

Corso di Laurea in Ingegneria e Scienze Informatiche

Navigazione e Guida Autonoma di Veicoli Aerospaziali: Filtro di Kalman e sue varianti

Tesi di laurea in:
METODI NUMERICI

Relatore
Prof.ssa Damiana Lazzaro

Candidato
Marco Buda

Abstract

Max 2000 characters, strict.

Optional. Max a few lines.

Contents

Abstract	iii
1 Introduction	1
2 State of the art	3
2.1 I sistemi dinamici	3
2.2 Il filtro di Kalman	4
2.3 Definizione del problema lineare	4
2.4 Costruzione del filtro lineare	5
2.5 Varianti non lineari	11
2.6 Limitazioni	12
3 Contribution	15
3.1 Fancy formulas here	15
	17
Bibliography	17

CONTENTS

List of Figures

2.1	Diagramma di classificazione delle varianti del filtro di Kalman. . .	13
-----	---	----

LIST OF FIGURES

List of Listings

listings/HelloWorld.java	15
------------------------------------	----

LIST OF LISTINGS

Chapter 1

Introduction

Write your intro here.

You can use acronyms that you defined previously, such as `cro:IoT`Internet of Thing (IoT). If you use acronyms twice, they will be written in full only once (indeed, you can mention the IoT now without it being fully explained). In some cases, you may need a plural form of the acronym. For instance, that you are discussing `cro:vm`Virtual Machines (VMs), you may need both VM and VMs.

Marco Buda : Add sidenotes in this way. They are named after the author of the thesis

Structure of the Thesis

Marco Buda : At the end, describe the structure of the paper

Chapter 2

State of the art

2.1 I sistemi dinamici

Un sistema dinamico è un qualunque sistema in cui sia individuabile uno stato che evolve come funzione del tempo: $x = f(t)$.

Per gli scopi di questa tesi, lo stato verrà rappresentato come un vettore $x \in \mathbb{R}^n$, che raccoglie le variabili di stato (es.: posizione, velocità...), e il tempo verrà considerato come una quantità continua $t \in \mathbb{R}$, eventualmente campionata in determinati istanti t_k .

Modellizzare la realtà a livello macroscopico comporta spesso che f sia non deterministica, per via di fenomeni microscopici. Questo introduce la necessità di effettuare misurazioni periodiche (*measurement*, $z \in \mathbb{R}^m$) per monitorare la reale evoluzione dello stato in presenza di disturbi (*process noise*, $w \in \mathbb{R}^n$).

Sfortunatamente, le misurazioni stesse sono soggette ad ulteriori disturbi (*measurement noise*, $v \in \mathbb{R}^m$).

Nasce, quindi, l'esigenza di algoritmi che raccolgano ed interpretino i dati misurati per determinare una stima (*state estimate*, \hat{x}) dello stato reale (*ground truth*, x). Questi algoritmi sono detti, appunto, stimatori.

2.2 Il filtro di Kalman

Il filtro di Kalman [1] [2] [3] è uno stimatore lineare ricorsivo che minimizza l'errore quadratico medio.

L'aspetto ricorsivo si rifà al fatto che ogni stima \hat{x}_k è determinata in base alla stima precedente \hat{x}_{k-1} e alla misurazione attuale z_k , senza richiedere l'utilizzo esplicito di $\hat{x}_0, \hat{x}_1, \dots, \hat{x}_{k-2}$ o z_0, \dots, z_{k-1} . Si tratta di uno dei principali vantaggi del filtro, in quanto riduce sia la complessità temporale che quella spaziale, senza compromettere l'ottimalità.

Come anticipato, l'ottimalità è definita dal fatto che, ad ogni passo k , l'algoritmo produce la stima \hat{x}_k che minimizza la quantità $\mathbb{E}[\|e_k\|^2]$, con $e_k = \hat{x}_k - x_k$.

