- approccio fisheriano VS. approccio bayesiano
- stima MAP - Maximum A Posteriori
- stima MMSE - Minimum Mean Square Error

stima MAP

stima del massimo - della probabilità - a posteriori

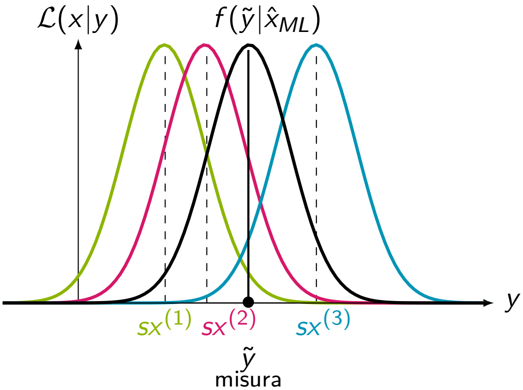
(o MAP - maximum a posteriori)

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}|\mathbf{y})$$

applicazione diretta del teorema di Bayes

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}|\mathbf{y}) = \arg\max_{\mathbf{x}\in\mathcal{X}} \frac{f(\mathbf{y}|\mathbf{x})f(\mathbf{x})}{f(\mathbf{y})} = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{y}|\mathbf{x})f(\mathbf{x})$$

esempio - caso scalare SISO: $x \in \mathbb{R}$, $y = sx + v \in \mathbb{R}$

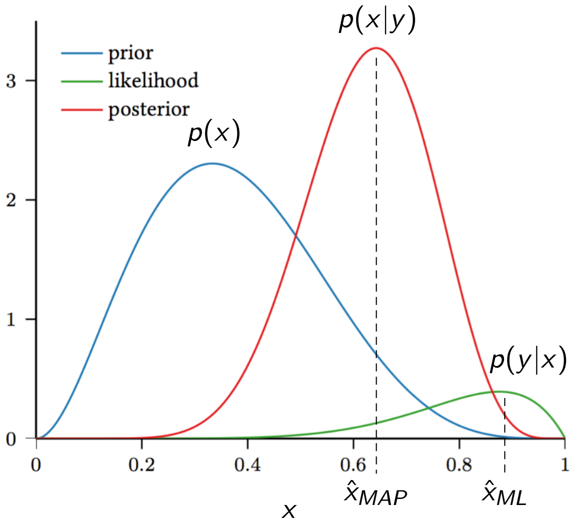


equivalenza con stima ML
se non si hanno infomazioni su $\mathbf{x}$ prima delle misure
allora la stima MAP coincide con la stima ML

$\Longrightarrow$ se non si hanno infomazioni su $\mathbf{x}$ prima delle misure,
la densità di probabilità a priori di $\mathbf{x}$ è uniforme $\rightarrow f(\mathbf{x}) = const.$

stima MAP: $\hat{\mathbf{x}}_{MAP} = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}|\mathbf{y}) = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{y}|\mathbf{x})$

$$\hat{\mathbf{x}}_{MAP} = \hat{\mathbf{x}}_{ML}$$


prior
likelihood
posterior
$p(x|y)$
$p(x)$
$p(y|x)$
3
2
1
0
0
0.2
0.4
0.6
0.8
1
$\hat{x}_{MAP}$
$\hat{x}_{ML}$
x

Stima Bayesiana

- approccio fisheriano VS. approccio bayesiano
- stima MAP - Maximum A Posteriori
- stima MMSE - Minimum Mean Square Error

stima MMSE

stima del minimo errore quadratico medio

(o MMSE - minimum mean square error)

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{X}} \mathbb{E}\left[\|\mathbf{x} - \mathbf{x}_r\|^2\right]$$

$$MSE = \mathbb{E}\left[\|\hat{\mathbf{x}} - \mathbf{x}_r\|^2\right]$$ errore quadratico medio - *mean square error*

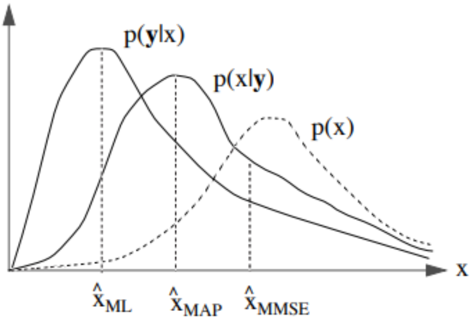
$$MSE = \mathbb{E}\left[\|\hat{\mathbf{x}} - \mathbf{x}_r\|^2\right] = \mathbb{E}\left[\mathbb{E}\left[\|\hat{\mathbf{x}}(\mathbf{y}) - \mathbf{x}_r\|^2 \middle| \mathbf{y}\right]\right]$$

▷ calcolare $\hat{\mathbf{x}}$ che minimizza $\mathbb{E}\left[\mathbb{E}\left[\|\hat{\mathbf{x}}(\mathbf{y}) - \mathbf{x}_r\|^2 \middle| \mathbf{y}\right]\right]$ per ogni $\mathbf{y}$

$$\begin{aligned}\frac{\partial\left(\mathbb{E}\left[\mathbb{E}\left[\|\hat{\mathbf{x}}-\mathbf{x}_r\|^2 \middle| \mathbf{y}\right]\right]\right)}{\partial \hat{\mathbf{x}}} &= \mathbb{E}\left[\frac{\partial\left(\|\hat{\mathbf{x}}-\mathbf{x}_r\|^2\right)}{\partial \hat{\mathbf{x}}} \middle| \mathbf{y}\right] = \mathbb{E}\left[\frac{\partial\left((\hat{\mathbf{x}}-\mathbf{x}_r)^\top(\hat{\mathbf{x}}-\mathbf{x}_r)\right)}{\partial \hat{\mathbf{x}}} \middle| \mathbf{y}\right] = \mathbb{E}\left[\frac{\partial\left(\hat{\mathbf{x}}^\top\hat{\mathbf{x}}-2\mathbf{x}_r^\top\hat{\mathbf{x}}+\mathbf{x}_r^\top\mathbf{x}_r\right)}{\partial \hat{\mathbf{x}}} \middle| \mathbf{y}\right] \\ &= \mathbb{E}\left[2\hat{\mathbf{x}} - 2\mathbf{x}_r | \mathbf{y}\right] = 2\mathbb{E}\left[(\hat{\mathbf{x}} - \mathbf{x}_r) | \mathbf{y}\right] = 2\mathbb{E}\left[\hat{\mathbf{x}} | \mathbf{y}\right] - 2\mathbb{E}\left[\mathbf{x}_r | \mathbf{y}\right] \\ &= 2\mathbb{E}\left[\hat{\mathbf{x}} | \mathbf{y}\right] - 2\mathbf{x}_r = 0 \quad \Leftrightarrow \quad \mathbb{E}\left[\hat{\mathbf{x}} | \mathbf{y}\right] = \mathbf{x}_r\end{aligned}$$

stimatore MMSE

$$\hat{\mathbf{x}} = \mathbb{E}\left[\mathbf{x} | \mathbf{y}\right]$$



stima MMSE: centro di massa della probabilità a posteriori
stima MAP: argomento corrispondente al massimo della probabilità a posteriori
stima ML: argomento corrispondente al massimo della likelihood

equivalenza con stima MAP
dato il modello dei dati $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$ con $\mathbf{H} \in \mathbb{R}^{p \times n}$,
se $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{v})$ e $\mathbf{x} \sim \mathcal{N}(\mathbf{x}_r, \boldsymbol{\Sigma}_\mathbf{x})$
allora la stima MMSE coincide con la stima MAP

$\implies$ sotto le ipotesi date
la probabilità $f(\mathbf{x}|\mathbf{y})$ è Gaussiana con

- media $\quad \mathbf{m}_{\mathbf{x}|\mathbf{y}} = \mathbb{E}[\mathbf{x}|\mathbf{y}] = \left(\mathbf{H}^\top \boldsymbol{\Sigma}_\mathbf{v}^{-1}\mathbf{H} + \boldsymbol{\Sigma}_\mathbf{x}^{-1}\right)^{-1} \left(\mathbf{H}^\top \boldsymbol{\Sigma}_\mathbf{v}^{-1}\mathbf{y} + \boldsymbol{\Sigma}_\mathbf{x}^{-1}\mathbf{x}_r\right)$
- varianza $\quad \boldsymbol{\Sigma}_{\mathbf{x}|\mathbf{y}} = \left(\mathbf{H}^\top \boldsymbol{\Sigma}_\mathbf{v}^{-1}\mathbf{H} + \boldsymbol{\Sigma}_\mathbf{x}^{-1}\right)^{-1}$

$$\hat{\mathbf{x}}_{MMSE} = \hat{\mathbf{x}}_{MAP}$$

caso gaussiano

modello dei dati

$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}, \quad \mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{v}}), \quad \mathbf{x} \sim \mathcal{N}(\mathbf{x}_r, \mathbf{\Sigma}_{\mathbf{x}})$

$$f(\mathbf{x}|\mathbf{y}) \sim \mathcal{N}(\mathbf{m}_{\mathbf{x}|\mathbf{y}}, \mathbf{\Sigma}_{\mathbf{x}|\mathbf{y}}) \qquad \begin{aligned} \mathbf{m}_{\mathbf{x}|\mathbf{y}} &= \left(\mathbf{H}^\top \mathbf{\Sigma}_{\mathbf{v}}^{-1} \mathbf{H} + \mathbf{\Sigma}_{\mathbf{x}}^{-1}\right)^{-1} \left(\mathbf{H}^\top \mathbf{\Sigma}_{\mathbf{v}}^{-1} \mathbf{y} + \mathbf{\Sigma}_{\mathbf{x}}^{-1} \mathbf{x}_r\right) \\ \mathbf{\Sigma}_{\mathbf{x}|\mathbf{y}} &= \left(\mathbf{H}^\top \mathbf{\Sigma}_{\mathbf{v}}^{-1} \mathbf{H} + \mathbf{\Sigma}_{\mathbf{x}}^{-1}\right)^{-1} \end{aligned}$$

$$\boxed{\hat{\mathbf{x}}_{MMSE} = \mathbf{m}_{\mathbf{x}|\mathbf{y}} = \left(\mathbf{H}^\top \mathbf{\Sigma}_{\mathbf{v}}^{-1} \mathbf{H} + \mathbf{\Sigma}_{\mathbf{x}}^{-1}\right)^{-1} \left(\mathbf{H}^\top \mathbf{\Sigma}_{\mathbf{v}}^{-1} \mathbf{y} + \mathbf{\Sigma}_{\mathbf{x}}^{-1} \mathbf{x}_r\right)}$$

$$\bullet \boldsymbol{\Sigma}_{\mathbf{x}} - \boldsymbol{\Sigma}_{\mathbf{x}|\mathbf{y}} = \boldsymbol{\Sigma}_{\mathbf{x}} - \left(\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\mathbf{H} + \boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\right)^{-1}$$

$$= \boldsymbol{\Sigma}_{\mathbf{x}} - \left(\boldsymbol{\Sigma}_{\mathbf{x}} - \boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{H}^\top(\boldsymbol{\Sigma}_{\mathbf{v}} + \mathbf{H}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{H}^\top)^{-1}\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{x}}\right)$$

$$= \boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{H}^\top(\boldsymbol{\Sigma}_{\mathbf{v}} + \mathbf{H}\boldsymbol{\Sigma}_{\mathbf{x}}\mathbf{H}^\top)^{-1}\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{x}}$$

$\Rightarrow \boldsymbol{\Sigma}_{\mathbf{x}} - \boldsymbol{\Sigma}_{\mathbf{x}|\mathbf{y}}$ (almeno) sdp: la misura diminuisce in media l'incertezza

