ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA

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

II Sessione di Laurea
Anno Accademico 2024-2025

Abstract

Max 2000 characters, strict.

Optional. Max a few lines.

Contents

List of Figures

List of Listings

Chapter 1

Introduction

Write your intro here.

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

You can use acronyms that your defined previously, such as cro:IoTInternet 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:vmVirtual Machines (VMs), you may need both VM and VMs.

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 [Kal60] 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 $\hat{x}_0, \hat{x}_1, ..., \hat{x}_{k-2}$ o $z_0, ..., z_{k-1}$. 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 $e_k = \hat{x}_k - x_k$.

2.2.1 Definizione del problema lineare

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

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

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:

$$z_k = H_k x_k + v_k \tag{2.2}$$

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 dei vincoli di non correlazione fra le variabili aleatorie:

$$
\begin{gathered}
\mathrm{Cov}(w_k, u_j) = 0_{n\times l}, \quad \forall\, k, \forall\, 0 \le j < k \\
\mathrm{Cov}(w_k, w_j) = 0_{n\times n}, \quad \forall\, k, \forall\, 0 \le j < k \\
\mathrm{Cov}(w_k, v_j) = 0_{n\times m}, \quad \forall\, k, \forall\, 1 \le j \le k \\
\mathrm{Cov}(w_k, x_0) = 0_{n\times n}, \quad \forall\, k
\end{gathered}
$$

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.

2.2.2 Costruzione del filtro lineare

Il primo passo è descrivere l'espressione che calcoli la stima $\hat{x}_k$. Al passo $k \ge 1$ 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 $\hat{x}_0 = \mathbb{E}[x_0]$, o una sua stima.

Una condizione implicita sul filtro richiede che anche l'errore sulle stime generate 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 $k \ge 1$ 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:

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

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

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

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

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

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

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

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

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

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

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

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

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

$$= \hat{x}_k^- + \mathcal{K}_k (z_k - H_k \hat{x}_k^-)$$

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

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

$$= A_k(\hat{x}_{k-1} - x_{k-1}) - w_{k-1} = A_k e_{k-1} - w_{k-1}$$

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 $P_k^- = \mathrm{Cov}(e_k^-, e_k^-)$ e $P_k = \mathrm{Cov}(e_k, e_k)$.

Si ha:

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

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 $u_0, ..., u_{k-2}$, $w_0, ..., w_{k-2}$, $v_1, ..., v_{k-1}$ 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.

I suggest referencing stuff as follows: fig. 2.1 or Figure 2.1

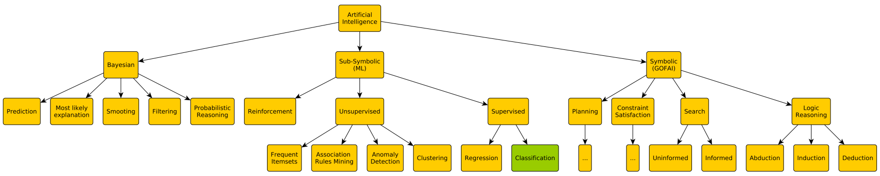

Figure 2.1: Some random image

2.3 Some cool topic

Chapter 3

Contribution

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

3.1 Fancy formulas here

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

Bibliography

[Kal60] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 03 1960.

Acknowledgements

Optional. Max 1 page.