

TEORIA DEI SISTEMI E DEL CONTROLLO

LM in Ingegneria Informatica e Ingegneria Elettronica

<http://www.dii.unimore.it/~lbiagiotti/TeoriaSistemiControllo.html>

Stima dello stato in presenza di disturbi: il filtro di Kalman

Ing. Luigi Biagiotti

e-mail: luigi.biagiotti@unimore.it

<http://www.dii.unimore.it/~lbiagiotti>

Tecnologia e incertezza

- Nella pratica, l'uscita di un sistema è una misura, ovvero l'uscita di uno o più sensori, tipicamente affetta da rumore dovuto a vari fattori come disturbi, limiti costruttivi e quantizzazione dell'informazione
- I sistemi di attuazione possono essere inaccurati e questa inaccuratezza può essere descritta come un rumore sull'ingresso
- I modelli sono per definizione inaccurati e la loro imperfezione può essere modellata come un rumore di processo

Stima dello stato

- Nella realizzazione di uno stimatore dello stato in catena chiusa, cioè

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + L(y(t) - C\hat{x}(t))$$

la presenza del rumore è l'unica ragione che impedisce l'impiego di guadagni arbitrariamente alti (e quindi di stime arbitrariamente veloci). Infatti, l'osservatore è tipicamente implementato in un calcolatore digitale, ove non vi è alcuna limitazione all'impiego di guadagni elevati.

- **DOMANDA:** assumendo di avere tutte le informazioni circa il rumore (di processo e di misura) è possibile scegliere la matrice L in maniera ottima, rispetto ad un ragionevole criterio?
- **RISPOSTA:** Sì, ricorrendo al *filtro di Kalman-Bucy* (comunemente noto come *filtro di Kalman*)

Il filtro di Kalman

- Il filtro di Kalman è un osservatore ottimo (sotto certe ipotesi sul rumore e rispetto a un certo criterio di ottimalità)

- Si consideri il sistema dinamico lineare, non stazionario e continuo:

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B_u(t)u(t) + B_w(t)w(t) \\ y(t) = C(t)x(t) + v(t) \end{cases}$$

$w(t)$ rumore di processo (dovuto ad esempio a disturbi in ingresso)

$v(t)$ rumore di misura (dovuto ad esempio a imperfezioni del sensore di misura)

$y(t)$ uscita misurata (cioè uscita effettiva del sistema più il rumore di misura)

con stato iniziale $x(t_0) = x_0$

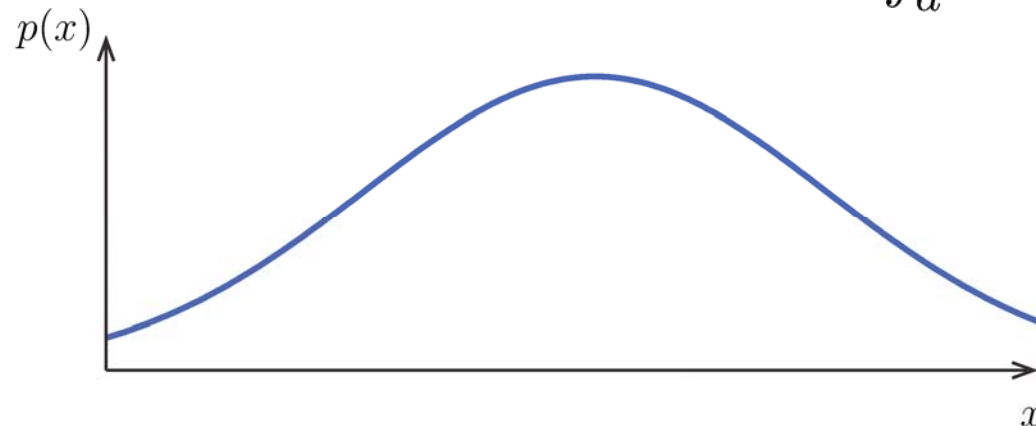
Processi stocastici

- Il rumore sulla misura e l'incertezza sul modello può essere modellato come un **processo stocastico**, ovvero come una variabile casuale caratterizzato da alcuni parametri (media, varianza, ...) che possono cambiare nel tempo
- Il modo in cui vengono descritte le incertezze agenti sul sistema è cruciale per sviluppare gli strumenti per l'osservazione dello stato.
- Il filtro di Kalman utilizza processi Gaussiani per modellare le incertezze agenti sul sistema. Questa descrizione è spesso (ma non sempre!) la migliore per molte applicazioni

Processi stocastici

- Una **variabile aleatoria continua** è una variabile che assume un valore casuale in un insieme continuo di valori ammissibili. Ciascun valore ammissibile è caratterizzato da una certa probabilità di essere assunto.
- Formalmente, se X è una variabile aleatoria continua:
 - X può assumere valori in un intervallo continuo $[x_1, x_2]$.
 - $p(X = x)$, o $p(x)$ (o $f(x)$), è la funzione di densità di probabilità che la variabile aleatoria X assuma il valore x