- $\boldsymbol{\Sigma}_{\mathbf{x}|\mathbf{y}}^{-1}\hat{\mathbf{x}}_{MMSE} = \mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\mathbf{y} + \boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{x}_r$

$$= \mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\mathbf{y} + \boldsymbol{\Sigma}_{\mathbf{x}}^{-1}\mathbf{x}_r + \mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}\mathbf{H}\mathbf{x}_r - \mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}\mathbf{H}\mathbf{x}_r$$

$$\Rightarrow \; \hat{\mathbf{x}}_{MMSE} = \mathbf{x}_r + \boldsymbol{\Sigma}_{\mathbf{x}|\mathbf{y}}\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\underbrace{(\mathbf{y} - \mathbf{H}\mathbf{x}_r)}_{\text{innovazione}}$$

siano

$\hat{\mathbf{x}}(t) \in \mathbb{R}^n$	stima MMSE basata sulle misure correnti
$\boldsymbol{\Sigma}_{\hat{\mathbf{x}}}(t) \in \mathbb{R}^{n \times n}$	varianza della stima MMSE basata sulle misure correnti
$\mathbf{y}_{t+1} \in \mathbb{R}^p$	misure successive

$$\begin{aligned}\boldsymbol{\Sigma}_{\hat{\mathbf{x}}}(t+1) &= \left(\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} \mathbf{H} + \boldsymbol{\Sigma}_{\hat{\mathbf{x}}}(t)^{-1}\right)^{-1} \\ \hat{\mathbf{x}}(t+1) &= \hat{\mathbf{x}}(t) + \boldsymbol{\Sigma}_{\hat{\mathbf{x}}}(t+1)\mathbf{H}^\top \boldsymbol{\Sigma}_{v}^{-1}\left(\mathbf{y}_{t+1} - \mathbf{H}\hat{\mathbf{x}}(t)\right)\end{aligned}$$

Stima secondo Kalman

- filtro di Kalman per sistemi a tempo discreto
  - derivazione delle formule
  - interpretazione come stimatore ad anello chiuso
  - proprietà del filtro di Kalman
- filtro di Kalman per sistemi a tempo continuo
- controllo ottimo LQG

problema di stima dello stato di un sistema dinamico

⚠ $\mathbf{z}_k = \mathbf{z}(k), \quad \mathbf{z}^k = \{\mathbf{z}_0 \ldots \mathbf{z}_k\}$

sistema LTI a tempo discreto

$$\mathbf{x}(t+1) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t)$$
$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t)$$
$$\mathbf{x}(0) = \mathbf{x}_0$$

$\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^p$

$\mathbf{w} \in \mathbb{R}^n$: errore di processo

$\mathbf{v} \in \mathbb{R}^p$: errore di misura

stima dello stato: inferire informazioni sullo stato $\mathbf{x}_t$ dalle le osservazioni $\mathbf{y}^{t'}$

- *predizione* se $t > t'$
- *filtraggio* se $t = t'$
- *regolarizzazione* se $t < t'$

la determinazione di $\mathbf{x}_t$ sarebbe banale se non fosse per la presenza degli errori
$\rightarrow$ sfruttare le equazioni di stato **&** le equazioni di misura

ipotesi sul sistema dinamico

sistema LTI a tempo discreto

$$\mathbf{x}(t+1) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t)$$
$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t)$$
$$\mathbf{x}(0) = \mathbf{x}_0$$

$\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^p$

$\mathbf{w} \in \mathbb{R}^n$: errore di processo

$\mathbf{v} \in \mathbb{R}^p$: errore di misura

1. errore/rumore di processo
processo stocastico Gaussiano, stazionario, bianco e a media nulla

- $\mathbb{E}\left[\mathbf{w}(t)\right] = \mathbf{0} \ \forall t$
- $\mathbb{E}\left[\mathbf{w}(t)\mathbf{w}(t)^\top\right] = \mathbf{\Sigma}_\mathbf{w} \ \forall t$
- $\mathbb{E}\left[\mathbf{w}(t_1)\mathbf{w}(t_2)^\top\right] = \mathbf{0} \ \forall t_1 \neq t_2$

$$\mathbf{w}(t) \in \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_\mathbf{w})$$

sistema LTI a tempo discreto

$$\mathbf{x}(t+1) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t)$$
$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t)$$
$$\mathbf{x}(0) = \mathbf{x}_0$$

$\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^p$

$\mathbf{w} \in \mathbb{R}^n$: errore di processo

$\mathbf{v} \in \mathbb{R}^p$: errore di misura

2. errore/rumore di misura
processo stocastico Gaussiano, stazionario, bianco e a media nulla

- $\mathbb{E}[\mathbf{v}(t)] = \mathbf{0}\ \forall t$
- $\mathbb{E}\left[\mathbf{v}(t)\mathbf{v}(t)^\top\right] = \mathbf{\Sigma}_\mathbf{v}\ \forall t$
- $\mathbb{E}\left[\mathbf{v}(t_1)\mathbf{v}(t_2)^\top\right] = \mathbf{0}\ \forall t_1 \neq t_2$

$$\boxed{\mathbf{v}(t) \in \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_\mathbf{v})}$$

sistema LTI a tempo discreto

$$\mathbf{x}(t+1) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t)$$
$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t)$$
$$\mathbf{x}(0) = \mathbf{x}_0$$

$\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^p$
$\mathbf{w} \in \mathbb{R}^n$: errore di processo
$\mathbf{v} \in \mathbb{R}^p$: errore di misura

3. errore/rumore di processo e di misura
processi stocastici Gaussiani, stazionari, bianco e a media nulla & *indipendenti*

$$\mathbb{E}\left[\mathbf{w}(t_1)\mathbf{v}(t_2)^\top\right] = \mathbf{0} \ \forall t_1 \neq t_2$$

4. stato iniziale
variabile aleatoria Gaussiana di media e varianza note

$$\mathbf{x}(0) \in \mathcal{N}(\mathbf{x}_0, \mathbf{\Sigma}_0)$$

▶ stima dello stato in base a $\mathbf{y}^t = \{\mathbf{y}(0) \dots \mathbf{y}(t)\}$

noto $\mathbf{y}^t = \{\mathbf{y}(0) \dots \mathbf{y}(t)\}$

$\tilde{\mathbf{x}}(t+1|t) = \hat{\mathbf{x}}(t+1|t) - \mathbf{x}(t+1)$ errore di stima a $t+1$

$\mathbf{P}(t+1|t) = \mathbb{E}\left[\tilde{\mathbf{x}}(t+1|t)\tilde{\mathbf{x}}(t+1|t)^\top\right]$ varianza della dell'errore

allora, in assenza di nuove misure & in base alla dinamica del sistema

$$\boxed{\hat{\mathbf{x}}(t+1|t) = \mathbf{F}\hat{\mathbf{x}}(t|t) + \mathbf{G}\mathbf{u}(t)}$$ aggiornamento della stima

$$\tilde{\mathbf{x}}(t+1|t) = \mathbf{F}\hat{\mathbf{x}}(t|t) - \mathbf{F}\mathbf{x}(t) - \mathbf{w}(t) = \mathbf{F}\tilde{\mathbf{x}}(t|t) - \mathbf{w}(t)$$

$$\begin{aligned}
\mathbf{P}(t+1|t) &= \mathbb{E}\left[\left(\mathbf{F}\tilde{\mathbf{x}}(t|t) - \mathbf{w}(t)\right)\left(\mathbf{F}\tilde{\mathbf{x}}(t|t) - \mathbf{w}(t)\right)^\top\right] \\
&= \mathbb{E}\left[\mathbf{F}\tilde{\mathbf{x}}(t|t)\tilde{\mathbf{x}}(t|t)^\top\mathbf{F}^\top - \mathbf{F}\tilde{\mathbf{x}}(t|t)\mathbf{w}(t)^\top - \mathbf{w}(t)\tilde{\mathbf{x}}(t|t)^\top\mathbf{F}^\top + \mathbf{w}(t)\mathbf{w}(t)^\top\right] \\
&= \mathbb{E}\left[\mathbf{F}\tilde{\mathbf{x}}(t|t)\tilde{\mathbf{x}}(t|t)^\top\mathbf{F}^\top + \mathbf{w}(t)\mathbf{w}(t)^\top\right]
\end{aligned}$$

$$\boxed{\mathbf{P}(t+1|t) = \mathbf{F}\mathbf{P}(t|t)\mathbf{F}^\top + \boldsymbol{\Sigma}_{\mathbf{w}}}$$

aggiornamento della varianza di stima

▶ stima dello stato dato $\mathbf{y}_{t+1} = \mathbf{y}(t+1)$

dato $\mathbf{y}_{t+1} = \mathbf{y}(t+1)$

$$\tilde{\mathbf{x}}(t+1|t+1) = \hat{\mathbf{x}}(t+1|t+1) - \mathbf{x}(t+1) \qquad \text{errore di stima a } t+1$$

$$\mathbf{P}(t+1|t+1) = \mathbb{E}\left[\tilde{\mathbf{x}}(t+1|t+1)\tilde{\mathbf{x}}(t+1|t+1)^\top\right] \qquad \text{varianza della dell'errore}$$

allora, sfruttando la nuova misura & la stima MMSE ricorsiva

$$\mathbf{P}(t+1|t+1) = \left(\mathbf{H}^\top \mathbf{\Sigma}_\mathbf{v}^{-1}\mathbf{H} + \mathbf{P}(t+1|t)^{-1}\right)^{-1}$$

$$\hat{\mathbf{x}}(t+1|t+1) = \hat{\mathbf{x}}(t+1|t) + \mathbf{P}(t+1|t+1)\mathbf{H}^\top \mathbf{\Sigma}_\mathbf{v}^{-1}\left(\mathbf{y}(t+1) - \mathbf{H}\hat{\mathbf{x}}(t+1|t)\right)$$

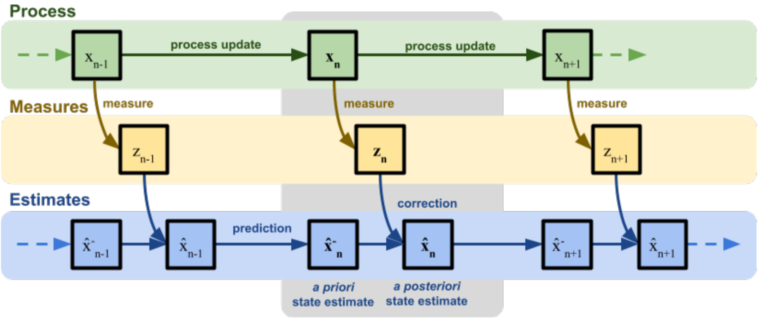
filtro di Kalman

stime a priori (*aggiornamento temporale*)

$$\mathbf{P}(t+1|t) = \mathbf{F}\mathbf{P}(t|t)\mathbf{F}^\top + \mathbf{\Sigma}_\mathbf{w}$$
$$\hat{\mathbf{x}}(t+1|t) = \mathbf{F}\hat{\mathbf{x}}(t|t) + \mathbf{G}\mathbf{u}(t)$$

stime a posteriori (*aggiornamento rispetto alle misure*)

$$\mathbf{P}(t+1|t+1) = \left(\mathbf{H}^\top\mathbf{\Sigma}_\mathbf{v}^{-1}\mathbf{H} + \mathbf{P}(t+1|t)^{-1}\right)^{-1}$$
$$\hat{\mathbf{x}}(t+1|t+1) = \hat{\mathbf{x}}(t+1|t) + \mathbf{P}(t+1|t+1)\mathbf{H}^\top\mathbf{\Sigma}_\mathbf{v}^{-1}(\mathbf{y}(t+1) - \mathbf{H}\hat{\mathbf{x}}(t+1|t))$$



Rudolf Emil Kálmán (1930 – 2016)
ingegnere e matematico
statunitense

Right: Rudolf Kalman, inventor of the Kalman Filter, receives the National Medal of Science from President Obama.

Left: Margaret Hamilton, one of the leading minds who developed the Apollo on-board flight software, sits in the Apollo module. The navigation system interface is above her head, below the circular attitude indicator, running Kalman's filter.