2.3 Definizione del problema lineare

Il filtro di Kalman a tempo discreto può essere applicato a sistemi dinamici che siano rappresentabili con un modello ricorsivo lineare:

$$x_k = A_k x_{k-1} + B_k u_{k-1} + w_{k-1} \quad (2.1)$$

dove A_k è la matrice $n \times n$ di evoluzione dello stato dal passo $k-1$ al passo k in assenza di disturbi e “azioni di controllo”, B_k è la matrice $l \times n$ di contributo delle “azioni di controllo” (*control input*, es.: forza di gravità, propulsione...) $u_{k-1} \in \mathbb{R}^l$ che agiscono dal passo $k-1$ al passo k e $w_{k-1} \in \mathbb{R}^n$ è il *process noise* che agisce dal passo $k-1$ al passo k .

In aggiunta, è richiesto che le misurazioni siano lineari rispetto allo stato:

$$z_k = H_k x_k + v_k \quad (2.2)$$

dove H_k è la matrice $m \times n$ che determina le variabili osservate in base allo stato che si vuole stimare e $v_k \in \mathbb{R}^m$ è il *measurement noise*.

Sono posti vincoli di non correlazione sulle variabili aleatorie:

$$\begin{aligned}
\text{Cov}(w_k, v_j) &= 0_{n \times m}, & \forall k, \forall j \\
\text{Cov}(w_k, u_j) &= 0_{n \times l}, & \forall k, \forall 0 \leq j \leq k-1 \\
\text{Cov}(v_k, u_j) &= 0_{m \times l}, & \forall k, \forall 0 \leq j \leq k-2 \\
\text{Cov}(w_k, w_j) &= 0_{n \times n}, & \forall k, \forall j \neq k \\
\text{Cov}(v_k, v_j) &= 0_{m \times m}, & \forall k, \forall j \neq k \\
\text{Cov}(w_k, x_0) &= 0_{n \times n}, & \forall k \\
\text{Cov}(v_k, x_0) &= 0_{m \times n}, & \forall k
\end{aligned}$$

Infine, nelle formulazioni standard del filtro, si richiede che w e v siano privi di *bias*, ossia che $\mathbb{E}[w_k] = \underline{0}$ e $\mathbb{E}[v_k] = \underline{0}$, ad ogni passo k .

2.4 Costruzione del filtro lineare

Nella letteratura sono riportate diverse derivazioni e dimostrazioni di ottimalità del filtro di Kalman [4, pp. 107-113] [5]. Si presenta qui una costruzione algebrica il più elementare possibile.

Il primo passo è descrivere l'espressione che calcoli la stima \hat{x}_k . Al passo $k \geq 1$ sono disponibili le informazioni riguardo \hat{x}_{k-1} , u_{k-1} e z_k , per cui l'espressione lineare avrà la forma generica:

$$\hat{x}_k = \mathcal{A}_k \hat{x}_{k-1} + \mathcal{B}_k u_{k-1} + \mathcal{K}_k z_k$$

Per quanto riguarda il passo $k = 0$, la scelta della stima iniziale \hat{x}_0 è libera, purché sia deterministica. Se si hanno informazioni sulla distribuzione di x_0 , è consigliabile scegliere $\hat{x}_0 = \mathbb{E}[x_0]$.

Una condizione implicita sul filtro richiede che anche l'errore sulle stime gen-

erate e_k sia privo di *bias*, ossia, ad ogni passo k :

$$\mathbb{E}[e_k] = \mathbb{E}[\hat{x}_k - x_k] = \underline{0}$$

Da questa condizione si otterranno informazioni sulle matrici incognite \mathcal{A} , \mathcal{B} e \mathcal{K} . In effetti, considerando per $k \geq 1$ e sostituendo \hat{x}_k con la sua definizione, si ottiene:

$$\mathbb{E}[\mathcal{A}_k \hat{x}_{k-1} + \mathcal{B}_k u_{k-1} + \mathcal{K}_k z_k - x_k] = \underline{0}$$

Successivamente, sostituendo z_k con la sua definizione:

$$\mathbb{E}[\mathcal{A}_k \hat{x}_{k-1} + \mathcal{B}_k u_{k-1} + \mathcal{K}_k (H_k x_k + v_k) - x_k] = \underline{0}$$