$$\Pr(x \in (a, b)) = \int_a^b p(x) dx$$



E' più probabile essere al centro della campana

Processi stocastici

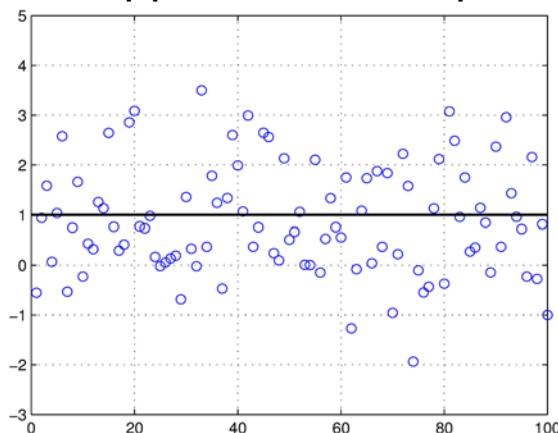
- **Valore atteso** (Expected Value):
 - generalizza il concetto di valore medio in un fenomeno aleatorio
 - nel caso di variabili aleatorie continue, se l'intervallo su cui è definita la variabile aleatoria X è \mathbb{R} e $p(x)$ indica la funzione densità di probabilità si ha che

$$E[X] = \int_{-\infty}^{\infty} xp(x)dx$$

- **Varianza** di un segnale intorno al suo valore atteso

$$\sigma^2 = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - E[X])^2 p(x) dx$$

- rappresenta la dispersione di una variabile aleatoria attorno al suo valore atteso

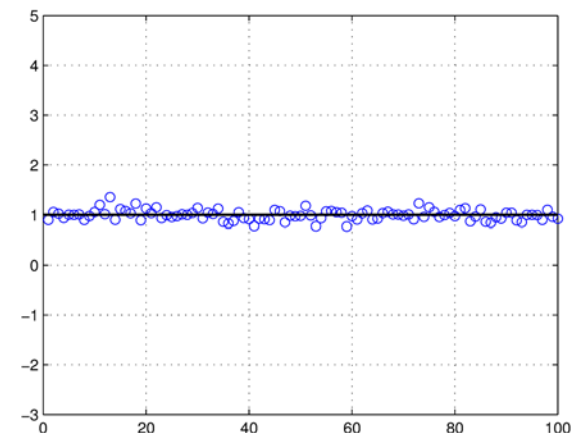


$$E[X] = 1$$

$$\sigma^2 = 1$$

$$E[X] = 1$$

$$\sigma^2 = 0.1$$



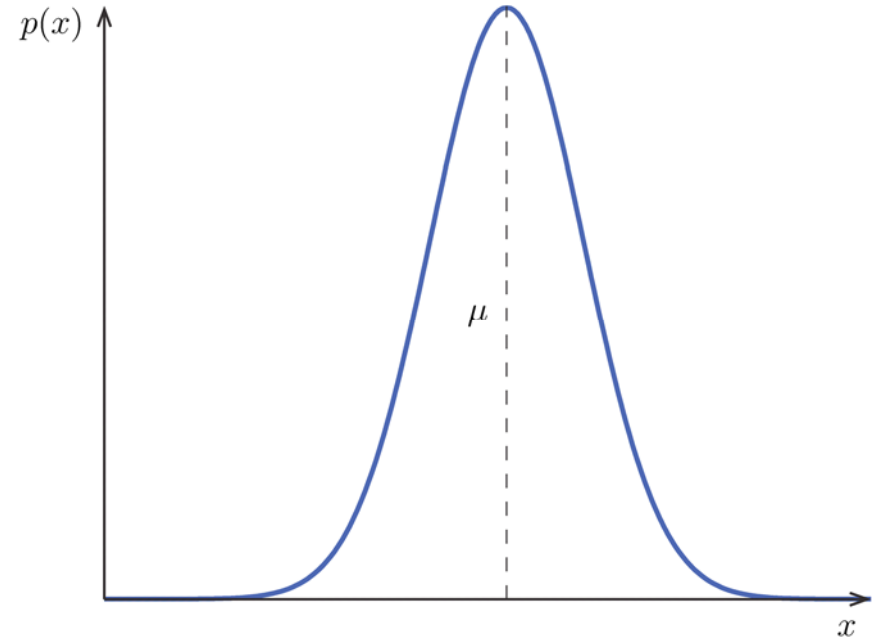
Processi stocastici

- Una variabile aleatoria continua si dice variabile gaussiana (o variabile normale o semplicemente gaussiana) se la sua densità di probabilità è una curva di Gauss del tipo

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

- Si dimostra che

- $E[X] = \mu$
- $E[(X - \mu)^2] = \sigma^2$



- Il valore atteso si dice anche valore medio o media della gaussiana
- Più la campana è “stretta” minore è la varianza
- Una variabile aleatoria Gaussiana è completamente descritta dal suo valore atteso e dalla sua varianza

Processi stocastici

- Un **vettore aleatorio** $X = (x_1, \dots, x_n)^T$ è un vettore le cui componenti sono variabili aleatorie
- Il valore atteso di un vettore aleatorio X è dato da

$$E[X] = \begin{bmatrix} E[x_1] \\ \vdots \\ E[x_n] \end{bmatrix}$$

- La matrice di covarianza Σ di un vettore aleatorio X è data da

$$\Sigma = E[(X - E[X])(X - E[X])^T] \quad \text{E' la generalizzazione della varianza!}$$

