



UNIVERSITÀ POLITECNICA DELLE MARCHE

FACOLTA' DI INGEGNERIA
Corso di Laurea in Ingegneria Informatica e dell'Automazione

**FILTRAGGIO INTRINSECO SU
GRUPPI DI LIE CON APPLICAZIONI
SULLA STIMA DELL'ASSETTO**

**INTRINSIC FILTERING ON LIE
GROUPS WITH APPLICATION TO
ATTITUDE ESTIMATION**

Relatore:
Chiar.mo Prof.
SIMONE FIORI

Presentata da:
SANDRO MATERAZZI

Sessione Straordinaria
Anno Accademico 2020/21

1 Introduzione

Questo documento ha lo scopo di riprodurre l'esperimento svolto da Axel Barrau e Silvère Bonnabel riguardo un approccio probabilistico al problema di filtrare un sistema con rumori in uno spazio matriciale (gruppo di Lie) con proprietà invarianti[1]. Sulla base del filtro di Kalman per sistemi lineari introducono una classe di filtri e ne studiano le proprietà, mostrando che l'equazione dell'errore è una catena di Markov che non dipende dalla stima del sistema. Quando il guadagno del filtro è fissato, la distribuzione dell'errore converge ad una distribuzione stazionaria sotto determinate condizioni. Introducono inoltre due metodi: IEKF e IEnKF, dove per il primo la matrice delle covarianze è dimostrato convergere. Si applicano poi questi metodi per la stima dell'assetto di un velivolo tramite simulazioni con dati sintetici.

2 Filtro di Kalman e Filtro di Kalman Esteso

Si danno ora alcune nozioni teoriche fondamentali sul filtro di Kalman e successivamente sul filtro di Kalman Esteso così da poter poi comprendere su quali basi è fondato il loro esperimento[2].

2.1 Il problema del filtraggio

Un sistema fisico è controllato da una serie di input esterni o da controlli e i suoi output sono misurati da apparecchi o sensori, quindi la conoscenza del sistema è data solamente dagli input e dalle misure sugli output.

Le misurazioni degli output e la struttura stessa del sistema portano degli errori e dell'incertezza nel sistema, la causa è la sensibilità dei sensori e delle componenti stesse, quindi tramite le misurazioni e gli input è necessario ottenere una stima dello stato del sistema che ottimizza un determinato criterio o proprietà, questo è il ruolo del filtro.

Il problema di filtraggio generale può essere formulato in questo modo, sia

$$\begin{cases} x(k+1) = f(x(k), u(k), w(k)) \\ y(k) = h(x(k), v(k)) \end{cases} \quad (1)$$

la dinamica dello stato di un sistema non-lineare tempo-variante, dove:

- $x \in R^n$ è il vettore dello stato del sistema,
- $f(., ., .)$ definisce la dinamica del sistema,
- $u \in R^m$ è il vettore di controllo,
- w è il vettore che rappresenta l'errore delle componenti del sistema,
- $y \in R^r$ è il vettore delle misurazioni,
- $h(., .)$ è la funzione di misurazione,
- v è il vettore che rappresenta l'errore nelle misurazioni.
- k è il passo temporale discreto

Qualsiasi tipo di filtro punta ad ottenere una stima ottima delle quantità desiderate (in questo caso lo stato del sistema) da dei dati presi in un ambiente con dei disturbi. La miglior stima possibile corrisponde alla minimizzazione dell'errore della stima dello stato.

Da un punto di vista Bayesiano il filtro propaga la funzione di densità di probabilità condizionata delle quantità desiderate, condizionata dalla conoscenza dei dati presi dalle misurazioni, ovvero il filtro valuta e propaga la funzione di densità di probabilità condizionata per valori crescenti di k ,

$$p(x(k) | y(1), \dots, y(k), u(0), \dots, u(k-1))$$

questa funzione descrive quanto ci si può fidare del valore di $x(k)$ che si ottiene. Si considera, per un istante di tempo k , la sequenza di input e misurazioni passate sono descritti dai due insiemi

$$U_0^{k-1} = u_0, u_1, \dots, u_{k-1}$$

$$Y_1^k = y_1, y_2, \dots, y_k.$$

L'intera evoluzione del sistema e il processo di filtraggio è descritta dai seguenti passi:

– Dato x_0 :

- vengono applicati i disturbi (w_0),
- vengono applicati i controlli (u_0),
- il sistema si muove nello stato x_1 ,
- si fa una misurazione y_1 ,