Sostituendo x_k con la sua relazione ricorsiva e manipolando i termini:

$$\begin{aligned} & \mathbb{E}[\mathcal{A}_k \hat{x}_{k-1} + \mathcal{B}_k u_{k-1} + \mathcal{K}_k (H_k (A_k x_{k-1} + B_k u_{k-1} + w_{k-1}) + v_k) - \\ & \quad - (A_k x_{k-1} + B_k u_{k-1} + w_{k-1})] = \underline{0} \\ \Rightarrow & \mathbb{E}[\mathcal{A}_k \hat{x}_{k-1} + \mathcal{B}_k u_{k-1} + \mathcal{K}_k H_k A_k x_{k-1} + \mathcal{K}_k H_k B_k u_{k-1} + \mathcal{K}_k H_k w_{k-1} + \\ & \quad + \mathcal{K}_k v_k - A_k x_{k-1} - B_k u_{k-1} - w_{k-1}] = \underline{0} \\ \Rightarrow & \mathbb{E}[\mathcal{A}_k \hat{x}_{k-1} + \mathcal{B}_k u_{k-1} + \mathcal{K}_k H_k A_k x_{k-1} + \mathcal{K}_k H_k B_k u_{k-1} + \mathcal{K}_k H_k w_{k-1} + \\ & \quad + \mathcal{K}_k v_k - A_k x_{k-1} - B_k u_{k-1} - w_{k-1} - \mathcal{A}_k x_{k-1} + \mathcal{A}_k x_{k-1}] = \underline{0} \\ \Rightarrow & \mathbb{E}[\mathcal{A}_k (\hat{x}_{k-1} - x_{k-1}) + (\mathcal{K}_k H_k A_k - A_k + \mathcal{A}_k) x_{k-1} + \\ & \quad + (\mathcal{K}_k H_k B_k - B_k + \mathcal{B}_k) u_{k-1} + (\mathcal{K}_k H_k - I) w_{k-1} + \mathcal{K}_k v_k] = \underline{0} \\ \Rightarrow & \mathcal{A}_k \mathbb{E}[\hat{x}_{k-1} - x_{k-1}] + (\mathcal{K}_k H_k A_k - A_k + \mathcal{A}_k) \mathbb{E}[x_{k-1}] + \\ & \quad + (\mathcal{K}_k H_k B_k - B_k + \mathcal{B}_k) \mathbb{E}[u_{k-1}] + (\mathcal{K}_k H_k - I) \mathbb{E}[w_{k-1}] + \mathcal{K}_k \mathbb{E}[v_k] = \underline{0} \end{aligned}$$

Sfruttando l'ipotesi che le quantità $\hat{x}_{k-1} - x_{k-1}$, w_{k-1} e v_k siano prive di *bias*:

$$(\mathcal{K}_k H_k A_k - A_k + \mathcal{A}_k) \mathbb{E}[x_{k-1}] + (\mathcal{K}_k H_k B_k - B_k + \mathcal{B}_k) \mathbb{E}[u_{k-1}] = \underline{0}$$

Non potendo fare assunzioni su x e u , ne segue che:

$$\mathcal{K}_k H_k A_k - A_k + \mathcal{A}_k = \underline{0} \quad , \quad \mathcal{K}_k H_k B_k - B_k + \mathcal{B}_k = \underline{0}$$

$$\Rightarrow \mathcal{A}_k = (I - \mathcal{K}_k H_k) A_k, \quad \mathcal{B}_k = (I - \mathcal{K}_k H_k) B_k$$

Risulta pratico definire la seguente quantità, per $k \geq 1$, come una stima *a priori*, ossia che non tenga conto della misurazione z_k :

$$\hat{x}_k^- = A_k \hat{x}_{k-1} + B_k u_{k-1}$$

Così facendo, si vanno a distinguere due fasi all'interno di ogni passo: una fase di predizione, o *time update*, in cui si calcola \hat{x}_k^- in base alle conoscenze sul modello, e una fase di correzione, o *measurement update*, in cui si combina \hat{x}_k^- con la misurazione z_k per ottenere una stima ottimale \hat{x}_k .