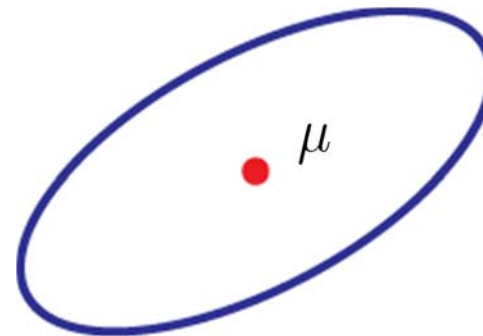
- Nel caso scalare la definizione di matrice di covarianza coincide con quella di varianza
- I termini sulla diagonale sono le varianze delle componenti del vettore X
- I termini fuori dalla diagonale indicano la correlazione che c'è tra le componenti di X
- Se un vettore aleatorio X è costituito da variabili aleatorie incorrelate, la matrice di covarianza è diagonale

Processi stocastici

- Un vettore aleatorio si dice **Gaussiano** (o normale) se la sua densità di probabilità è del tipo

$$p(X) = \frac{1}{\sqrt{(2\pi)^2 \det(\Sigma)}} e^{-\frac{1}{2}(X - \mu)^T \Sigma (X - \mu)}$$

- E' possibile mostrare che nel 95% dei casi il vettore si trova entro un **elissoide di incertezza** centrato in μ e descritto da $X^T \Sigma X = 1$



- Si dimostra che:

$$E[X] = \mu$$

$$E[(X - \mu)(X - \mu)^T] = \Sigma$$

Il filtro di Kalman: ipotesi - 1

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B_u(t)u(t) + B_w(t)w(t) \\ y(t) = C(t)x(t) + v(t) \end{cases}$$

- Si assume che sia w che v siano processi stocastici Gaussiani, stazionari, bianchi e a media nulla.
- $w(t)$ e $v(t)$ sono variabili casuali a distribuzione gaussiana $\forall t$.

$$E[w(t)] = 0, \quad \forall t$$

$$E[v(t)] = 0, \quad \forall t$$

$$E[w(t)w^T(t)] = \underline{W(t)}, \quad \forall t$$

$$E[v(t)v^T(t)] = \underline{V(t)}, \quad \forall t$$

$$E[w(t_1)w^T(t_2)] = 0, \quad \forall t_1 \neq t_2$$

$$E[v(t_1)v^T(t_2)] = 0, \quad \forall t_1 \neq t_2$$

Matrici simmetriche e definite positive (per definizione di covarianza)

Il filtro di Kalman: ipotesi - 2

- I rumori di processo e di misura sono fra loro indipendenti, ossia

$$E[w(t_1)v^T(t_2)] = 0, \quad \forall t_1, t_2$$

- Si assume inoltre che lo stato iniziale sia una variabile casuale Gaussiana di media e covarianza note:

$$E[x_0] = \bar{x}_0$$

$$E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = P_0$$

- Inoltre si assume che i processi stocastici $w(t)$ e $v(t)$ siano incorrelati con il vettore aleatorio x_0

Teorema (Filtro di Kalman)

- Se le coppie (A, B_w) e (A, C) sono rispettivamente controllabile e osservabile, allora l'osservatore

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + B_u(t)u(t) + L(t)(y(t) - C(t)\hat{x}(t))$$

è stabile e **ottimo**, nel senso che minimizza l'errore quadratico medio di stima

$$E[e^T(t)e(t)] \quad \text{con} \quad e(t) = x(t) - \hat{x}(t)$$

se il guadagno viene scelto come

$$L(t) = P(t)C^T(t)V^{-1}(t)$$

dove $P(t)$ è la soluzione **dell'equazione di Riccati**

$$\dot{P}(t) = A(t)P(t) + P(t)A^T(t) - P(t)C^T(t)V^{-1}(t)C(t)P(t) + B_w(t)W(t)B_w^T(t)$$

$$\text{con} \quad P(t_0) = P_0$$

Teorema (Filtro di Kalman)

- La stima fornita dall'osservatore ottimo, è non polarizzata (cioè il valor medio dell'errore di stima è nullo)

$$E[e(t)] = 0$$

- La matrice di covarianza dell'errore coincide con $P(t)$:

$$E[e(t)e^T(t)] = P(t)$$

- Il valore quadratico medio di stima è dato da

$$E[e^T(t)e(t)] = \text{tr}[P(t)]$$

Matrice dei guadagni di Kalman

- La matrice dei guadagni $L(t)$ costituisce un compromesso tra due esigenze distinte:
 - L'opportunità di utilizzare le misure disponibili per correggere la stima dello stato
 - La necessità di non peggiorare la stima corrente a causa degli errori sulla misura dell'uscita

$$L(t) = \underline{P(t)} \underline{C^T(t) V^{-1}(t)}$$

Guadagni dell'osservatore tanto più elevati quanto più elevato è l'errore sulla stima corrente

Proporzionalità tra i guadagni dell'osservatore e l'affidabilità delle misure sull'uscita

Filtro di Kalman nel caso stazionario

- Nel caso di sistemi e di processi stocastici stazionari

$$\begin{cases} \dot{x}(t) &= Ax(t) + B_u u(t) + B_w w(t) \\ y(t) &= Cx(t) + v(t) \end{cases}$$

con stato iniziale $x(t_0) = x_0$

e matrici di covarianza costanti $W(t) = W$ e $V(t) = V$

- il guadagno di Kalman risulta $L = PC^T V^{-1}$

dove P è l'unica soluzione simmetrica e definita positiva dell'equazione algebrica di Riccati

$$AP + PA^T - PC^T V^{-1} CP + B_w W B_w^T = 0$$

Osservazione

- L'equazione di Riccati appena vista si può ottenere dall'equazione algebrica di Riccati per il controllo LQ:

$$A^T S + AS - SB_u R^{-1} B_u^T S + C^T Q C = 0$$

effettuando le sostituzioni seguenti

$$A \rightarrow A^T$$

$$B_u \rightarrow C^T$$

$$C \rightarrow B_w^T$$

$$Q \rightarrow W$$

$$R \rightarrow V$$

Controllo LQG

- Se il progetto di controllo è di tipo LQ (a tempo infinito) e l'osservatore è un filtro di Kalman, si parla di controllo **LQG** (*Lineare Quadratico Gaussiano*)