Stima secondo Kalman

- filtro di Kalman per sistemi a tempo discreto
  - derivazione delle formule
  - interpretazione come stimatore ad anello chiuso
  - proprietà del filtro di Kalman
- filtro di Kalman per sistemi a tempo continuo
- controllo ottimo LQG

TdS: stimatore ad anello chiuso

$$\Sigma: \begin{cases} x(t+1) = Fx(t) + Gu(t) \\ y(t) = Hx(t) \end{cases}$$

stimatore ad anello chiuso

$$\boxed{\hat{\Sigma}: \begin{cases} \hat{x}(t+1) = F\hat{x}(t) + Gu(t) - L(y(t) - H\hat{x}(t)) \\ \hat{y}(t) = \hat{x}(t) \end{cases}}$$

$L \in \mathbb{R}^{n \times p}$ = guadagno dello stimatore

errore di stima: $e(t) \triangleq x(t) - \hat{x}(t)$

$$e(t+1) = x(t+1) - \hat{x}(t+1) = F(x(t) - \hat{x}(t)) + LH(x(t) - \hat{x}(t)) = (F + LH)e(t)$$

errore di stima $e(t)$ tende a zero se $F + LH$ è asintoticamente stabile (e in questo caso F può anche essere instabile) !!!

- se il sistema (la coppia (F, H)) è **osservabile**
  allora è *sempre* possible calcolare un guadagno L
  in grado di rendere $F + LH$ asintoticamente stabile

- se il sistema (la coppia (F, H)) non è osservabile
  ma il sistema (la coppia (F, H)) è **rivelabile**
  allora è *ancora* possible calcolare un guadagno L
  in grado di rendere $F + LH$ asintoticamente stabile

▶ **come calcolare il guadagno L?**

sintesi mediante assegnazione degli autovalori

POCO INTUITIVO

linee guida

- autovalori (sufficientemente) grandi $\rightarrow$ rapidità di convergenza della stima
- autovalori (sufficientemente) piccoli $\rightarrow$ insensibilità al rumore di misura

▶ **assumendo di avere infomazioni sul rumore come calcolare il guadagno L?**

filtro di Kalman
(anche sistemi multidimensionali MIMO)

filtro di Kalman

stime a priori (*aggiornamento temporale*)

$$\mathbf{P}(t+1|t) = \mathbf{F}\mathbf{P}(t|t)\mathbf{F}^\top + \mathbf{\Sigma_w}$$
$$\hat{\mathbf{x}}(t+1|t) = \mathbf{F}\hat{\mathbf{x}}(t|t) + \mathbf{G}\mathbf{u}(t)$$

stime a posteriori (*aggiornamento rispetto alle misure*)

$$\mathbf{P}(t+1|t+1) = \left(\mathbf{H}^\top \mathbf{\Sigma_v}^{-1}\mathbf{H} + \mathbf{P}(t+1|t)^{-1}\right)^{-1}$$
$$\hat{\mathbf{x}}(t+1|t+1) = \hat{\mathbf{x}}(t+1|t) + \mathbf{P}(t+1|t+1)\mathbf{H}^\top \mathbf{\Sigma_v}^{-1}\left(\mathbf{y}(t+1) - \mathbf{H}\hat{\mathbf{x}}(t+1|t)\right)$$

interpretazione come stimatore ad anello chiuso

premessa: lemma di inversione $(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1}+\mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}$

$$
\begin{aligned}
\mathbf{P}(t+1|t+1) &= \left(\mathbf{H}^\top \boldsymbol{\Sigma}_\mathbf{v}^{-1}\mathbf{H} + \mathbf{P}(t+1|t)^{-1}\right)^{-1} \\
&= \mathbf{P}(t+1|t) - \mathbf{P}(t+1|t)\mathbf{H}^\top \left(\boldsymbol{\Sigma}_\mathbf{v} + \mathbf{HP}(t+1|t)\mathbf{H}^\top\right)^{-1}\mathbf{HP}(t+1|t) \\
&= \mathbf{P}(t+1|t) - \mathbf{P}(t+1|t)\mathbf{H}^\top \left(\mathbf{S}(t+1)\right)^{-1}\mathbf{HP}(t+1|t)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{L}'(t+1) &= \mathbf{P}(t+1|t+1)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} \\
&= \mathbf{P}(t+1|t)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} - \mathbf{P}(t+1|t)\mathbf{H}^\top \left(\mathbf{S}(t+1)\right)^{-1} \mathbf{HP}(t+1|t)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} \\
&= \mathbf{P}(t+1|t)\mathbf{H}^\top \left(\mathbf{S}(t+1)\right)^{-1} \left(\mathbf{S}(t+1)\boldsymbol{\Sigma}_{\mathbf{v}}^{-1} - \mathbf{HP}(t+1|t)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1}\right) \\
&\vdots
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{S}(t+1)\boldsymbol{\Sigma}_{\mathbf{v}}^{-1} - \mathbf{HP}(t+1|t)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} &= \left(\boldsymbol{\Sigma}_{\mathbf{v}} + \mathbf{HP}(t+1|t)\mathbf{H}^\top\right) \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} - \mathbf{HP}(t+1|t)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} \\
&= \boldsymbol{\Sigma}_{\mathbf{v}}\boldsymbol{\Sigma}_{\mathbf{v}}^{-1} + \mathbf{HP}(t+1|t)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} - \mathbf{HP}(t+1|t)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} = \mathbf{I}
\end{aligned}
$$

$$
\begin{aligned}
&\vdots \\
&= \mathbf{P}(t+1|t)\mathbf{H}^\top \left(\mathbf{S}(t+1)\right)^{-1} = \mathbf{P}(t+1|t)\mathbf{H}^\top \left(\boldsymbol{\Sigma}_{\mathbf{v}} + \mathbf{HP}(t+1|t)\mathbf{H}^\top\right)^{-1}
\end{aligned}
$$

$$
\begin{aligned}
\hat{\mathbf{x}}(t+1|t+1) &= \hat{\mathbf{x}}(t+1|t) + \mathbf{P}(t+1|t+1)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} \left(\mathbf{y}(t+1) - \mathbf{H}\hat{\mathbf{x}}(t+1|t)\right) \\
&= \hat{\mathbf{x}}(t+1|t) + \mathbf{L}'(t+1) \left(\mathbf{y}(t+1) - \mathbf{H}\hat{\mathbf{x}}(t+1|t)\right) \\
\hat{\mathbf{x}}(t+1|t) &= \mathbf{F}\hat{\mathbf{x}}(t|t) + \mathbf{G}\mathbf{u}(t) \\
\hat{\mathbf{x}}(t|t) &= \hat{\mathbf{x}}(t|t-1) + \mathbf{P}(t|t)\mathbf{H}^\top \boldsymbol{\Sigma}_{\mathbf{v}}^{-1} \left(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t|t-1)\right) \\
&= \hat{\mathbf{x}}(t|t-1) + \mathbf{L}'(t) \left(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t+1|t)\right)
\end{aligned}
$$

$$
\boxed{
\begin{aligned}
\hat{\mathbf{x}}(t+1|t) &= \mathbf{F}\hat{\mathbf{x}}(t|t-1) + \mathbf{F}\mathbf{L}'(t) \left(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t|t-1)\right) + \mathbf{G}\mathbf{u}(t) \\
&= \mathbf{F}\hat{\mathbf{x}}(t|t-1) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}(t) \left(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t|t-1)\right)
\end{aligned}
}
$$

$$
\mathbf{L}(t) = \mathbf{F}\mathbf{L}'(t) = \mathbf{F}\mathbf{P}(t|t-1)\mathbf{H}^\top \left(\boldsymbol{\Sigma}_{\mathbf{v}} + \mathbf{H}\mathbf{P}(t|t-1)\mathbf{H}^\top\right)^{-1} \text{ guadagno tempo variante}
$$

$$
\begin{aligned}
\mathbf{P}(t+1|t+1) &= \left(\mathbf{H}^\top \mathbf{\Sigma}_\mathbf{v}^{-1}\mathbf{H} + \mathbf{P}(t+1|t)^{-1}\right)^{-1} \\
&= \mathbf{P}(t+1|t) - \mathbf{P}(t+1|t)\mathbf{H}^\top \left(\mathbf{S}(t+1)\right)^{-1} \mathbf{H}\mathbf{P}(t+1|t) \\
\mathbf{P}(t+1|t) &= \mathbf{F}\mathbf{P}(t|t)\mathbf{F}^\top + \mathbf{\Sigma}_\mathbf{w} \\
\mathbf{P}(t|t) &= \left(\mathbf{H}^\top \mathbf{\Sigma}_\mathbf{v}^{-1}\mathbf{H} + \mathbf{P}(t|t-1)^{-1}\right)^{-1} \\
&= \mathbf{P}(t|t-1) - \mathbf{P}(t|t-1)\mathbf{H}^\top \left(\mathbf{S}(t)\right)^{-1} \mathbf{H}\mathbf{P}(t|t-1)
\end{aligned}
$$

$$
\boxed{\mathbf{P}(t+1|t) = \mathbf{F}\mathbf{P}(t|t-1)\mathbf{F}^\top - \mathbf{F}\mathbf{P}(t|t-1)\mathbf{H}^\top \left(\mathbf{S}(t)\right)^{-1} \mathbf{H}\mathbf{P}(t|t-1)\mathbf{F}^\top + \mathbf{\Sigma}_\mathbf{w}}
$$

$$
\begin{aligned}
\left(\mathbf{S}(t)\right. &= \left.\boldsymbol{\Sigma}_{\mathbf{v}} + \mathbf{H}\mathbf{P}(t|t-1)\mathbf{H}^\top\right) \\
\mathbf{L}(t) &= \mathbf{F}\mathbf{P}(t|t-1)\mathbf{H}^\top \left(\mathbf{S}(t)\right)^{-1} \\
\hat{\mathbf{x}}(t+1|t) &= \mathbf{F}\hat{\mathbf{x}}(t|t-1) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}(t)\left(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t|t-1)\right) \\
\mathbf{P}(t+1|t) &= \mathbf{F}\mathbf{P}(t|t-1)\mathbf{F}^\top - \mathbf{F}\mathbf{P}(t|t-1)\mathbf{H}^\top \left(\mathbf{S}(t)\right)^{-1} \mathbf{H}\mathbf{P}(t|t-1)\mathbf{F}^\top + \boldsymbol{\Sigma}_{\mathbf{w}}
\end{aligned}
$$

Stima secondo Kalman

- filtro di Kalman per sistemi a tempo discreto
  - derivazione delle formule
  - interpretazione come stimatore ad anello chiuso
  - proprietà del filtro di Kalman
- filtro di Kalman per sistemi a tempo continuo
- controllo ottimo LQG

ottimalità

Dato il sistema

$$\mathbf{x}(t+1) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0$$
$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t)$$

tale che

- ipotesi di gaussianità: $\mathbf{x}(0) \sim \mathcal{N}(\mathbf{x}_0, \mathbf{\Sigma}_0)$, $\mathbf{w}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_\mathbf{w})$, $\mathbf{v}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_\mathbf{v})$;
- ipotesi di incorrelazione: $\mathbf{w}(t)$, $\mathbf{v}(t)$ sono mutuamente incorrelati e incorrelati con $\mathbf{x}(0)$.