L'espressione per ricavare \hat{x}_k , ossia la stima *a posteriori*, diventa ora:

$$\begin{aligned} \hat{x}_k &= (I - \mathcal{K}_k H_k) A_k \hat{x}_{k-1} + (I - \mathcal{K}_k H_k) B_k u_{k-1} + \mathcal{K}_k z_k = \\ &= (I - \mathcal{K}_k H_k) (A_k \hat{x}_{k-1} + B_k u_{k-1}) + \mathcal{K}_k z_k = (I - \mathcal{K}_k H_k) \hat{x}_k^- + \mathcal{K}_k z_k = \\ &= \hat{x}_k^- + \mathcal{K}_k (z_k - H_k \hat{x}_k^-) \end{aligned}$$

Proseguendo, si definisca $e_k^- = \hat{x}_k^- - x_k$ come l'errore sulla stima *a priori* e si consideri:

$$\begin{aligned} e_k^- &= \hat{x}_k^- - x_k = (A_k \hat{x}_{k-1} + B_k u_{k-1}) - (A_k x_{k-1} + B_k u_{k-1} + w_{k-1}) = \\ &= A_k (\hat{x}_{k-1} - x_{k-1}) - w_{k-1} = A_k e_{k-1} - w_{k-1} \end{aligned}$$

Si osservi che anche e_k^- è privo di *bias*:

$$\mathbb{E}[e_k^-] = \mathbb{E}[A_k e_{k-1} - w_{k-1}] = A_k \mathbb{E}[e_{k-1}] - E[w_{k-1}] = \underline{0}$$

Si definiscano, per $k \geq 1$, le auto-covarianze degli errori $P_k^- = \text{Cov}(e_k^-, e_k^-)$ e $P_k = \text{Cov}(e_k, e_k)$.

Si ha:

$$\begin{aligned}
P_k^- &= \mathbb{E} [(e_k^- - \mathbb{E}[e_k^-])(e_k^- - \mathbb{E}[e_k^-])^T] = \mathbb{E} [e_k^- (e_k^-)^T] = \\
&= \mathbb{E} [(A_k e_{k-1} - w_{k-1})(A_k e_{k-1} - w_{k-1})^T] = \\
&= \mathbb{E} [(A_k e_{k-1} - w_{k-1}) (e_{k-1}^T A_k^T - w_{k-1}^T)] = \\
&= A_k \mathbb{E} [e_{k-1} (e_{k-1})^T] A_k^T - A_k \mathbb{E} [e_{k-1} (w_{k-1})^T] + \\
&\quad - \mathbb{E} [w_{k-1} (e_{k-1})^T] A_k^T + \mathbb{E} [w_{k-1} (w_{k-1})^T] = \\
&= A_k \mathbb{E} [(e_{k-1} - \underline{0})(e_{k-1} - \underline{0})^T] A_k^T - A_k \mathbb{E} [(e_{k-1} - \underline{0})(w_{k-1} - \underline{0})^T] + \\
&\quad - \mathbb{E} [(w_{k-1} - \underline{0})(e_{k-1} - \underline{0})^T] A_k^T + \mathbb{E} [(w_{k-1} - \underline{0})(w_{k-1} - \underline{0})^T] = \\
&= A_k \mathbb{E} [(e_{k-1} - \mathbb{E}[e_{k-1}])(e_{k-1} - \mathbb{E}[e_{k-1}])^T] A_k^T + \\
&\quad - A_k \mathbb{E} [(e_{k-1} - \mathbb{E}[e_{k-1}])(w_{k-1} - \mathbb{E}[w_{k-1}])^T] + \\
&\quad - \mathbb{E} [(w_{k-1} - \mathbb{E}[w_{k-1}])(e_{k-1} - \mathbb{E}[e_{k-1}])^T] A_k^T + \\
&\quad + \mathbb{E} [(w_{k-1} - \mathbb{E}[w_{k-1}])(w_{k-1} - \mathbb{E}[w_{k-1}])^T] = \\
&= A_k \text{Cov}(e_{k-1}, e_{k-1}) A_k^T - A_k \text{Cov}(e_{k-1}, w_{k-1}) + \\
&\quad - \text{Cov}(w_{k-1}, e_{k-1}) A_k^T + \text{Cov}(w_{k-1}, w_{k-1})
\end{aligned}$$

Si osservi che i termini centrali si annullano se e_{k-1} e w_{k-1} sono non correlati. In effetti, analizzando ricorsivamente l'errore e_{k-1} , si trova che le uniche variabili aleatorie da cui esso dipende sono u_0, \dots, u_{k-2} , w_0, \dots, w_{k-2} , v_1, \dots, v_{k-1} e x_0 , ossia variabili con cui w_{k-1} è non correlato per ipotesi.