- Dato il sistema

$$\begin{cases} \dot{x}(t) = Ax(t) + B_u u(t) + B_w w(t) \\ y(t) = Cx(t) + v(t) \end{cases} \quad \begin{array}{l} w(t) \text{ e } v(t) \text{ rumori bianchi} \\ \text{Gaussiani con matrici di} \\ \text{covarianza } W \text{ e } V \end{array}$$

il controllore ottimo che minimizza il (valore atteso del) funzionale

$$J = E \left[\int_0^{\infty} x^T(t) Q x(t) + u^T(t) R u(t) dt \right]$$

ha la forma

$$\begin{cases} u = -K \hat{x}(t) \\ \dot{\hat{x}}(t) = A \hat{x}(t) + B_u u(t) + L(y(t) - C \hat{x}(t)) \end{cases}$$

dove L è il guadagno dell'osservatore ottimo (filtro di Kalman) progettato ignorando il controllo e K è il guadagno del controllo ottimo progettato ignorando il rumore (*principio di separazione*)

Sistemi tempo discreti: filtro di Kalman

- La formulazione del problema del filtro di Kalman per i sistemi a tempo-discreto è analogo al caso continuo.

- Si consideri il sistema dinamico lineare, non stazionario e continuo:

$$\begin{cases} x(k+1) = A(k)x(k) + B_u(k)u(k) + B_w(k)\underline{w(k)} \\ y(k) = C(k)x(k) + \underline{v(k)} \end{cases}$$

rumore di processo (dovuto ad esempio a disturbi in ingresso)

rumore di misura (dovuto ad esempio a imperfezioni del sensore di misura)

uscita misurata (cioè uscita effettiva del sistema più il rumore di misura)

con stato iniziale $x(k_0) = x_0$

Ipotesi - 1

$$\begin{cases} x(k+1) = A(k)x(k) + B_u(k)u(k) + B_w(k)w(k) \\ y(k) = C(k)x(k) + v(k) \end{cases}$$

- Si assume che sia w che v siano processi stocastici Gaussiani, stazionari, bianchi e a media nulla.
- $w(k)$ e $v(k)$ sono variabili casuali a distribuzione gaussiana $\forall t$.

$$E[w(k)] = 0, \quad \forall k$$

$$E[v(k)] = 0, \quad \forall k$$

$$E[w(k)w^T(k)] = \underline{W(k)}, \quad \forall k$$

$$E[v(k)v^T(k)] = \underline{V(k)}, \quad \forall k$$

$$E[w(k_1)w^T(k_2)] = 0, \quad \forall k_1 \neq k_2$$

$$E[v(k_1)v^T(k_2)] = 0, \quad \forall k_1 \neq k_2$$

Matrici simmetriche e definite positive (per definizione di covarianza)

Ipotesi - 2

- I rumori di processo e di misura sono fra loro indipendenti, ossia

$$E[w(k_1)v^T(k_2)] = 0, \quad \forall k_1, k_2$$

- Si assume inoltre che lo stato iniziale sia una variabile casuale Gaussiana di media e covarianza note:

$$E[x_0] = \bar{x}_0$$

$$E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = P_0$$

- Inoltre si assume che i processi stocastici $w(k)$ e $v(k)$ siano incorrelati con il vettore aleatorio x_0

Teorema (Filtro di Kalman)

- Se le coppie (A, B_w) e (A, C) sono rispettivamente controllabile e osservabile, allora l'osservatore

$$\hat{x}(k+1) = A(k)\hat{x}(k) + B(k)u(k) + L(k)(y(k) - C(k)\hat{x}(k))$$

è stabile e **ottimo**, nel senso che minimizza l'errore quadratico medio di stima

$$E[e^T(k)e(k)] \quad \text{con} \quad e(k) = x(k) - \hat{x}(k)$$

se il guadagno viene scelto come

$$L(k) = A(k)P(k)C^T(k) \left[C(k)P(k)C^T(k) + V(k) \right]^{-1}$$

dove $P(k)$ è la soluzione **dell'equazione di Riccati**

$$P(k+1) = -A(k)P(k)C^T(k) \left[V(k) + C(k)P(k)C^T(k) \right]^{-1} C(k)P(k)A^T(k) + \\ + A(k)P(k)A^T(k) + B_w(k)W(k)B_w^T(k)$$

$$\text{con} \quad P(k_0) = P_0$$

Teorema (Filtro di Kalman)

- La stima fornita dall'osservatore ottimo, è non polarizzata (cioè il valor medio dell'errore di stima è nullo)

$$E[e(k)] = 0$$

- La matrice di covarianza dell'errore coincide con $P(k)$:

$$E[e(k)e^T(k)] = P(k)$$

- Il valore quadratico medio di stima è dato da

$$E[e^T(k)e(k)] = \text{tr}[P(k)]$$

Esempio: il filtro di Kalman nell'identificazione di un sistema

- Si consideri il sistema SISO tempo-discreto descritto da

$$G(z) = \frac{b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}} = \frac{Y(s)}{U(s)}$$