Il filtro di Kalman interpretato come stimatore ad anello chiuso e inizializzato con $\hat{\mathbf{x}}(0) = \mathbf{x}_0$ e $\mathbf{P}(0) = \mathbf{\Sigma}_0$ risulta essere l'osservatore che minimizza ad ogni istante $t > 0$ l'errore quadratico medio (MSE) di stima dello stato.

guadagno costante

se $\mathbf{L}(t) = \mathbf{L}$ (guadagno costante)
allora $\mathbf{P}(t+1|t) = \mathbf{P}$: no aggiornamento della varianza

$\mathbf{P} = \mathbf{P}^\top$ sdp: soluzione della EARD

$$\mathbf{P} = \mathbf{FPF}^\top - \mathbf{FPH}^\top(\Sigma_\mathbf{v} + \mathbf{HPH}^\top)^{-1}\mathbf{HPF}^\top + \boldsymbol{\Sigma}_\mathbf{w}$$

se $(\mathbf{F}, \mathbf{G})$ stabilizzabile, $\mathbf{GG}^\top = \boldsymbol{\Sigma}_\mathbf{w}$ e $(\mathbf{F}, \mathbf{H})$ rivelabile
allora il filtro di Kalman è asintoticamente stabile

dualità controllo-stima ottima

$$\mathbf{x}(t+1) = \mathbf{F}_s\mathbf{x}(t) + \mathbf{G}_s\mathbf{u}(t) \qquad \mathbf{x}(t) \in \mathbb{R}^n,\ \mathbf{u}(t) \in \mathbb{R}^m,\ \mathbf{y}(t) \in \mathbb{R}^p$$
$$\mathbf{y}(t) = \mathbf{H}_s\mathbf{x}(t)$$

stima alla Kalman nel caso stazionario

$$\mathbf{L} = \mathbf{F}_s\mathbf{P}\mathbf{H}^\top \left(\mathbf{\Sigma}_\mathbf{v} + \mathbf{H}\mathbf{P}\mathbf{H}^\top\right)^{-1}$$
$$\hat{\mathbf{x}}(t+1|t) = \mathbf{F}_s\hat{\mathbf{x}}(t|t-1) + \mathbf{G}_s\mathbf{u}(t) + \mathbf{L}\left(\mathbf{y}(t) - \mathbf{H}_s\hat{\mathbf{x}}(t|t-1)\right)$$
$$\mathbf{P} = \mathbf{F}_s\mathbf{P}\mathbf{F}_s^\top - \mathbf{F}_s\mathbf{P}\mathbf{H}_s^\top \left(\mathbf{\Sigma}_\mathbf{v} + \mathbf{H}_s\mathbf{P}\mathbf{H}_s^\top\right)^{-1} \mathbf{H}_s\mathbf{P}\mathbf{F}_s^\top + \mathbf{\Sigma}_\mathbf{w}$$

$$\mathbf{x}(t+1) = \mathbf{F}_c\mathbf{x}(t) + \mathbf{G}_c\mathbf{u}(t) \qquad\qquad \mathbf{x}(t) \in \mathbb{R}^n,\ \mathbf{u}(t) \in \mathbb{R}^m,\ \mathbf{y}(t) \in \mathbb{R}^p$$
$$\mathbf{y}(t) = \mathbf{H}_c\mathbf{x}(t)$$

controllo LQ a orizzonte infinito

$$\mathbf{u}^\star_\infty(t) = \arg\min_{t\in[0,+\infty)} J_\infty(\mathbf{u}(t)), \quad J_\infty(\mathbf{u}(t)) = \sum_{t=0}^{+\infty}\left(\mathbf{x}(t)^\top\mathbf{Q}\mathbf{x}(t) + \mathbf{u}(t)^\top\mathbf{R}\mathbf{u}(t)\right)$$
$$\mathbf{Q} \in \mathbb{R}^{n\times n} \text{ sdp},\ \mathbf{R} \in \mathbb{R}^{m\times m} \text{ dp}$$

se $(\mathbf{F}_c, \mathbf{G}_c)$ stabilizzabile, $(\mathbf{F}_c, \mathbf{H}_c)$ rivelabile con $\mathbf{Q} = \mathbf{H}_c\mathbf{H}_c^\top$, allora

$$\mathbf{u}^\star_\infty(t) = -\mathbf{K}^\star_\infty\mathbf{x}(t) \quad \text{con} \quad \mathbf{K}^\star_\infty = \left(\mathbf{R} + \mathbf{G}_c^\top M_\infty \mathbf{G}_c\right)^{-1}\mathbf{G}_c^\top M_\infty \mathbf{F}_c$$

dove $\mathbf{M}_\infty = \mathbf{M}_\infty^\top$ è l'unica soluzione della EARD

$$\mathbf{M} = \mathbf{F}_c^\top\mathbf{M}\mathbf{F}_c - \mathbf{F}_c^\top\mathbf{M}\mathbf{G}_c\left(\mathbf{R} + \mathbf{G}_c^\top\mathbf{M}\mathbf{G}_c\right)^{-1}\mathbf{G}_c^\top\mathbf{M}\mathbf{F}_c + \mathbf{Q}$$

guadagno controllore: $\mathbf{K}_\infty^\star = \left(\mathbf{R} + \mathbf{G}_c^\top \mathbf{M}_\infty \mathbf{G}_c\right)^{-1} \mathbf{G}_c^\top \mathbf{M}_\infty \mathbf{F}_c$ dove $\mathbf{M}_\infty = \mathbf{M}_\infty^\top$ soddisfa

$$\mathbf{M} = \mathbf{F}_c^\top \mathbf{M} \mathbf{F}_c - \mathbf{F}_c^\top \mathbf{M} \mathbf{G}_c \left(\mathbf{R} + \mathbf{G}_c^\top \mathbf{M} \mathbf{G}_c\right)^{-1} \mathbf{G}_c^\top \mathbf{M} \mathbf{F}_c + \mathbf{H}_c^\top \mathbf{H}_c$$

guadagno stimatore: $\mathbf{L} = \mathbf{F}_s \mathbf{P} \mathbf{H}_s^\top \left(\boldsymbol{\Sigma}_\mathbf{v} + \mathbf{H}_s \mathbf{P} \mathbf{H}_s^\top\right)^{-1}$ dove $\mathbf{P} = \mathbf{P}^\top$ soddisfa

$$\mathbf{P} = \mathbf{F}_s \mathbf{P} \mathbf{F}_s^\top - \mathbf{F}_s \mathbf{P} \mathbf{H}_s^\top \left(\boldsymbol{\Sigma}_\mathbf{v} + \mathbf{H}_s \mathbf{P} \mathbf{H}_s^\top\right)^{-1} \mathbf{H}_s \mathbf{P} \mathbf{F}^\top + \mathbf{G}_s \mathbf{G}_s^\top$$

dualità: sia $\mathbf{G}_{\mathbf{LQ}}\left(\mathbf{F}_c, \mathbf{G}_c, \mathbf{H}_c, \mathbf{R}\right) = \mathbf{K}_\infty^\star$ e $\mathbf{G}_{\mathbf{KF}}\left(\mathbf{F}_s, \mathbf{G}_s, \mathbf{H}_s, \boldsymbol{\Sigma}_\mathbf{v}\right) = \mathbf{L}$, si osserva che

$$\mathbf{G}_{\mathbf{LQ}}\left(\mathbf{F}_c, \mathbf{G}_c, \mathbf{H}_c, \mathbf{R}\right) = \mathbf{G}_{K\mathbf{F}}\left(\mathbf{F}_s^\top, \mathbf{H}_s^\top, \mathbf{G}_s^\top, \boldsymbol{\Sigma}_\mathbf{v}\right)$$

interpretazione come predittore

aggiornamento rispetto alle misure

$$\mathbf{P}(t|t) = \mathbf{P}(t|t-1) - \mathbf{P}(t|t-1)\mathbf{H}^\top \left(\Sigma_{\mathbf{v}} + \mathbf{H}\mathbf{P}(t|t-1)\mathbf{H}^\top\right)^{-1} \mathbf{H}\mathbf{P}(t|t-1)$$

$$\hat{\mathbf{x}}(t|t) = \hat{\mathbf{x}}(t|t-1) + \mathbf{P}(t|t-1)\mathbf{H}^\top \mathbf{\Sigma}_{\mathbf{v}}^{-1} \left(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t|t-1)\right)$$

aggiornamento temporale

$$\mathbf{P}(t+1|t) = \mathbf{F}\mathbf{P}(t|t)\mathbf{F}^\top + \mathbf{\Sigma}_{\mathbf{w}}$$

$$\hat{\mathbf{x}}(t+1|t) = \mathbf{F}\hat{\mathbf{x}}(t|t) + \mathbf{G}\mathbf{u}(t)$$

correzione

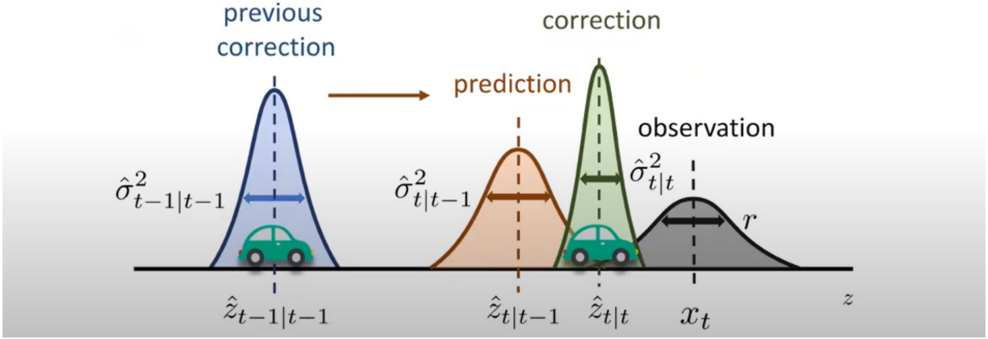
$\mathbf{e}(t) = \mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t\|t-1)$	innovazione
$\mathbf{S}(t) = \boldsymbol{\Sigma}_{\mathbf{v}} + \mathbf{H}\mathbf{P}(t\|t-1)\mathbf{H}^\top$	varianza dell'innovazione
$\mathbf{L}'(t) = \mathbf{P}(t\|t-1)\mathbf{H}^\top \left(\mathbf{S}(t)\right)^{-1}$	guadagno di correzione
$\hat{\mathbf{x}}(t\|t) = \hat{\mathbf{x}}(t\|t-1) + \mathbf{L}'(t)\mathbf{e}(t)$	correzione della stima
$\mathbf{P}(t\|t) = \mathbf{P}(t\|t-1) - \mathbf{L}'(t)\mathbf{S}(t)\mathbf{L}'(t)^\top$	correzione della varianza

predizione

$\hat{\mathbf{x}}(t+1\|t) = \mathbf{F}\hat{\mathbf{x}}(t\|t) + \mathbf{G}\mathbf{u}(t)$	predizione della stima
$\mathbf{P}(t+1\|t) = \mathbf{F}\mathbf{P}(t\|t)\mathbf{F}^\top + \boldsymbol{\Sigma}_{\mathbf{w}}$	predizione della varianza

State equation: $z_t = z_{t-1} + u\Delta t + w_t$, where $w_t \sim \mathcal{N}(0, q)$

Observation equation: $x_t = z_t + v_t$, where $v_t \sim \mathcal{N}(0, r)$



Stima secondo Kalman

- filtro di Kalman per sistemi a tempo discreto
  - derivazione delle formule
  - interpretazione come stimatore ad anello chiuso
  - proprietà del filtro di Kalman
- filtro di Kalman per sistemi a tempo continuo
- controllo ottimo LQG

ipotesi sul sistema dinamico

sistema LTI a tempo continuo

$$\dot{\mathbf{x}}(t) = \mathbf{Fx(t)} + \mathbf{Gu(t)} + \mathbf{w(t)}$$
$$\mathbf{y(t)} = \mathbf{Hx(t)}+\mathbf{v(t)}$$
$$\mathbf{x(0)} = \mathbf{x_0}$$

$\mathbf{x} \in \mathbb{R}^{\mathbf{n}}$, $\mathbf{u} \in \mathbb{R}^{\mathbf{m}}$, $\mathbf{y} \in \mathbb{R}^{\mathbf{p}}$

$\mathbf{w} \in \mathbb{R}^n$: errore di processo

$\mathbf{v} \in \mathbb{R}^p$: errore di misura

1. errore/rumore di processo
processo stocastico Gaussiano, stazionario, bianco e a media nulla