Ricordando la definizione di P_{k-1} e definendo Q_{k-1} come la auto-covarianza di w_{k-1} , si trova, dunque, l'espressione:

$$P_k^- = A_k P_{k-1} A_k^T + Q_{k-1}$$

La scelta iniziale di P_0 è pressoché libera. Se si hanno informazioni sulla distribuzione di x_0 , è consigliabile utilizzare una stima della sua auto-covarianza. In ogni caso, è necessario avere $P_0 \neq 0_{n \times n}$ e semidefinita positiva.

Per trovare un'espressione per P_k , si consideri inizialmente:

$$\begin{aligned}\hat{x}_k &= (I - \mathcal{K}_k H_k) \hat{x}_k^- + \mathcal{K}_k z_k \\ \Rightarrow e_k &= \hat{x}_k - x_k = (I - \mathcal{K}_k H_k) \hat{x}_k^- + \mathcal{K}_k z_k - x_k\end{aligned}$$

Sostituendo z_k con la sua definizione:

$$\begin{aligned}e_k &= (I - \mathcal{K}_k H_k) \hat{x}_k^- + \mathcal{K}_k (H_k x_k + v_k) - x_k = \\ &= (I - \mathcal{K}_k H_k) \hat{x}_k^- + \mathcal{K}_k H_k x_k + \mathcal{K}_k v_k - x_k = \\ &= (I - \mathcal{K}_k H_k) \hat{x}_k^- - (I - \mathcal{K}_k H_k) x_k + \mathcal{K}_k v_k = \\ &= (I - \mathcal{K}_k H_k) (\hat{x}_k^- - x_k) + \mathcal{K}_k v_k = (I - \mathcal{K}_k H_k) e_k^- + \mathcal{K}_k v_k\end{aligned}$$

Dunque, si ha:

$$\begin{aligned}P_k &= \mathbb{E} [(e_k - \mathbb{E}[e_k])(e_k - \mathbb{E}[e_k])^T] = \mathbb{E} [e_k (e_k)^T] = \\ &= \mathbb{E} \left[((I - \mathcal{K}_k H_k) e_k^- + \mathcal{K}_k v_k) ((I - \mathcal{K}_k H_k) e_k^- + \mathcal{K}_k v_k)^T \right] = \\ &= \mathbb{E} \left[((I - \mathcal{K}_k H_k) e_k^- + \mathcal{K}_k v_k) ((e_k^-)^T (I - \mathcal{K}_k H_k)^T + v_k^T \mathcal{K}_k^T) \right] = \\ &= (I - \mathcal{K}_k H_k) \mathbb{E} [e_k^- (e_k^-)^T] (I - \mathcal{K}_k H_k)^T + (I - \mathcal{K}_k H_k) \mathbb{E} [e_k^- (v_k)^T] \mathcal{K}_k^T + \\ &\quad + \mathcal{K}_k \mathbb{E} [v_k (e_k^-)^T] (I - \mathcal{K}_k H_k)^T + \mathcal{K}_k \mathbb{E} [v_k (v_k)^T] \mathcal{K}_k^T\end{aligned}$$

Anche in questo caso, l'espressione si riduce alle covarianze. I termini centrali si annullano, poiché le uniche variabili aleatorie ad influenzare e_k^- sono u_0, \dots, u_{k-2} , w_0, \dots, w_{k-1} , v_1, \dots, v_{k-1} e x_0 , ossia variabili con cui v_k è non correlato per ipotesi. Ricordando la definizione di P_k^- e definendo R_k come la auto-covarianza di v_k , si trova l'espressione provvisoria:

$$P_k = (I - \mathcal{K}_k H_k) P_k^- (I - \mathcal{K}_k H_k)^T + \mathcal{K}_k R_k \mathcal{K}_k^T$$

Si osservi che minimizzare l'errore quadratico medio $\mathbb{E}[\|e_k\|^2]$ equivale a minimizzare la traccia di P_k . In effetti:

$$\mathbb{E}[\|e_k\|^2] = \mathbb{E}[e_k^T e_k] = \mathbb{E}[\text{tr}(e_k^T e_k)] = \mathbb{E}[\text{tr}(e_k (e_k)^T)] = \text{tr}(\mathbb{E}[e_k (e_k)^T]) = \text{tr}(P_k)$$