- Il modello può essere scritto nella forma

$$y(k) = -a_1 y(k-1) - a_2 y(k-2) - \dots - a_n y(k-n) + b_1 u(k-1) + b_2 u(k-2) + \dots + b_n u(k-n)$$

- Si indica con $\alpha = [-a_1, -a_2, \dots, -a_n, b_1, b_2, \dots, b_n]$ il vettore dei parametri e si suppone che tali parametri siano soggetti a perturbazioni descrivibili mediante relazioni del tipo

$$\alpha_i(k+1) = \alpha_i(k) + w_i(k) \quad w_i(t) \text{ rumore bianco Gaussiano}$$

- Si suppone che le misure disponibili (di ingresso e uscita) siano affette da errori

$$y(k) = -a_1 y(k-1) - a_2 y(k-2) - \dots - a_n y(k-n) + b_1 u(k-1) + b_2 u(k-2) + \dots + b_n u(k-n) + v(k)$$

$$v(t) \text{ rumore bianco Gaussiano}$$

Esempio: il filtro di Kalman nell'identificazione di un sistema

- Il problema della stima dei parametri α_i del sistema a partire dalle misure $y(k)$, $u(k)$ può essere basato sul filtro di Kalman considerando il processo con stato $x(k) = \alpha(k)$ descritto dal modello

$$\begin{aligned}x(k+1) &= x(k) + w(k) \\ y(k) &= C(k)x(k) + v(k)\end{aligned}$$

ove

$$C = [y(k-1), \dots, y(k-n), u(k-1), \dots, u(k-n)]$$

- Il filtro di Kalman fornisce la stima

$$\hat{x}(k+1) = (I - L(k)C(k))\hat{x}(k) + L(k)y(k)$$

con

$$L(k) = P(k)C^T(k) [C(k)P(k)C^T(k) + V(k)]^{-1}$$

$$P(k+1) = P(k) - P(k)C^T(k) [V(k) + C(k)P(k)C^T(k)]^{-1} C(k)P(k) + W(k)$$

$$V(k) = E[v(k)^2]$$

$$W(k) = E[w(k)^2]$$

- In applicazioni batch fornisce la stima ottenibile utilizzando i minimi quadrati ordinari

Filtro di Kalman: origini e storia

- Il **filtro di Kalman** è essenzialmente un set di equazioni matematiche che implementano uno stimatore del tipo **predittore-correttore** che è ottimo nel senso che minimizza la covarianza dell'errore di stima
- Dal momento della sua introduzione nel 1960 da parte di Rudolph E. Kalman, il filtro di Kalman è stato oggetto di numerose ricerche e applicazioni, in particolare nell'ambito dei veicoli spaziali a guida autonoma e della navigazione assistita
- Questo è probabilmente dovuto agli sviluppi dei sistemi di calcolo digitale che hanno reso possibile l'uso pratico del filtro di Kalman, e alla semplicità e robustezza dello stesso
- Nella pratica, **raramente le condizioni per ottenere l'ottimalità sono verificate**, ma il filtro sembra continuare a funzionare correttamente in molte di queste situazioni
- Molte forme diverse e differenti implementazioni del filtro sono state proposte in letteratura, anche per i sistemi non lineari (*Extended Kalman Filter, EKF*)

Filtro di Kalman discreto: formulazione standard

- Il filtro di Kalman prova a risolvere il problema della stima dello stato di un processo tempo-discreto governato dalle equazioni alle differenze lineari stocastiche

$$x_k = A x_{k-1} + B u_k + w_k$$

$$y_k = C x_k + v_k$$

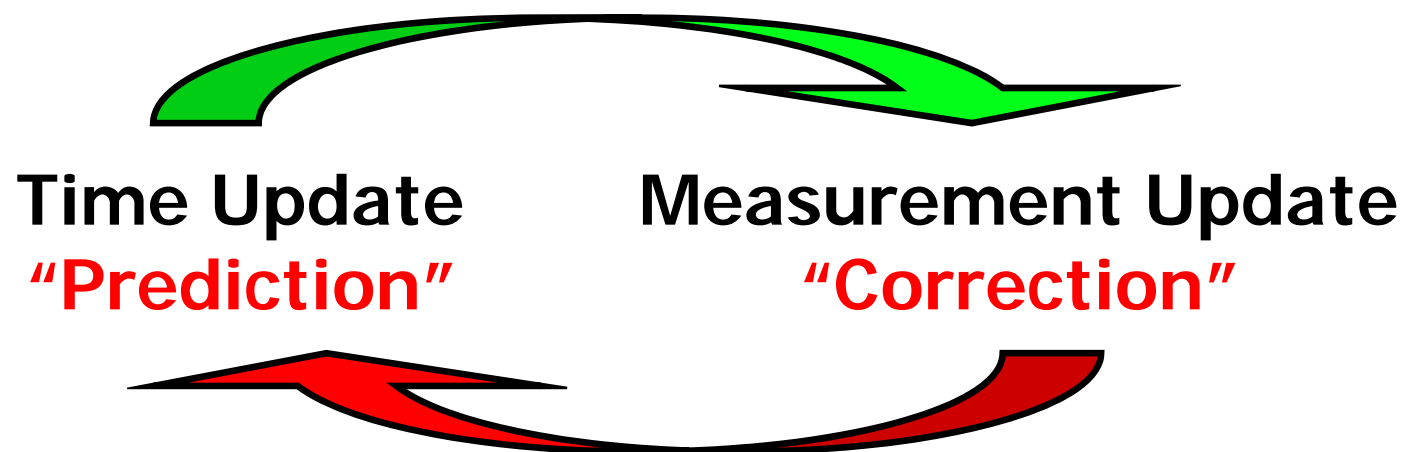
- w_k e v_k rappresentano il rumore di processo e di misura, rispettivamente. Si assume che siano variabili stocastiche indipendenti, bianche e Gaussiane con covarianza Q e R rispettivamente
- Si definisce \hat{x}_k^- (notare l'apice «-») lo stato **a priori** stimato all'istante k , sulla base della conoscenza del processo anteriore all'istante k , e \hat{x}_k la stima dello stato **a posteriori** all'istante k , data la misura y_k allo stesso istante.
- Si definiscono gli errori di stima **a priori** e **a posteriori** e le rispettive covarianze come