- $\mathbb{E}\left[\mathbf{w}(t)\right] = \mathbf{0} \ \forall t$
- $\mathbb{E}\left[\mathbf{w}(t)\mathbf{w}(t)^\top\right] = \mathbf{\Sigma_w} \ \forall t$
- $\mathbb{E}\left[\mathbf{w}(t_1)\mathbf{w}(t_2)^\top\right] = \mathbf{0} \ \forall t_1 \neq t_2$

$$\boxed{\mathbf{w}(t) \in \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_w})}$$

sistema LTI a tempo continuo

$$\dot{\mathbf{x}}(t) = \mathbf{Fx}(\mathbf{t}) + \mathbf{Gu}(\mathbf{t}) + \mathbf{w}(\mathbf{t})$$
$$\mathbf{y}(\mathbf{t}) = \mathbf{Hx}(\mathbf{t}) + \mathbf{v}(\mathbf{t})$$
$$\mathbf{x}(\mathbf{0}) = \mathbf{x_0}$$

$\mathbf{x} \in \mathbb{R}^{\mathbf{n}}$, $\mathbf{u} \in \mathbb{R}^{\mathbf{m}}$, $\mathbf{y} \in \mathbb{R}^{\mathbf{p}}$

$\mathbf{w} \in \mathbb{R}^{n}$: errore di processo

$\mathbf{v} \in \mathbb{R}^{p}$: errore di misura

2. errore/rumore di misura

processo stocastico Gaussiano, stazionario, bianco e a media nulla

- $\mathbb{E}\left[\mathbf{v}(t)\right] = \mathbf{0}\ \forall t$
- $\mathbb{E}\left[\mathbf{v}(t)\mathbf{v}(t)^\top\right] = \mathbf{\Sigma_v}\ \forall t$
- $\mathbb{E}\left[\mathbf{v}(t_1)\mathbf{v}(t_2)^\top\right] = \mathbf{0}\ \forall t_1 \neq t_2$

$$\mathbf{v}(t) \in \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_v})$$

sistema LTI a tempo continuo

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(\mathbf{t}) + \mathbf{G}\mathbf{u}(\mathbf{t}) + \mathbf{w}(\mathbf{t})$$
$$\mathbf{y}(\mathbf{t}) = \mathbf{H}\mathbf{x}(\mathbf{t}) + \mathbf{v}(\mathbf{t})$$
$$\mathbf{x}(\mathbf{0}) = \mathbf{x_0}$$

$\mathbf{x} \in \mathbb{R}^{\mathbf{n}}, \ \mathbf{u} \in \mathbb{R}^{\mathbf{m}}, \ \mathbf{y} \in \mathbb{R}^{\mathbf{p}}$

$\mathbf{w} \in \mathbb{R}^n$: errore di processo

$\mathbf{v} \in \mathbb{R}^p$: errore di misura

3. errore/rumore di processo e di misura
processi stocastici Gaussiani, stazionari, bianco e a media nulla & *indipendenti*

$$\mathbb{E}\left[\mathbf{w}(t_1)\mathbf{v}(t_2)^\top\right] = \mathbf{0} \ \forall t_1 \neq t_2$$

4. stato iniziale
variabile aleatoria Gaussiana di media e varianza note

$$\mathbf{x}(\mathbf{0}) \in \mathcal{N}(\mathbf{x_0}, \mathbf{\Sigma_0})$$

filtro di Kalman-Bucy

$$\mathbf{L}(t) = \mathbf{P}(\mathbf{t})\mathbf{H}^\top \boldsymbol{\Sigma}_\mathbf{v}^{-1}$$
$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}\hat{\mathbf{x}}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}(t)\left(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t)\right)$$
$$\dot{\mathbf{P}}(t) = \mathbf{F}\mathbf{P}(\mathbf{t}) + \mathbf{P}(\mathbf{t})\mathbf{F}^\top - \mathbf{P}(\mathbf{t})\mathbf{H}^\top \boldsymbol{\Sigma}_\mathbf{v}^{-1}\mathbf{H}\mathbf{P}(\mathbf{t}) + \boldsymbol{\Sigma}_\mathbf{w}$$

- il filtro di Kalman a tempo continuo è uno stimatore ad anello chiuso (alla Luengerger)
- nel caso a tempo continuo non esiste distinzione tra stima predittiva $(\hat{\mathbf{x}}(t+1|t))$ e stima filtrata $(\hat{\mathbf{x}}(t|t))$: ad ogni istante temporale t viene propagata un'unica stima $\hat{\mathbf{x}}(t)$ basata sulle osservazioni $\mathbf{y}(\tau)$, $\tau < t$ fino all'istante t

proprietà del filtro di Kalman-Bucy

realizzabilità

la necessità di acquisire ed elaborare misure con continuità rende di fatto **impossibile** l'implementazione e l'utilizzo del filtro di Kalman-Bucy in generale

eccezioni

- il sistema è tempo-invariante, i disturbi sono stazionari e il guadagno $\mathbf{L}(\mathbf{t}) \equiv \mathbf{L}$ viene mantenuto costante: il filtro di Kalman-Bucy si riduce a un filtro LTI analogico realizzabile con componenti elettronici lineari
- le osservazioni vengono acquisite ad istanti discreti: il filtro di Kalman-Bucy è implementato nella sua versione ibrida

guadagno costante

allora $\mathbf{P(t)} = \mathbf{P}$: no aggiornamento della varianza

$\mathbf{P} = \mathbf{P}^\top$ sdp: unica soluzione della EAR

$$\mathbf{0} = \mathbf{FP} + \mathbf{PF}^\top - \mathbf{PH}^\top \boldsymbol{\Sigma}_\mathbf{v}^{-1} \mathbf{HP} + \boldsymbol{\Sigma}_\mathbf{w}$$

se $(\mathbf{F}, \mathbf{G})$ stabilizzabile, $\mathbf{GG}^\top = \boldsymbol{\Sigma}_\mathbf{w}$ e $(\mathbf{F}, \mathbf{H})$ rivelabile
allora il filtro di Kalman è asintoticamente stabile

esempio: caso continuo VS. caso discreto

sistema del primo ordine *a tempo continuo*

$$\begin{aligned} \dot{x}(t) &= fx(t) + w(t) \qquad & w(t) &\sim \mathcal{N}(0, q = \sigma_w^2 \geq 0) \\ y(t) &= x(t) + v(t) & v(t) &\sim \mathcal{N}(0, r = \sigma_v^2 > 0) \end{aligned}$$

sistema osservabile &, se $q > 0$, raggiungibile
$\rightarrow$ sistema rivelabile &, se $q > 0$ oppure se $q = 0$ e $f < 0$, stabilizzabile

EAR: $\quad 2fp - \frac{1}{r}p^2 + q = 0 \quad \Rightarrow \quad \begin{cases} p_{1,2} = fr \pm \sqrt{f^2r^2 + qr} & \text{se } q > 0 \\ p_1 = 0, p_2 = 2fr & \text{se } q = 0 \end{cases}$

se $q > 0$ (sistema osservabile e raggiungibile) allora $p = p_1 = fr + \sqrt{f^2 r^2 + qr} > 0$

$$L = \frac{1}{r} p = f + \sqrt{f^2 + q/r}$$

$$\dot{\hat{x}} = f \hat{x}(t) + L(y(t) - \hat{x}(t)) = -\sqrt{f^2 + q/r}\, \hat{x}(t) + \left(f + \sqrt{f^2 + q/r} \right) y(t)$$

$$\dot{\tilde{x}}(t) = -\sqrt{f^2 + q/r}\, \tilde{x}(t) + w(t) - \left(f + \sqrt{f^2 + q/r} \right) v(t) \quad \text{esp. stabile}$$

se $q = 0$ e $f < 0$ (sistema stabilizzabile) allora $p = p_1 = 0$

$$\boxed{\text{stimatore in catena aperta:} \quad \dot{\hat{x}} = f\hat{x}(t) + L(y(t) - \hat{x}(t)) = f\hat{x}(t)}$$

sistema del primo ordine *a tempo discreto*

$$
\begin{aligned}
&x(t+1) = fx(t) + w(t) \qquad && w(t) \sim \mathcal{N}(0, q = \sigma_w^2 \geq 0) \\
&y(t) = x(t) + v(t) && v(t) \sim \mathcal{N}(0, r = \sigma_v^2 \geq 0)
\end{aligned}
$$

sistema osservabile &, se $q > 0$, raggiungibile
$\rightarrow$ sistema rivelabile &, se $q > 0$ oppure se $q = 0$ e $|f| < 1$, stabilizzabile

EARD: $\quad p = f^2 p - \frac{f^2 p^2}{r+p} + q$

$$
\Rightarrow \quad \begin{cases} p_{1,2} = \frac{1}{2}\left((f^2-1)r + q \pm \sqrt{(r - f^2 r - q) + 4qr}\right) & \text{se } q > 0 \\ p_1 = 0, p_2 = (f^2-1)r & \text{se } q = 0 \end{cases}
$$

se $q > 0$ (sist. oss e ragg.) allora $p = p_1 = \frac{1}{2}\left((f^2-1)r + q + \sqrt{(r - f^2 r - q) + 4qr}\right) > 0$

$$\boxed{\begin{aligned} \hat{x}(t+1|t) &= \frac{fr}{r+p}\hat{x}(t|t-1) + \frac{fp}{r+p}y(t) \\ \hat{x}(t|t) &= \frac{r}{r+p}\hat{x}(t|t-1) + \frac{p}{r+p}y(t) \end{aligned}}$$

$\tilde{x}(t)$ esponenzialmente stabile

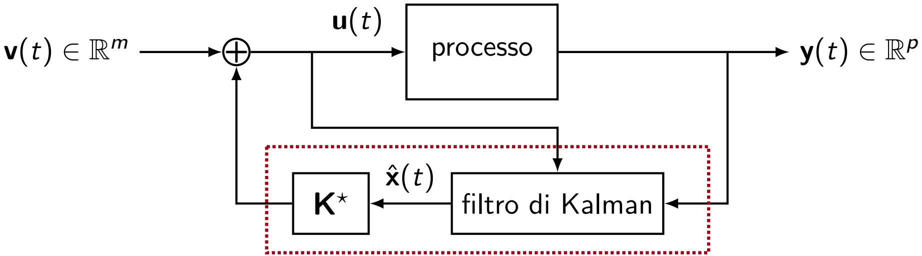
se $q = 0$ e $|f| < 1$ (sistema stabilizzabile) allora $p = p_1 = 0$

$$\boxed{\text{stimatore in catena aperta:} \quad \hat{x}(t+1|t) = f\hat{x}(t|t), \quad \hat{x}(t|t) = \hat{x}(t|t-1)}$$

Stima secondo Kalman

- filtro di Kalman per sistemi a tempo discreto
  - derivazione delle formule
  - interpretazione come stimatore ad anello chiuso
  - proprietà del filtro di Kalman
- filtro di Kalman per sistemi a tempo continuo
- controllo ottimo LQG

LQG = LQR + KF


$\mathbf{v}(t) \in \mathbb{R}^m$
$\mathbf{u}(t)$
processo
$\mathbf{y}(t) \in \mathbb{R}^p$
$\hat{\mathbf{x}}(t)$
$\mathbf{K}^\star$
filtro di Kalman

caso a tempo discreto

$$\mathbf{x}(t+1) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t)$$
$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t)$$
$$\mathbf{x}(0) = \mathbf{x}_0$$