Dunque, si calcoli:

$$\begin{aligned}
\text{tr}(P_k) &= \text{tr}\left((I - \mathcal{K}_k H_k) P_k^- (I - \mathcal{K}_k H_k)^T + \mathcal{K}_k R_k \mathcal{K}_k^T\right) = \\
&= \text{tr}\left((I - \mathcal{K}_k H_k) P_k^- (I - H_k^T \mathcal{K}_k^T) + \mathcal{K}_k R_k \mathcal{K}_k^T\right) = \\
&= \text{tr}\left(P_k^- - P_k^- H_k^T \mathcal{K}_k^T - \mathcal{K}_k H_k P_k^- + \mathcal{K}_k H_k P_k^- H_k^T \mathcal{K}_k^T + \mathcal{K}_k R_k \mathcal{K}_k^T\right)
\end{aligned}$$

Essendo P_k^- simmetrica, vale $P_k^- H_k^T \mathcal{K}_k^T = (P_k^-)^T H_k^T \mathcal{K}_k^T = (\mathcal{K}_k H_k P_k^-)^T$, per cui:

$$\text{tr}(P_k) = \text{tr}(P_k^-) - 2 \text{tr}(\mathcal{K}_k H_k P_k^-) + \text{tr}(\mathcal{K}_k H_k P_k^- H_k^T \mathcal{K}_k^T) + \text{tr}(\mathcal{K}_k R_k \mathcal{K}_k^T)$$

Per ricercare \mathcal{K}_k che minimizzi $\text{tr}(P_k)$ si ponga:

$$\begin{aligned}
\frac{\partial \text{tr}(P_k)}{\partial \mathcal{K}_k} &= 0_{n \times n} \\
\Rightarrow \frac{\partial \text{tr}(P_k^-)}{\partial \mathcal{K}_k} - 2 \frac{\partial \text{tr}(\mathcal{K}_k H_k P_k^-)}{\partial \mathcal{K}_k} + \frac{\partial \text{tr}(\mathcal{K}_k H_k P_k^- H_k^T \mathcal{K}_k^T)}{\partial \mathcal{K}_k} + \\
&\quad + \frac{\partial \text{tr}(\mathcal{K}_k R_k \mathcal{K}_k^T)}{\partial \mathcal{K}_k} = 0_{n \times n}
\end{aligned}$$

Si osservi che $\text{tr}(P_k^-)$ è costante in K_k , per cui il primo termine si annulla. Per calcolare i termini restanti, si utilizzino le seguenti identità, con la seconda valida se N è simmetrica:

$$\frac{\partial \text{tr}(MN)}{\partial M} = N^T, \quad \frac{\partial \text{tr}(MNM^T)}{\partial M} = 2MN$$

Si ottiene:

$$\begin{aligned}
-2(H_k P_k^-)^T + 2\mathcal{K}_k (H_k P_k^- H_k^T) + 2\mathcal{K}_k R_k &= 0_{n \times n} \\
\Rightarrow -2P_k^- H_k^T + 2\mathcal{K}_k H_k P_k^- H_k^T + 2\mathcal{K}_k R_k &= 0_{n \times n} \\
\Rightarrow \mathcal{K}_k = P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1}
\end{aligned}$$

Esaminando la matrice Hessiana di $\text{tr}(P_k)$, la quale risulta essere semidefinita positiva, si può verificare che il valore trovato rappresenta un minimo globale.

La quantità \mathcal{K}_k è detta matrice dei guadagni di Kalman (*Kalman gain*) e può essere pensata come un indice di affidabilità delle misurazioni rispetto al modello teorico. In effetti, per $R_k \rightarrow 0_{m \times m}$ si ha $\mathcal{K}_k \rightarrow H^{-1}$, per cui $\hat{x}_k \rightarrow z_k$, mentre per $P_k^- \rightarrow 0_{n \times n}$ si ha $\mathcal{K}_k \rightarrow 0_{n \times m}$, per cui $\hat{x}_k \rightarrow \hat{x}_k^-$.