$$e_k^- = x_k - \hat{x}_k^- \quad P_k^- = E[e_k^- e_k^{-T}]$$

$$e_k = x_k - \hat{x}_k \quad P_k = E[e_k e_k^T]$$

Filtro di Kalman discreto: l'algoritmo

- Il filtro di Kalman stima lo stato del processo in certi istanti di tempo e quindi realizza un feedback sulla base delle misure (rumorose)
- Le equazioni del filtro di Kalman appartengono a due gruppi: predizione dello stato e aggiornamento basato sulle misure
- Le equazioni di predizione dello stato proiettano in avanti lo stato corrente e la covarianza dell'errore di stima al fine di ottenere una stima a priori per il successivo istante temporale
- Le equazioni di aggiornamento dello stato realizzano il meccanismo in retroazione, cioè incorporano le nuove misure nella stima a priori per ottenere una stima a posteriori migliorata



Filtro di Kalman discreto: l'algoritmo

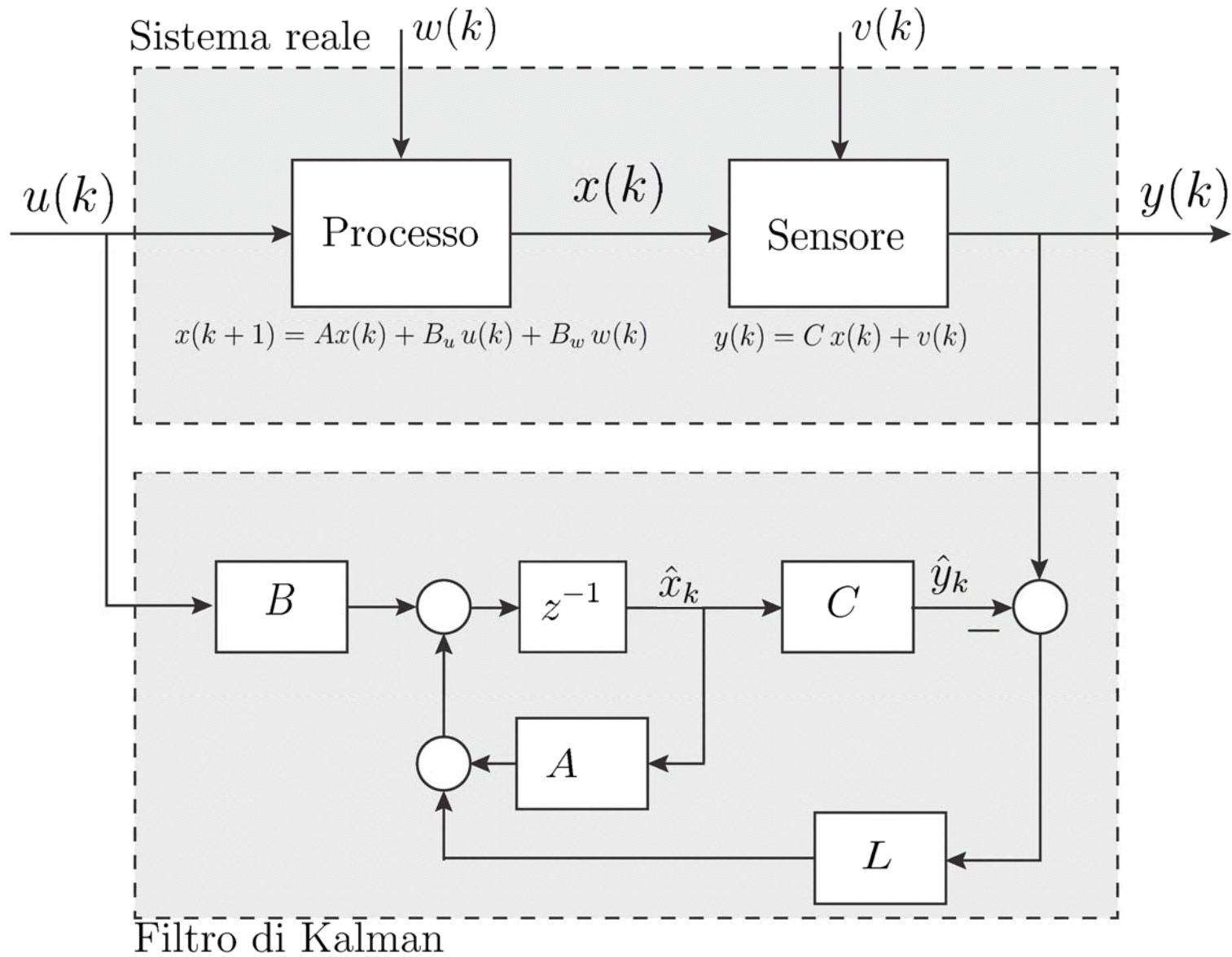
- **Predizione** - le equazioni proiettano lo stato e la covarianza dell'errore di stima in avanti dall'istante temporale $k-1$ all'istante k