$\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^p$

$\mathbf{w} \in \mathbb{R}^n$: errore di processo

$\mathbf{v} \in \mathbb{R}^p$: errore di misura

- ipotesi di gaussianità: $\mathbf{x}(0) \sim \mathcal{N}(\mathbf{x}_0, \mathbf{\Sigma}_0)$, $\mathbf{w}(t) \sim \mathcal{N}(0, \mathbf{\Sigma}_\mathbf{w})$, $\mathbf{v}(t) \sim \mathcal{N}(0, \mathbf{\Sigma}_\mathbf{v})$
- ipotesi di incorrelazione: $\mathbf{w}(t)$, $\mathbf{v}(t)$ sono mutuamente incorrelati e incorrelati con $\mathbf{x}(0)$

$$J_T(t) = \mathbb{E}\left[\sum_{t=0}^{T-1}\left(\mathbf{x}(t)^\top \mathbf{Q}\mathbf{x}(t) + \mathbf{u}(t)^\top \mathbf{R}\mathbf{u}(t)\right) + \mathbf{x}(T)^\top \mathbf{S}\mathbf{x}(T)\right]$$

$$\mathbf{S}, \mathbf{Q} \in \mathbb{R}^{n\times n} \text{ sdp}, \ \mathbf{R} \in \mathbb{R}^{m\times m} \text{ dp}$$

controllo LQG

$$\mathbf{u}_T^\star(t) = \arg\min_{t\in[0,T-1]} J_T(\mathbf{u}(t))$$

$$\mathbf{u}_T^\star(t) = -\mathbf{K}_T^\star(t)\hat{\mathbf{x}}(t|t-1)$$

$$\mathbf{K}_T^\star(t) = \left(\mathbf{R} + \mathbf{G}^\top \mathbf{M}_T(t+1)\mathbf{G}\right)^{-1} \mathbf{G}^\top \mathbf{M}_T(t+1)\mathbf{F}$$

$$\hat{\mathbf{x}}(t+1|t) = \mathbf{F}\hat{\mathbf{x}}(t|t-1) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}(t)\left(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t|t-1)\right)$$

$\mathbf{M}_T(t)$ soluzione della ERD
$\mathbf{L}(t)$ guadagno del filtro di Kalman

principio di separazione

caso a tempo continuo

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{w}(t)$$
$$\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t)$$
$$\mathbf{x}(0) = \mathbf{x}_0$$

$\mathbf{x} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^p$

$\mathbf{w} \in \mathbb{R}^n$: errore di processo

$\mathbf{v} \in \mathbb{R}^p$: errore di misura

- ipotesi di gaussianità: $\mathbf{x}(0) \sim \mathcal{N}(\mathbf{x}_0, \mathbf{\Sigma}_0)$, $\mathbf{w}(t) \sim \mathcal{N}(0, \mathbf{\Sigma}_\mathbf{w})$, $\mathbf{v}(t) \sim \mathcal{N}(0, \mathbf{\Sigma}_\mathbf{v})$
- ipotesi di incorrelazione: $\mathbf{w}(t)$, $\mathbf{v}(t)$ sono mutuamente incorrelati e incorrelati con $\mathbf{x}(0)$

$$J_T(t) = \mathbb{E}\left[\int_{t=0}^{T} \left(\mathbf{x}(t)^\top \mathbf{Q}\mathbf{x}(t) + \mathbf{u}(t)^\top \mathbf{R}\mathbf{u}(t)\right) + \mathbf{x}(T)^\top \mathbf{S}\mathbf{x}(T)\right]$$

$$\mathbf{S}, \mathbf{Q} \in \mathbb{R}^{n \times n} \text{ sdp}, \ \mathbf{R} \in \mathbb{R}^{m \times m} \text{ dp}$$

controllo LQG

$$\mathbf{u}_T^\star(t) = \arg\min_{t\in[0,T]} J_T(\mathbf{u}(t))$$

$$\mathbf{u}_T^\star(t) = -\mathbf{K}_T^\star(t)\hat{\mathbf{x}}(t)$$

$\mathbf{K}_T^\star(t) = \mathbf{R}^{-1}\mathbf{G}^\top\mathbf{M}_T(t)$
$\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}\hat{\mathbf{x}}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}(t)\left(\mathbf{y}(t) - \mathbf{H}\hat{\mathbf{x}}(t)\right)$

$\mathbf{M}_T(t)$ soluzione della EDR
$\mathbf{L}(t)$ guadagno del filtro di Kalman

principio di separazione

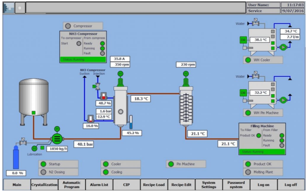
cenni di
Controllo Adattativo

- principali aspetti del controllo adattativo
- ruolo dell'identificazione
- tipici schemi di controllo adattativo
- controllo adattativo di un quadrotor

introduzione al controllo adattativo

$P(s)$ è un'*approssimazione* del processo reale

$\Longrightarrow$ incertezza dei parametri (non lineari)
$\Longrightarrow$ dinamiche inaspettate



controllo robusto o H_∞

regolatore *unico* che garantisce prestazioni adeguate
per *tutti* i parametri in un certo intervallo

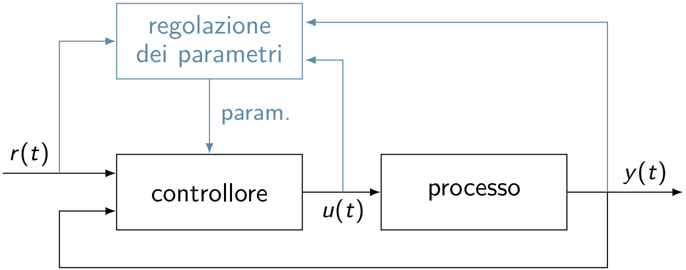
△ affidabile & definitivo
▽ spesso conservativo (worst-case)

identificazione & controllo adattativo

regolatore basato su una stima (offline/online)
dei parametri (non noti e tempo-invarianti/tempo varianti)

△ tenente conto delle prestazione limite del sistema
▽ complesso, fragile, dipendente dell'identificabilità del sistema

controllo adattativo (o adattivo): metodo di controllo utilizzato da un controllore che può *modificare il suo comportamento* in risposta a cambiamenti nella dinamica del processo e nel carattere dei disturbi al fine di ottimizzare le prestazioni



un controllore adattivo è un controllore con parametri regolabili e un meccanismo per la regolazione dei parametri

storia del controllo adattativo

anni '50

necessità di controllare *super aerei*
caratterizzati da ampi range lavorativi
in termini di velocità e altezza

anni '60

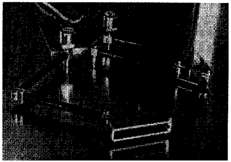
sviluppo della teoria del controllo
(rappresentazione in spazio di stato, teoria della stabilità, teoria del controllo stocastico, identificazione di sistemi, dynamic programming)

anni '70-80

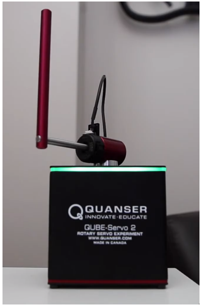
rinascita del controllo adattativo
(stabilità dei sistemi adattativi ma sotto ipotesi molto restrittive,
idea di combinare l'identificazione dei sistemi e il controllo robusto)

anni '80-90

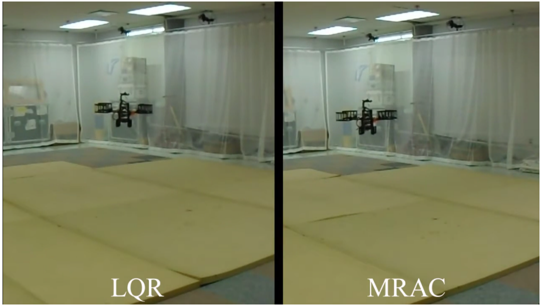
epoca d'oro del controllo adattativo
(esperimenti in ambito accademico e industriale)



vantaggi del controllo adattativo



https://www.youtube.com/watch?v=ypCf_VODmik



https://www.youtube.com/watch?v=BIpnjvQ1uhI

cenni di
Controllo Adattativo

- principali aspetti del controllo adattativo
- ruolo dell'identificazione
- tipici schemi di controllo adattativo
- controllo adattativo di un quadrotor

identificazione

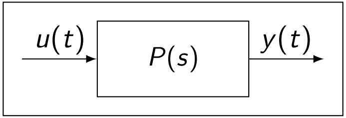
identificazione (di sistemi): metodologia per *costruire modelli matematici di sistemi dinamici* utilizzando misurazioni dei segnali di ingresso e di uscita del sistema

- sistema lineare/non lineare
- approccio parametrico/non parametrico
  - i metodi non parametrici tentano di stimare un modello generico (risposte al gradino, all'impulso, in frequenza)
  - i metodi parametrici stimano i parametri in un modello specificato (funzioni di trasferimento, matrici spazio degli stati)

in seguito, identificazione (parametrica) della funzione di trasferimento

▶ tipi di identificazione

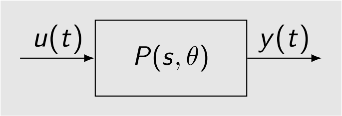
white box



modello di cui si conosce
la struttura e
i parametri

modello ricavato in base
ai principi fisici e
ai dati di progetto

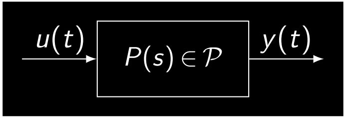
gray box



modello di cui si conosce
la struttura ma
non tutti i parametri

modello ricavato in base
ai principi fisici e
ai dati sperimentali

black box



modello di cui
NON si conosce
la struttura e i parametri

modello ricavato in base
ai dati sperimentali

- **identificazione gray box**

gray box

$$u(t) \longrightarrow \boxed{P(s,\theta)} \longrightarrow y(t)$$

modello di cui si conosce
la struttura ma
non tutti i parametri

modello ricavato in base
ai principi fisici e
ai dati sperimentali

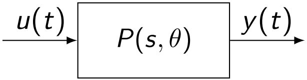
esperimenti

- implementare degli specifici ingressi $u(t)$, $t \in [0, T]$
- registrare le corrispondenti uscite $y(t)$, $t \in [0, T]$
- stimare i parametri ignoti θ sfruttando $(u(t), y(t))$

come scegliere gli ingressi $u(t)$, $t \in [0, T]$?
qual è la bontà della stima $\hat{\theta}$ che si può ottenere?
è possibile determinare l'incertezza della stima $\hat{\theta}$?

▶ identificazione diretta e indiretta

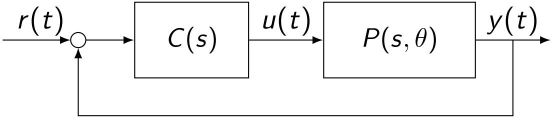
diretta



la stima di θ si ricava sfruttando
i dati $(u(t), y(t))$, $t \in [0, T]$

la retroazione è ignorata

indiretta



la stima di θ si ricava sfruttando
i dati $(r(t), y(t))$, $t \in [0, T]$

il controllore è noto

▶ **identificazione offline e online**

offline

stima statica

for $t \in [0, T]$
  si raccolgono i dati $(u(t), y(t))$
end
si stimano i parametri (costanti) $\hat{\theta}$

online

stima dinamica

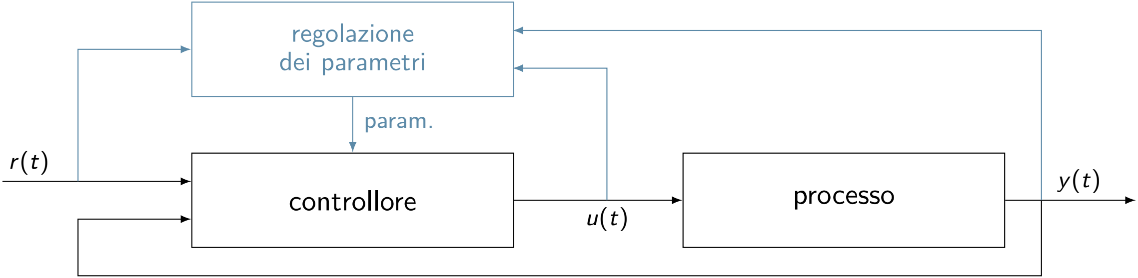
for $t \in [0, T]$
  si raccolgono i dati $(u(t), y(t))$
  si stimano i parametri $\hat{\theta}(t)$
end

cenni di

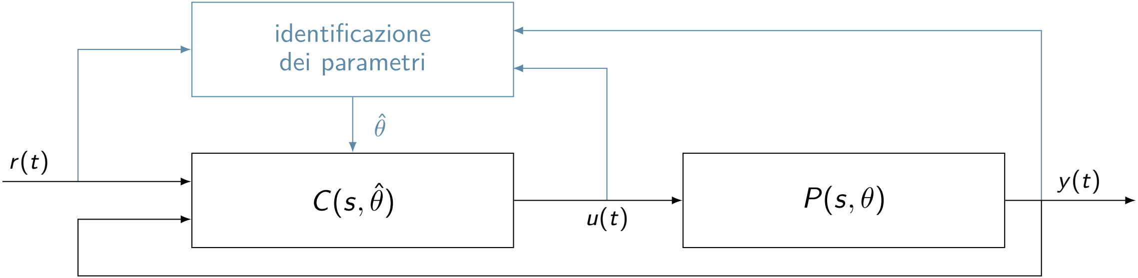
Controllo Adattativo

- principali aspetti del controllo adattativo
- ruolo dell'identificazione
- tipici schemi di controllo adattativo
- controllo adattativo di un quadrotor