Sostituendo il valore trovato, l'espressione per P_k diventa:

$$\begin{aligned} P_k &= P_k^- - P_k^- H_k^T \mathcal{K}_k^T - \mathcal{K}_k H_k P_k^- + \mathcal{K}_k H_k P_k^- H_k^T \mathcal{K}_k^T + \mathcal{K}_k R_k \mathcal{K}_k^T = \\ &= (I - \mathcal{K}_k H_k) P_k^- - P_k^- H_k^T \mathcal{K}_k^T + \mathcal{K}_k (H_k P_k^- H_k^T + R_k) \mathcal{K}_k^T = \\ &= (I - \mathcal{K}_k H_k) P_k^- - P_k^- H_k^T \mathcal{K}_k^T + \\ &\quad + P_k^- H_k^T \underbrace{(H_k P_k^- H_k^T + R_k)^{-1} (H_k P_k^- H_k^T + R_k)}_I \mathcal{K}_k^T = \\ &= (I - \mathcal{K}_k H_k) P_k^- - P_k^- H_k^T \mathcal{K}_k^T + P_k^- H_k^T \mathcal{K}_k^T = \\ &= (I - \mathcal{K}_k H_k) P_k^- \end{aligned}$$

Riassumendo, l'algoritmo si basa sulle seguenti equazioni:

$$\begin{aligned} \hat{x}_0 &= \mathbb{E}[x_0] \\ P_0 &= \text{Cov}(x_0, x_0) \end{aligned}$$

$$\begin{aligned} \hat{x}_k^- &= A_k \hat{x}_{k-1} + B u_{k-1} \\ P_k^- &= A_k P_{k-1} A_k^T + Q_{k-1} \\ \mathcal{K}_k &= P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + \mathcal{K}_k (z_k - H_k \hat{x}_k^-) \\ P_k &= (I - \mathcal{K}_k H_k) P_k^- \end{aligned}$$

2.5 Varianti non lineari

Sulla base della formulazione originale di Kalman, sono state sviluppate numerose varianti del filtro, con l'obiettivo principale di estenderne il campo di applicazione, specie a modelli non lineari. Si presenta qui un elenco delle varianti più conosciute, successivamente schematizzate in fig. 2.1:

- KF ("Standard" Kalman Filter) [1] [2]: Semplice ed ottimale per sistemi

lineari.

- EKF (Extended Kalman Filter) [2]: Basato sulla linearizzazione di una funzione di transizione non lineare, con rischio di divergenza.
 - EKF2 o SO-EKF (Second-Order Extended Kalman Filter)
 - MEKF (Multiplicative Extended Kalman Filter) [6]: Specifico per la stima dell'assetto (orientamento) espresso come quaternione.
- Filtri Sigma-Point: Accurati e robusti in sistemi altamente dinamici, con costo di esecuzione maggiore.
 - UKF (Unscented Kalman Filter) [7]
 - QKF (Quadrature Kalman Filter), come il GHKF (Gauss-Hermite Kalman Filter)
 - CKF (Cubature Kalman Filter)
- Metodo Monte Carlo: Basato sull'utilizzo di campioni.
 - EnKF (Ensemble Kalman Filter)
 - PF (Particle Filter): Efficace anche quando i disturbi hanno distribuzione non Gaussiana.

Nell'utilizzo delle varianti EKF e Sigma-Point, spesso si sceglie di memorizzare P_k sotto forma di radice quadrata $\sqrt{P_k}$, per favorire la stabilità numerica dell'algoritmo. Si parla, in questo caso, di SRKF (Square-Root Kalman Filters).

2.6 Limitazioni

Nonostante oggi assuma un ruolo centrale nell'ingegneria aerospaziale e in diversi altri settori, il filtro di Kalman presenta alcune limitazioni intrinseche:

- Per sistemi non lineari con distribuzioni non Gaussiane, non è garantita, in generale, l'ottimalità.