$$\begin{aligned}\hat{x}_k^- &= A\hat{x}_{k-1} + B u_k \\ P_k^- &= AP_{k-1}A^T + Q\end{aligned}$$

- **Aggiornamento** – prima viene calcolata la matrice dei guadagni di Kalman L_k , quindi le misure y_k sono usate per generare una stima dello stato a posteriori. Alla fine, viene calcolata una stima a posteriori della covarianza dell'errore

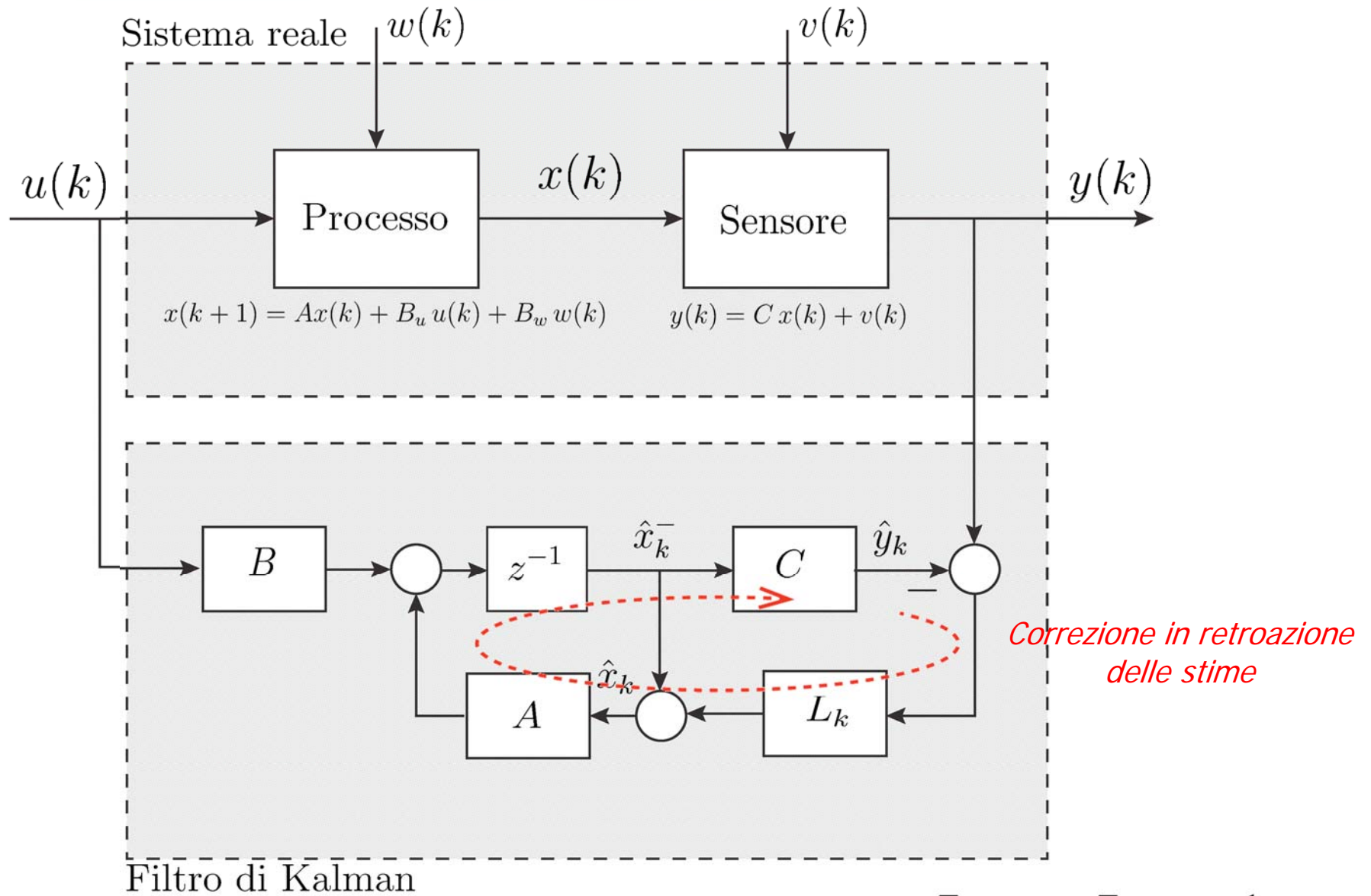
$$\begin{aligned}L_k &= P_k^- C^T (C P_k^- C^T + R)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + L_k \underbrace{(y_k - C\hat{x}_k^-)}_{\text{Innovazione}} \\ P_k &= (I - L_k C) P_k^-\end{aligned}$$