* Qual'è la migliore stima di x_1 ? \implies E' ottenuta da $p(x_1 | Y_1^1 U_0^0)$,

- vengono applicati i disturbi (w_1),
- vengono applicati i controlli (u_1),
- il sistema si muove nello stato x_2 ,
- si fa una misurazione y_2 ,

* Qual'è la migliore stima di x_2 ? \implies E' ottenuta da $p(x_2 | Y_1^2 U_0^1)$,

- ... ,
- ... ,
- ... ,
- ... ,

* Qual'è la migliore stima di x_k ? \implies E' ottenuta da $p(x_k | Y_1^k U_0^{k-1})$,

- vengono applicati i disturbi (w_k),
- vengono applicati i controlli (u_k),
- il sistema si muove nello stato x_{k+1} ,
- si fa una misurazione y_{k+1} ,

* Qual'è la migliore stima di x_{k+1} ? \implies E' ottenuta da $p(x_{k+1} | Y_1^{k+1} U_0^k)$.

Quindi il filtro propaga la funzione di densità di probabilità condizionata per valori crescenti di k , dove per ogni k , si ottiene la miglior stima possibile di \hat{x}_k che ottimizza una determinata proprietà

$$\begin{array}{lll} p(x_1 | Y_1^1 U_0^0) & \longrightarrow & \hat{x}_1 \\ p(x_2 | Y_1^2 U_0^1) & \longrightarrow & \hat{x}_2 \\ \dots & \longrightarrow & \dots \\ p(x_{k-1} | Y_1^{k-1} U_0^{k-2}) & \longrightarrow & \hat{x}_{k-1} \\ p(x_k | Y_1^k U_0^{k-1}) & \longrightarrow & \hat{x}_k \end{array}$$

dove \hat{x}_k è la miglior stima possibile dello stato x_k . In base al criterio con cui si sceglie di ottimizzare la stima si possono avere vettori diversi, la stima può essere in base a:

- la **media**, ovvero il centro di massa della probabilità, che corrisponde al criterio di minimizzazione dell'errore quadratico medio,
- la **moda**, ovvero il valore di x che ha la probabilità maggiore, che corrisponde al criterio di Massimo a Posteriori(MAP),
- la **mediana**, dove la stima è il valore di x che ha a destra e sinistra la stessa probabilità.

Si è quindi esposto il problema di filtraggio generale. Sotto particolari condizioni legate alla linearità della dinamica del sistema e ai vettori casuali coinvolti, le funzioni di densità di probabilità propagate dal filtro sono Gaussiane per ogni k e sono completamente definite dal vettore media e dalla matrice delle covarianze. Invece di propagare tutta la funzione di probabilità, il filtro propaga ricorsivamente solo il primo e il secondo momento della funzione di densità di probabilità condizionata. Il filtro generale si semplifica nel **Filtro di Kalman**, che valuta la stima in base all'errore quadratico medio minimo dello stato del sistema.

2.2 Filtro di Kalman

Si considera ora un sistema con dinamica lineare tempo-variante, le equazioni diventano

$$\begin{cases} x(k+1) = A_k x_k + B_k u_k + G w_k \\ y(k) = C x_k + v_k \end{cases} \quad (2)$$

dove $x(k) \in R^n$, $u(k) \in R^m$, $w(k) \in R^n$, $v(k) \in R^r$, $y(k) \in R^r$, w_k e v_k sono rumori bianchi gaussiani con media $E[w_k] = E[v_k] = 0$ e la matrice delle covarianze congiunta

$$E\left[\begin{pmatrix} w \\ v \end{pmatrix} \begin{pmatrix} w_k^T & v_k^T \end{pmatrix}\right] = \begin{bmatrix} Q_k & 0 \\ 0 & R_k \end{bmatrix}. \quad (3)$$

Lo stato iniziale x_0 è un vettore Gaussiano randomico con media $E[x_0] = \bar{x}_0$ e matrice delle covarianze $E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = \Sigma_0$. Il problema di filtraggio può essere interpretato come la stima di un vettore parametrico randomico. Per il sistema (5) il filtro di Kalman è quello che ottiene il minor errore quadratico medio della stima. Infatti, quando $x(0)$ è un vettore Gaussiano, i disturbi sullo stato e sulle misurazioni $w(k)$ e $v(k)$ sono rumori bianchi Gaussiani e le dinamiche dello stato e delle osservazioni sono lineari; le funzioni di densità di probabilità condizionata sono Gaussiane $p(x_k | Y_1^k, U_0^{k-1})$ per ogni k ; la media, la moda e la mediana di questa funzione di densità di probabilità condizionata coincidono; il filtro di Kalman è il migliore tra tutti i filtri possibili e ottimizza ogni possibile criterio che può essere considerato.