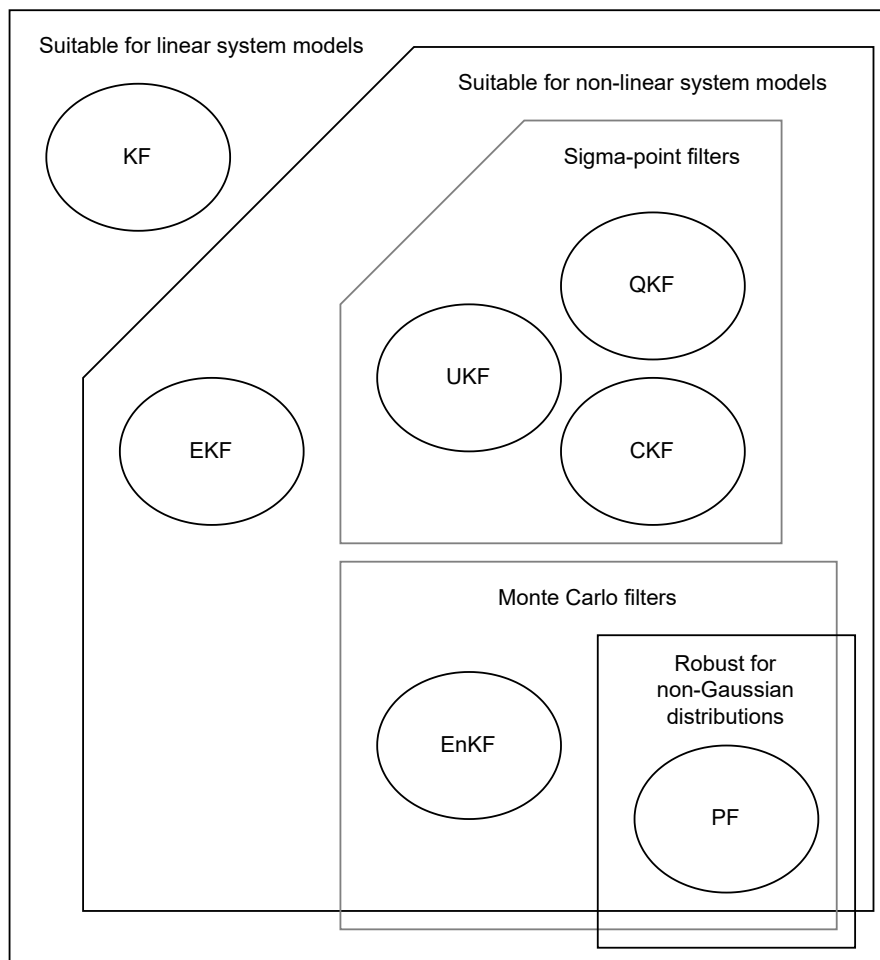


Figure 2.1: Diagramma di classificazione delle varianti del filtro di Kalman.

- Benché la conoscenza del sistema reale sia la base dell'efficacia del filtro, questa può anche costituire un limite sulla precisione, qualora modelli sufficientemente accurati non siano disponibili.

Chapter 3

Contribution

You may also put some code snippet (which is NOT float by default), eg: chapter 3.

3.1 Fancy formulas here

```
1 public class HelloWorld {
2     public static void main(String[] args) {
3         // Prints "Hello, World" to the terminal window.
4         System.out.println("Hello, World");
5     }
6 }
```

Bibliography

- [1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Journal of Basic Engineering*, vol. 82, pp. 35–45, Mar. 1960.
- [2] G. Welch and G. Bishop, “An introduction to the kalman filter,” Tech. Rep. 95-041, University of North Carolina at Chapel Hill, 1995.
- [3] Y. Pei, S. Biswas, D. S. Fussell, and K. Pingali, “An elementary introduction to kalman filtering,” *Communications of the ACM*, vol. 62, p. 122–133, Oct. 2019.
- [4] A. Gelb and T. A. S. Corporation, *Applied Optimal Estimation*. The MIT Press, 1974.
- [5] H. Masnadi-Shirazi, A. Masnadi-Shirazi, and M. A. Dastgheib, “A step by step mathematical derivation and tutorial on kalman filters.” ArXiv Preprint ArXiv:1910.03558, Oct. 2019.
- [6] F. L. Markley, “Multiplicative vs. additive filtering for spacecraft attitude determination,” in *Proceedings of the 6th Cranfield Conference on Dynamics and Control of Systems and Structures in Space*, pp. 467–474, 2004.
- [7] E. Wan and R. Van Der Merwe, “The unscented kalman filter for nonlinear estimation,” in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, pp. 153–158, 2000.

Acknowledgements

Optional. Max 1 page.