Struttura (standard) del filtro di Kalman



$$L(k) = A(k)P(k)C^T(k) [C(k)P(k)C^T(k) + V(k)]^{-1}$$

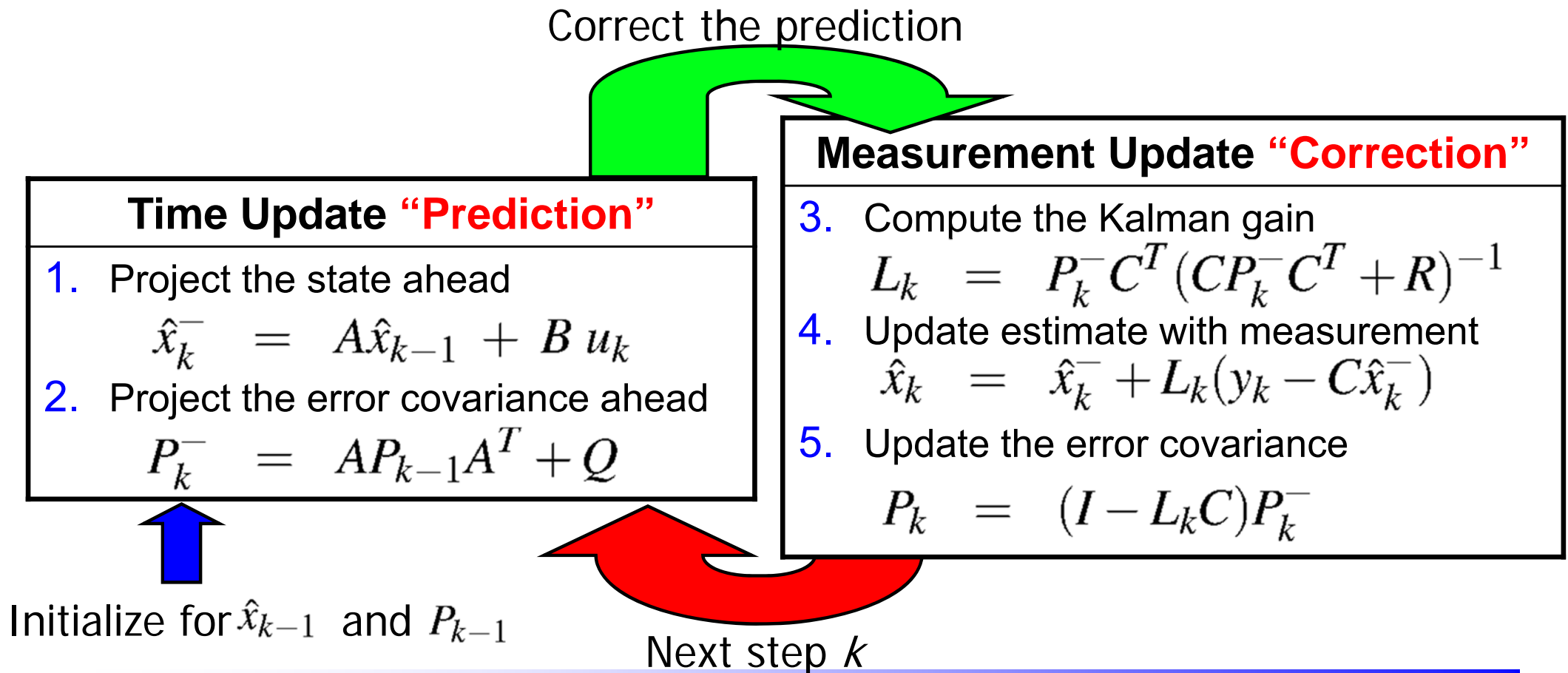
Struttura del filtro di Kalman



$$L_k = P_k^- C^T (C P_k^- C^T + R)^{-1}$$

Filtro di Kalman discreto: l'algoritmo

- Dopo ciascuna coppia predizione-correzione, il processo è ripetuto con la precedente *stima a posteriori* usata per proiettare o predire la nuova *stima a priori*
- La natura ricorsiva del filtro di Kalman lo rende adatto per implementazioni pratiche semplici ed efficaci



Filtro di Kalman discreto: l'algoritmo

- La differenza $(y_k - C\hat{x}_k^-)$ è chiamata **innovazione** o **residuo**. Il residuo riflette la differenza tra la misura predetta e la misura reale
- Se la covarianza delle misure tende a zero, il guadagno di Kalman pesa il residuo maggiormente

$$\lim_{R_k \rightarrow 0} L_k = C^T (CC^T)^{-1} = C^+$$

- Quando la stima a priori della covarianza dell'errore tende a zero, l'influenza del residuo diminuisce

$$\lim_{P_k^- \rightarrow 0} L_k = 0$$

- In altre parole, quando la covarianza delle misure tende a zero, la misura reale è sempre più affidabile, mentre la predizione è sempre meno sicura
- D'altro lato, quando la stima a priori della covarianza dell'errore va a zero, ci si affida sempre meno alla misura reale, mentre si fa sempre più affidamento sulla predizione

TEORIA DEI SISTEMI E DEL CONTROLLO

LM in Ingegneria Informatica e Ingegneria Elettronica

<http://www.dii.unimore.it/~lbiagiotti/TeoriaSistemiControllo.html>

Stima dello stato in presenza di disturbi: il filtro di Kalman

Ing. Luigi Biagiotti

e-mail: luigi.biagiotti@unimore.it

<http://www.dii.unimore.it/~lbiagiotti>