Sia:

$$p(x_k | Y_1^k, U_0^{k-1}) \sim N(\hat{x}(k|k), P(k|k)) \quad (4)$$

la funzione di densità di probabilità condizionata come una funzione di densità di probabilità Gaussiana. La stima dello stato $\hat{x}(k|k)$ è la media condizionata della funzione di densità di probabilità e la matrice delle covarianze $P(k|k)$ quantifica l'incertezza della stima:

$$\hat{x}(k|k) = E[x(k) | Y_k^1, U_0^{k-1}],$$

$$P(k|k) = E[(x(k) - \hat{x}(k|k))(x(k) - \hat{x}(k|k))^T | Y_k^1, U_0^{k-1}].$$

Il filtro di Kalman invece di propagare l'intera funzione di densità di probabilità condizionata, propaga solamente il primo ed il secondo momento. Questo metodo è ricorsivo, infatti per valutare $\hat{x}(k+1|k+1)$ il filtro di Kalman necessita solo della stima precedente $\hat{x}(k|k)$ e la nuova misurazione $y(k+1)$.

2.3 Dinamiche del Filtro di Kalman

Quando v_k , w_k e x_0 sono vettori Gaussiani, i vettori randomici x_k , x_{k+1} , Y_1^k sono congiuntamente Gaussiani. Come detto in precedenza il filtro di Kalman propaga la funzione di densità di probabilità Gaussiana $p(x_k | Y_1^k, U_0^{k-1})$ e quindi le dinamiche del filtro di Kalman definiscono la generica transizione da $p(x_k | Y_1^k, U_0^{k-1})$ a $p(x_{k+1} | Y_1^{k+1}, U_0^k)$, dove entrambe le funzioni di densità di probabilità sono Gaussiane e le informazioni disponibili su input e misurazioni ad un istante k e $k+1$ sono date dagli insiemi Y_1^k , Y_1^{k+1} , U_0^{k-1} , U_0^k . Questa procedura invece di essere svolta in modo diretto è divisa in due step: **prediction cycle** e **filtering** o **update cycle**.

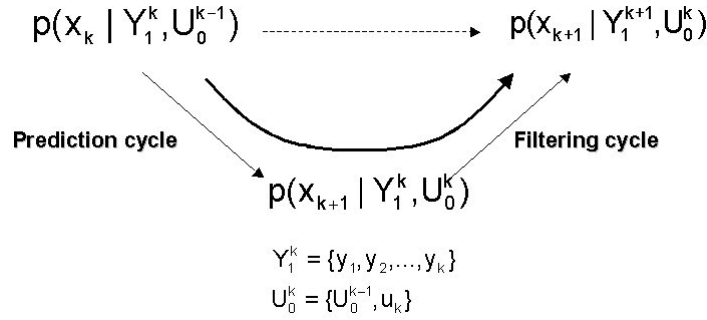


Figura 1: Ciclo di Predizione e Filtraggio nella dinamica del Filtro di Kalman.

La figura 4.2 mostra le dinamiche del filtro di Kalman, dove:

- $p(x_{k+1} | Y_1^{k+1}, U_0^k)$ rappresenta le informazioni che si hanno, per determinato istante $k+1$, dello stato $X(k+1)$ prima della misurazione $y(k+1)$;
- Il filtering cycle dice come aggiornare le informazioni di $x(k+1)$ dopo la misurazione $y(k+1)$.

Sia

$$p(x_k | Y_1^k, U_0^{k-1}) \sim N(\hat{x}(k|k), P(k|k)) \quad (5)$$

$$p(x_{k+1} | Y_1^k, U_0^k) \sim N(\hat{x}(k+1|k), P(k+1|k)) \quad (6)$$

dove $\hat{x}(k|k)$ e $\hat{x}(k+1|k)$ sono date da

$$\hat{x}(k|k) = E[x_k | Y_1^k, U_0^{k-1}] \quad (7)$$

$$\hat{x}(k+1|k) = E[x_{k+1} | Y_1^k, U_0^k] \quad (8)$$

mentre $P(k|k)$ e $P(k+1|k)$

$$P(k|k) = E[(x_k - \hat{x}(k|k))(x_k - \hat{x}(k|k))^T | Y_k^1, U_0^{k-1}] \quad (9)$$

$$P(k+1|k) = E[(x_{k+1} - \hat{x}(k+1|k))(x_{k+1} - \hat{x}(k+1|k))^T | Y_1^k, U_0^k]. \quad (10)$$

Per ricavare le dinamiche del filtro si suppone a questo punto che $p(x_k | Y_1^k, U_0^{k-1})$ è conosciuto, ovvero che sono dati $\hat{x}(k|k)$ e $P(k|k)$.