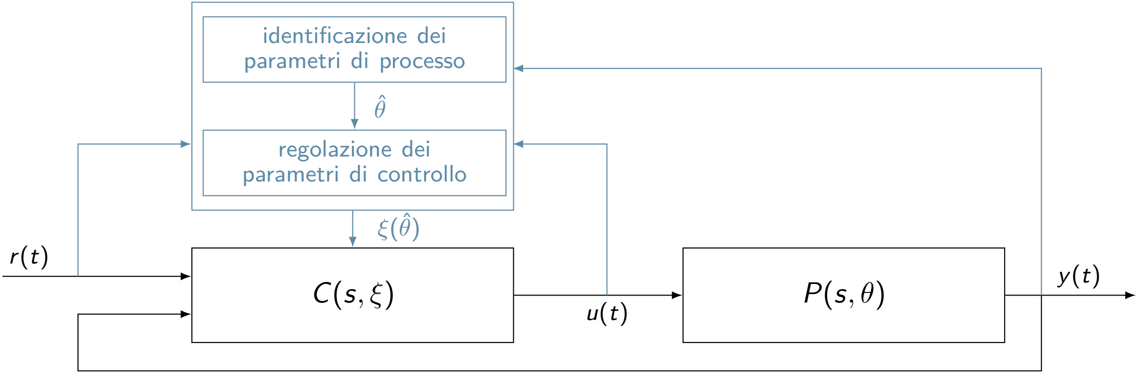
identificazione e controllo adattativo



regolazione diretta



regolazione indiretta



regolazione

- **feedback**: l'azione di controllo viene definita calcolando l'errore (differenza) tra riferimento $r(t)$ ed uscita $y(t)$ che filtrato attraverso un blocco adattativo restituisce l'azione di controllo

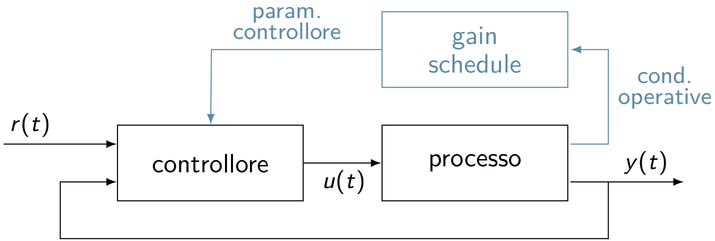
maggiore prontezza in caso di disturbi esterni (ma in assenza di ritardi)

- **feedforward**: l'azione di controllo viene definita sfruttando un'opportuna modellazione del sistema (ad esempio in forma tabellata) che viene modificata secondo un'opportuna legge adattativa in base alle misure effettuate in uscita

da preferire in caso di forti ritardi nel segnale d'uscita

gain scheduling

regolazione dei parametri del controllore in base alle possibili condizioni operative del processo identificabili in base all'andamento di alcune variabili misurabili

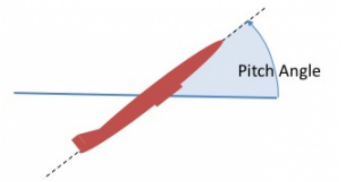


controllo in feedback in cui il guadagno dipende dalle **condizioni operative** del processo determinate osservando specifiche variabili di pianificazione (scheduling variables) la cui scelta è dettata dalla *fisica (ben nota) del sistema*

$\Longrightarrow$ identificazione *online* delle condizioni operative

regolazione *"offline"* dei parametri del controllore

esempio di applicazione



controllo dell'angolo di pitch in un aereo

https://www.youtube.com/watch?v=YiUjAV1bhKs&t=12s

$P(s, \theta) = \frac{\text{elevation angle}}{\text{pitch command}}$ sistema lineare a parametri variabili

parametri tempo varianti e dipendenti dalle condizioni operative

variabili di pianificazione: Mach number e/o pressione

$C(s, \{K_P, K_I, K_D\})$ controllore PID con guadagni variabili

implementazione

0. determinazione delle scheduling variables
1. linearizzazione del processo in corrispondenza ad ogni condizione operativa
2. calcolo dei parametri di controllo in corrispondenza ad ogni condizione operativa
3. selezione della gain scheduling architecture (gestione delle transizioni)
4. test delle performance del controllore per via simulativa

⚠ no feedback tra le performance del sistema in catena chiusa e i parametri del controllore (compensazione open loop)

stabilità del controllore $\propto$ transizioni tra le diverse condizioni operative

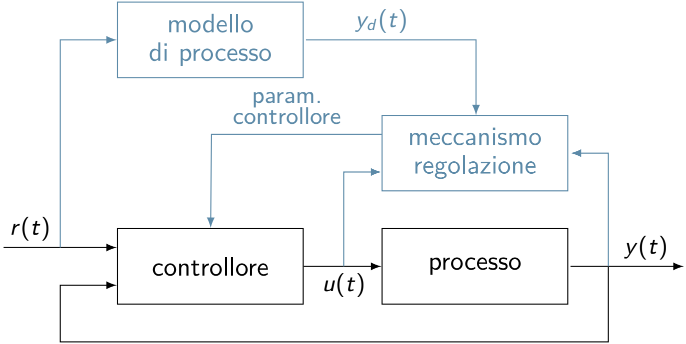
pros & cons

- △ utilizzo di tecniche di controllo lineare anche per sistemi non lineari
- △ parametri di controllo fissati a priori (no fenomeni di deriva o divergenza)

- ▽ processo di sintesi del controllore laboriosa
- ▽ non-ottimalità rispetto alle specifiche di controllo

Model Reference Adaptive Control

regolazione dei parametri del controllore in modo da minimizzare l'errore tra l'uscita del processo vera e desiderata



schema di controllo a **due loops** in cui non c'è una vera e propria identificazione

$\Longrightarrow$ *loop interno*: controllo in feedback standard

loop esterno: regolazione dei parametri del controllore

meccanismo di regolazione

- input: $u(t), y(t), y_d(t)$ ingresso e uscita reali del processo, uscita desiderata
  output: $\xi(t)$ parametri del controllore

- strategie $e(t) = y(t) - y_d(t)$

  - metodo del gradiente

    MIT rule $\frac{d\xi(t)}{dt} = -\gamma e(t)\frac{de(t)}{d\xi(t)}$

    minimizzazione di $J(\xi(t)) = \frac{1}{2}e(t)^2$

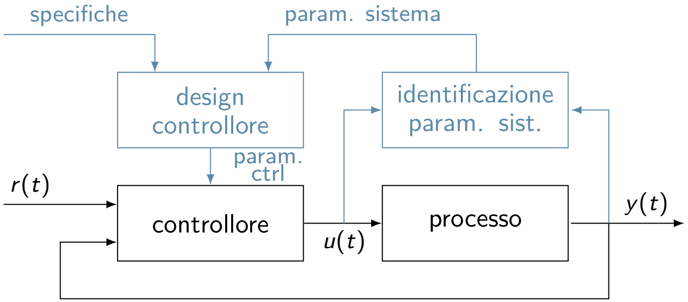
    γ: adaptive gain $\frac{de(t)}{d\xi(t)}$: sensitivity derivative

  - teoria della stabilità

    funzione di Lyapunov tale che $\lim_{t\to\infty} e(t) = 0$

Self Tuning Regulator

regolazione dei parametri del controllore in base al risultato dell'identificazione dei parametri di processo e delle specifiche



schema di controllo a **due loops** in cui l'identificazione dei parametri di sistema gioca un ruolo fondamentale

$\Longrightarrow$ *loop interno*: controllo in feedback standard

loop esterno: stima dei parametri del sistema + progettazione del controllore

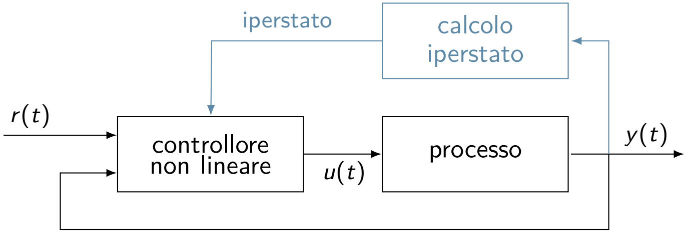
principio di funzionamento

- assunzioni
  - i parametri del sistema sono non noti ma *costanti o lentamente tempo-varianti*
  - *certainty equivalence principle*: i parametri di sistema stimati sono considerati corretti nella progettazione del controllore
- **explicit/indirect** STR
  - identificazione *online* dei parametri di sistema (stima parametrica ricorsiva)
  - progettazione *online* del controllore sulla base delle specifiche e dei risultati dell'identificazione

  implicit/direct STR
  - determinazione dei parametri del controllore in seguito alla riparametrizzazione del processo (unico blocco nel loop esterno)

dual control

sintesi del controllore in base al risultato della stima dell'iperstato (stato e parametri del processo)



schema di controllo non lineare basato sulla stima stocastico dello stato aumentato (iperstato) del processo

$\Longrightarrow$ identificazione *online* e *contemporanea* di stato e parametri di sistema

principio di funzionamento

- assunzione
  i parametri del sistema variano con una dinamica comparabile a quella del processo

- sintesi del controllore
  tecniche di ottimizzazione: minimizzazione di una funzione di costo che coivolge l'aspettazione di una funzione dell'iperstato stimato (aspetto stocastico)

teoria del controllo duale: branca della teoria del controllo che si occupa del controllo di sistemi le cui caratteristiche sono inizialmente sconosciute con un duplice scopo