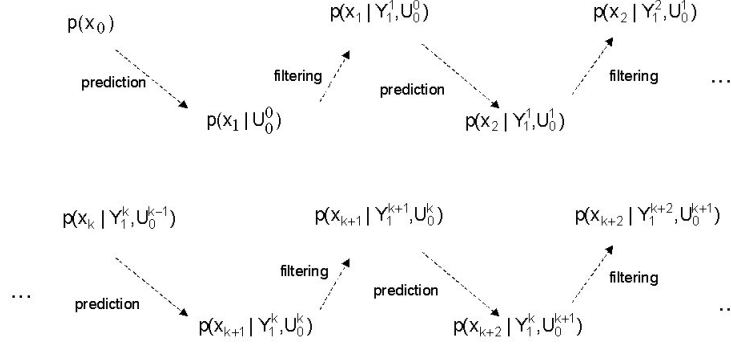


Figura 2: Consecutivi cicli di Predizione e Fitraggio nel Filtro di Kalman.

Si analizzano ora le dinamiche step by step:

Step 1 - State Prediction : calcolare $p(x_{k+1} | Y_1^k, U_0^k)$.

La densità funzione di densità di probabilità è completamente caratterizzata dalla media e dalla matrice delle covarianze. Applicando l'operatore media ad entrambi i membri della prima equazione del sistema (2) si ottiene

$$E[x_{k+1} | Y_1^k, U_0^k] = A_k E[x_k | Y_1^k, U_0^k] + B_k E[u_k | Y_1^k, U_0^k] + G E[w_k | Y_1^k, U_0^k]. \quad (11)$$

Considerando le equazioni (5) (6), w_k e Y_1^k sono vettori randomici indipendenti e w_k ha media 0, si ottiene:

$$\hat{x}(k+1|k) = A_k \hat{x}_k(k|k) + B_k u_k \quad (12)$$

Definiamo l'**errore di predizione** come

$$\tilde{x}(k+1|k) = x(k+1) - \hat{x}(k+1|k) \quad (13)$$

e sostituendo il valore di $x(k+1)$ e $\hat{x}(k+1|k)$ si ottiene

$$\tilde{x}(k+1|k) = A_k x_k + B_k u_k + G_k w_k - A_k \hat{x}(k|k) - B_k u_k = A_k \tilde{x}(k|k) + G_k w_k \quad (14)$$

dove l'**errore di filtraggio** è definito simile a (13)

$$\tilde{x}(k|k) = x(k) - \hat{x}(k|k). \quad (15)$$

Dato che $\tilde{x}(k|k)$ e w_k sono indipendenti, dalla (14) si ha

$$E[\tilde{x}(k+1|k)\tilde{x}(k+1|k)^T] = A_k E[\tilde{x}(k|k)\tilde{x}(k|k)^T] A_k^T + G_k Q G_k^T; \quad (16)$$

includendo in quest'ultima la (9) (10) risulta:

$$P(k+1|k) = A_k P(k|k) A_k^T + G_k Q G_k^T. \quad (17)$$

La stima prevista dello stato del sistema e la matrice delle covarianze associata, (12) e (17), corrispondono alla migliore conoscenza possibile che si può avere dello stato del sistema, ad un determinato istante $k+1$, prima della misurazione a quell'istante.

Step 2 - Measurement Prediction : calcolare $p(y_{k+1}|Y_1^k, U_0^k)$.
Dalla seconda equazione della (2) è evidente che

$$p(y_{k+1}|Y_1^k, U_0^k) = p(C_{k+1}x_{k+1}|Y_1^k, U_0^k), \quad (18)$$

che una funzione di densità di probabilità Gaussiana, così la **misurazione prevista** è data da

$$\hat{y}(k+1|k) = E[y_{k+1}|Y_1^k, U_0^k] = C_{k+1}\hat{x}_{k+1|k}. \quad (19)$$