1) action - controllare il sistema al meglio in base alla sua conoscenza attuale
2) investigation - sperimentare il sistema in modo da apprenderne il comportamento

cenni di
Controllo Adattativo

- principali aspetti del controllo adattativo
- ruolo dell'identificazione
- tipici schemi di controllo adattativo
- controllo adattativo di un quadrotor

modello del quadrotor

modello non lineare

$$\mathbf{x} = \begin{bmatrix} \mathbf{p} & \boldsymbol{\delta} & \mathbf{v} & \boldsymbol{\omega} \end{bmatrix}^\top \in \mathbb{R}^{12}$$

$$= \begin{bmatrix} x & y & z & \phi & \theta & \psi & \dot{x} & \dot{y} & \dot{z} & \omega_x & \omega_y & \omega_z \end{bmatrix}^\top$$

$$\mathbf{u} = \begin{bmatrix} \|\mathbf{f}_c\| & \boldsymbol{\tau}_c \end{bmatrix}^\top \in \mathbb{R}^4$$

$$= \begin{bmatrix} F & \tau_1 & \tau_2 & \tau_3 \end{bmatrix}^\top$$

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\boldsymbol{\delta}} \\ \dot{\mathbf{v}} \\ \dot{\boldsymbol{\omega}} \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \mathbf{T}\boldsymbol{\omega} \\ -g\mathbf{e}_3 + m^{-1}\mathbf{R}(\boldsymbol{\delta})\mathbf{f}_c \\ J^{-1}(-\boldsymbol{\omega} \times J\boldsymbol{\omega} + \boldsymbol{\tau}_c) \end{bmatrix} = \begin{bmatrix} \mathbf{v} \\ \mathbf{T}\boldsymbol{\omega} \\ -g\mathbf{e}_3 \\ J^{-1}(-\boldsymbol{\omega} \times J\boldsymbol{\omega}) \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{3\times 1} \\ \mathbf{0}_{3\times 1} \\ m^{-1}\mathbf{R}(\boldsymbol{\delta})\mathbf{f}_c \\ J^{-1}\boldsymbol{\tau}_c \end{bmatrix} = f_1(\mathbf{x}) + f_2(\mathbf{x}, \mathbf{u})$$

$$
f_1(\mathbf{x}) = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \omega_x + (\sin\phi\tan\theta)\omega_y + (\cos\phi\tan\theta)\omega_z \\ \cos\phi\,\omega_y - \sin\phi\,\omega_z \\ (\sin\phi\tan\theta)\omega_y + \frac{\cos\phi}{\cos\theta}\omega_z \\ 0 \\ 0 \\ -g \\ \frac{J_y - J_z}{J_x}\omega_y\omega_z \\ \frac{J_x - J_z}{J_y}\omega_x\omega_z \\ \frac{J_x - J_y}{J_z}\omega_x\omega_y \end{bmatrix}
\qquad
f_2(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{m}(\sin\psi\sin\phi + \cos\psi\sin\theta\cos\phi)\,F \\ \frac{1}{m}(-\cos\psi\sin\phi + \sin\psi\sin\theta\cos\phi)\,F \\ \frac{1}{m}(\cos\theta\cos\phi)\,F \\ \frac{1}{J_x}\tau_1 \\ \frac{1}{J_y}\tau_2 \\ \frac{1}{J_z}\tau_3 \end{bmatrix}
$$

modello linearizzato

$$\mathbf{x} = \begin{bmatrix}\mathbf{p} & \boldsymbol{\delta} & \mathbf{v} & \boldsymbol{\omega}\end{bmatrix}^\top = \begin{bmatrix}x & y & z & \phi & \theta & \psi & \dot{x} & \dot{y} & \dot{z} & \omega_x & \omega_y & \omega_z\end{bmatrix}^\top$$

$$\mathbf{u} = \begin{bmatrix}\|\mathbf{f}_c\| & \boldsymbol{\tau}_c\end{bmatrix}^\top = \begin{bmatrix}F & \tau_1 & \tau_2 & \tau_3\end{bmatrix}^\top$$

$$\dot{\mathbf{x}} = \mathbf{F}\mathbf{x} + \mathbf{G}\mathbf{u}$$

$$\mathbf{F} = \begin{bmatrix}\mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{F}_{32} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3}\end{bmatrix}, \quad \mathbf{F}_{32} = \begin{bmatrix}0 & g & 0 \\ -g & 0 & 0 \\ 0 & 0 & 0\end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix}\mathbf{0}_{3\times4} \\ \mathbf{0}_{3\times4} \\ \mathbf{G}_3 \\ \mathbf{G}_4\end{bmatrix}, \quad \mathbf{G}_3 = \begin{bmatrix}0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{m} & 0 & 0 & 0\end{bmatrix} \quad \mathbf{G}_4 = \begin{bmatrix}0 & \frac{1}{J_x} & 0 & 0 \\ 0 & 0 & \frac{1}{J_y} & 0 \\ 0 & 0 & 0 & \frac{1}{J_z}\end{bmatrix}$$

modello linearizzato di roll, pitch e yaw

$$\ddot{\phi} \approx \dot{\omega}_x = \frac{1}{J_x}\tau_1 \qquad \rightarrow \qquad P_\phi(s) = \frac{\phi(s)}{\tau_1(s)} = \frac{1}{J_x s^2}$$

$$\ddot{\theta} \approx \dot{\omega}_y = \frac{1}{J_2}\tau_2 \qquad \rightarrow \qquad P_\theta(s) = \frac{\theta(s)}{\tau_2(s)} = \frac{1}{J_y s^2}$$

$$\ddot{\psi} \approx \dot{\omega}_z = \frac{1}{J_z}\tau_3 \qquad \rightarrow \qquad P_\psi(s) = \frac{\psi(s)}{\tau_3(s)} = \frac{1}{J_z s^2}$$

doppio integratore

modello linearizzato di roll

$$\begin{bmatrix}\dot{\phi}(t)\\ \ddot{\phi}(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\phi(t)\\ \dot{\phi}(t)\end{bmatrix} + \begin{bmatrix}0\\ \frac{1}{J_x}\end{bmatrix}\tau_1(t)$$

$$\phi(t) = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}\phi(t)\\ \dot{\phi}(t)\end{bmatrix}$$

$$\iff$$

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t)$$

$$y(t) = \mathbf{c}^\top \mathbf{x}(t)$$

$$\begin{bmatrix}\phi(t+1)\\ \dot{\phi}(t+1)\end{bmatrix} = \begin{bmatrix}1 & T\\ 0 & 1\end{bmatrix}\begin{bmatrix}\phi(t)\\ \dot{\phi}(t)\end{bmatrix} + \begin{bmatrix}\frac{T^2}{2J_x}\\ \frac{T}{J_x}\end{bmatrix}\tau_1(t)$$

$$\phi(t) = \begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}\phi(t)\\ \dot{\phi}(t)\end{bmatrix}$$

$$\iff$$

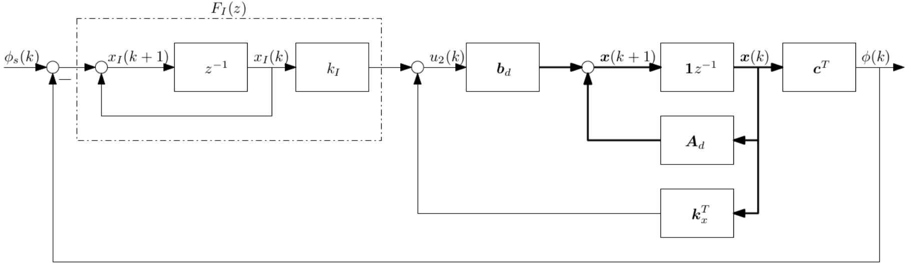
$$\mathbf{x}(t+1) = \mathbf{A}_d\mathbf{x}(t) + \mathbf{b}_d u(t)$$

$$y(t) = \mathbf{c}_d^\top \mathbf{x}(t)$$

controllo roll & pitch: non adattativo

azione integrale $F_I(z) = \frac{k_I}{z-1}$

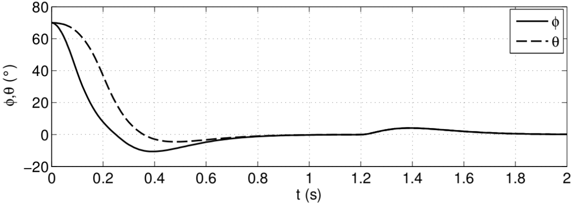
compensazione di disturbi costanti

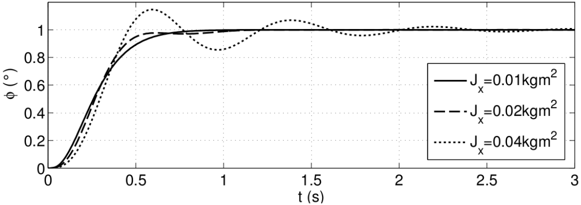


k_I e $\mathbf{k}_x^\top$ scelte in modo da garantire che

il sistema in catena chiusa abbia i poli in $z_i = 0.9$ ($T = 0.01$s)

(formula di ackerman)





- stabilizzazione in 0.8s senza overshoot
- regressione di disturbo di coppia (0.25Nm) agente a $t = 1.2$s

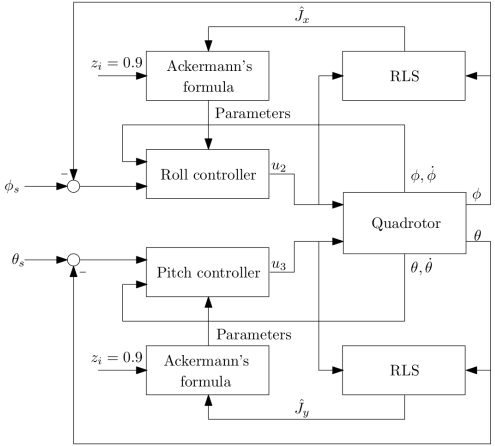
- degradazione delle performance in caso di variazioni parametriche

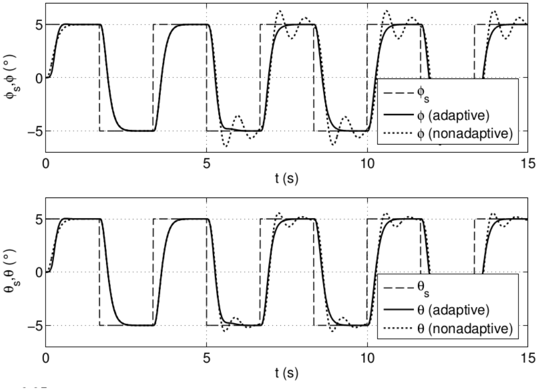
controllo roll & pitch: MIAC

Model Identification Adaptive Control ($\sim$ STR)

calcolo adattativo di k_I *e* $\mathbf{k}_x^\top$

- **stima RLS** (con forgetting factor esponenziale) dei parametri di inerzia
- **formula di Ackermann** per la sintesi del controllore (poli del sistema in catena chiusa in $z_i = 0.9$ ($T = 0.01$s))





a $t = 0$s
$J_x = J_y = 0.01$kg m^2 (valore nominale)
$\rightarrow$ controllo adattativo e non
performano ugale

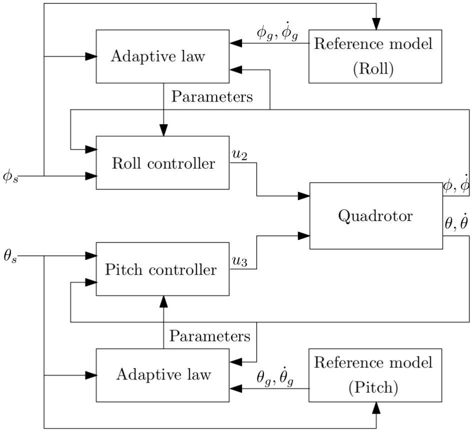
a $t = 5$s
$J_x = 0.04$kg m^2 e $J_y = 0.03$kg m^2
$\rightarrow$ controllo adattativo performa meglio

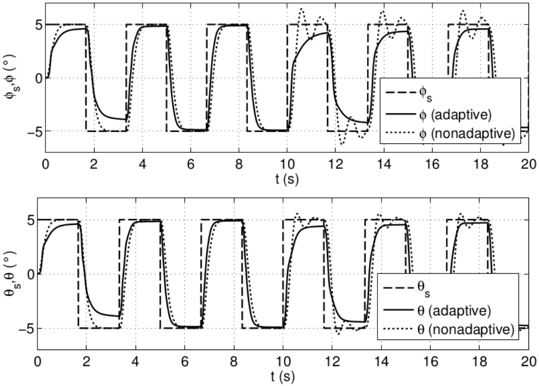
controllo roll & pitch: MRAC

Model Reference Adaptive Control

calcolo adattativo di k_I *e* $\mathbf{k}_x^\top$

- **modello di riferimento** del sistema in catena chiusa
- **Lyapunov-based adaptive law** per la sintesi del controllore (poli del sistema in catena chiusa in $z_i = 0.9$ ($T = 0.01$s))





a $t = 0$s

$J_x = J_y = 0.01$kg m^2 (valore nominale)

$\rightarrow$ controllo non adatt. performa meglio

ottimizzato sul modello nominale

a $t = 10$s

$J_x = 0.04$kg m^2 e $J_y = 0.03$kg m^2

$\rightarrow$ controllo adattativo performa meglio

Teoria dei Sistemi e Controllo ottimo (TSC)

Docente: Giulia Michieletto

Stima dello Stato & Controllo Adattativo

Corso di Laurea Magistrale in Ingegneria Meccatronica

A.A. 2023-2024

✉ giulia.michieletto@unipd.it