Si definisce l'errore della misurazione prevista come

$$\tilde{y}(k+1|k) = y(k+1) - \hat{y}(k+1|k) \quad (20)$$

e sostituendo i valori di $y(k+1)$ e $\hat{y}(k+1|k)$ si ottiene

$$\tilde{y}(k+1|k) = C_{k+1}\hat{x}_{k+1|k} + v_{k+1}; \quad (21)$$

la matrice delle covarianze associata alla (21) è data da

$$P_{\tilde{y}}(k+1|k) = C_{k+1}P(k+1|k)C_{k+1}^T + R_k \quad (22)$$

Moltiplicando x_{k+1} a destra per $\tilde{y}(k+1|k)^T$ e usando la (21) si ha

$$x_{k+1}\tilde{y}(k+1|k)^T = x_{k+1}\hat{x}(k+1|k)^T C_{k+1}^T + x_{k+1}v_{k+1}^T \quad (23)$$

da cui

$$E[x_{k+1}\tilde{y}(k+1|k)^T] = P(k+1|k)C_{k+1}^T. \quad (24)$$

Dati la stima prevista dello stato ad un istante $k+1$ conoscendo le misurazioni fino al momento k , $\hat{x}(k+1|k)$ dalla (12), e tenendo in considerazione che l'errore nelle misurazioni ha media zero, data dalla dinamica lineare, è chiaro che la misurazione prevista (19) segue la stessa dinamica delle osservazioni del sistema reale.

Step 3 - Filtering : calcolare $p(x_{k+1}|Y_1^{k+1}, U_0^k)$.

Per calcolare la media condizionata di x_{k+1} si nota che Y_1^k e $Y_1^k, \tilde{y}(k+1|k)$ sono equivalenti dal punto di vista delle informazioni contenute, perciò

$$E[x_{k+1}|Y_1^{k+1}, U_0^k] = E[x_{k+1}|Y_1^k, \tilde{y}(k+1|k), U_0^k]. \quad (25)$$

Inoltre Y_1^k e $\tilde{y}(k+1|k)$ sono indipendenti, quindi

$$\hat{x}(k+1|k+1) = E[x(k+1)|Y_1^k] + E[x_{k+1} \cdot \tilde{y}(k+1|k) P_{\tilde{y}(k+1|k)}^{-1} \tilde{y}(k+1|k)]$$

che è equivalente a

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + P(k+1|k)C_{k+1}^T [C_{k+1}P(k+1|k)C_{k+1}^T + R]^{-1} [y(k+1) - C_{k+1}\hat{x}(k+1|k)]. \quad (26)$$

Definendo il **Guadagno di Kalman**

$$K(k+1) = P(k+1|k)C_{k+1}^T[C_{k+1}P(k+1|k)C_{k+1}^T + R], \quad (27)$$

si può riscrivere l'equazione (26) come

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1|k)[y(k+1) - C_{k+1}\hat{x}(k+1|k)], \quad (28)$$

da cui si conclude che la stima dello stato filtrato è ottenuta dalla stima prevista:

$$\boxed{\text{Stima stato filtrato} = \text{Stima prevista stato} + (\text{Guadagno} * \text{Errore})}.$$

L'errore è dato da $[y(k+1) - C_{k+1}\hat{x}(k+1|k)]$, ovvero è la differenza tra la misurazione reale ottenuta all'istante $k+1$ e la misurazione prevista ottenuta dal valore dello stato predetto. Questo esprime la nuova informazione che la nuova misurazione $y(k+1)$ porta nel filtro relativamente allo stato $x(k+1)$.

Definendo l'errore di filtraggio come

$$\hat{x}(k+1|k+1) = x(k+1) - \tilde{x}(k+1|k+1) \quad (29)$$

e sostituendolo nella (26) si ottiene:

$$\begin{aligned} \tilde{x}(k+1|k+1) &= \tilde{x}(k+1|k) - P(k+1|k)C_{k+1}^T \\ &\quad [C_{k+1}P(k+1|k)C_{k+1}^T + R]^{-1}[C_{k+1}\tilde{x}(k+1|k) + v_{k+1}] \end{aligned} \quad (30)$$

da cui

$$\begin{aligned} P(k+1|k+1) &= P(k+1|k) - P(k+1|k)C_{k+1}^T \\ &\quad [C_{k+1}P(k+1|k)C_{k+1}^T + R]^{-1}C_{k+1}P(k+1|k). \end{aligned} \quad (31)$$

Riassumendo:

Condizioni iniziali

$$\hat{x}(0|-1) = \bar{x}_0 \quad (32)$$

$$P(0|-1) = \Sigma_0 \quad (33)$$

Predizioni

$$\hat{x}(k+1|k) = A_k\hat{x}(k|k) + B_k u_k \quad (34)$$

$$P(k+1|k) = A_k P(k|k) A_k^T + G_k Q G_k^T \quad (35)$$

Filtraggio

$$\hat{x}(k|k) = \hat{x}(k|k-1) + K(k)[y(k) - C_k\hat{x}(k|k-1)] \quad (36)$$

$$K(k) = P(k|k-1)C_k^T[C_k P(k|k-1)C_k^T + R]^{-1} \quad (37)$$

$$P(k|k) = [I - K(k)C_k]P(k|k-1) \quad (38)$$

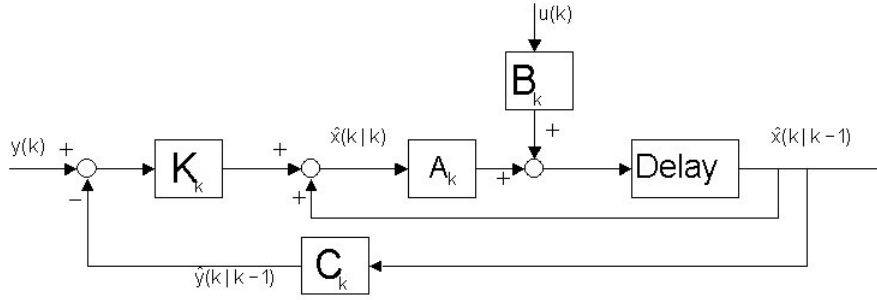


Figura 3: Schema a blocchi del Filtro di Kalman.

Il filtro di Kalman ha le seguenti le seguenti proprietà:

- il filtro di Kalman è un sistema lineare, tempo-variante discreto di dimensione finita, i cui input sono gli input del sistema u_1, u_1, \dots, u_{k-1} e le misurazioni del sistema y_0, y_1, \dots, y_k ; l'output è $\hat{x}(k|k-1) \hat{x}(k|k)$;
- la matrice di errore condizionata delle covarianze

$$P(k|k+1) = E[(x(k) - \hat{x}k|k+1))(x(k) - \hat{x}k|k+1))^T | Y_1^{k-1}, U_0^{k-1}] \quad (39)$$

è, in realtà, indipendente da Y_1^{k-1} , ciò significa che nessun set di misurazioni aiuta di più di un altro per eliminare l'incertezza rispetto allo stato $x(k)$. Anche il guadagno è indipendente da Y_1^{k-1} , perciò la matrice di errore delle covarianze $P(k|k-1)$ e il guadagno $K(k)$ possono essere calcolati prima che il filtro stesso sia attivo. Questo non è necessariamente vero per i filtri non lineari.

2.4 Il Filtro di Kalman Esteso

Si analizza ora il problema di filtraggio nel caso in cui il sistema sia non lineare. Si considera che il sistema non ha input esterni. La dinamica di un sistema non lineare è descritta dalle seguenti equazioni

$$\begin{cases} x_{k+1} = f_k(x_k) + w_k \\ y_k = h_k(x_k) + v_k \end{cases} \quad (40)$$

dove:

$$x_k \in R^n, f_k(x_k) : R^n \rightarrow R^n$$

$$y_k \in R^r, h_k(x_k) : R^n \rightarrow R^r$$

$$v_k \in R^r$$

$$w_k \in R^n$$

e $v_k w_k$ sono processi indipendenti bianchi Gaussiani random con media zero e matrice delle covarianze

$$E[v_k v_k^T] = R_k \quad E[w_k w_k^T] = Q_k \quad (41)$$

e x_0 è la condizione iniziale del sistema considerato come un vettore Gaussiano random

$$x_0 \sim N(\bar{x}_0, \Sigma_0). \quad (42)$$

Sia $Y_1^k = y_1, y_2, \dots, y_k$ un set di misurazioni. L'obiettivo del filtro è di ottenere una stima dello stato del sistema basato su queste misurazioni. Come visto precedentemente lo stimatore che minimizza l'errore quadratico medio calcola la media condizionata della funzione di densità di probabilità di x_k dato Y_1^k . Escluse particolari situazioni, il calcolo richiede la conoscenza completa della funzione di densità di probabilità. Un caso particolare è quando la dinamica del sistema è lineare, la condizione iniziale è un vettore random Gaussiano e i disturbi sulle misurazioni processi bianchi Gaussiani mutualmente indipendenti con media zero. Di conseguenza le funzioni di densità di probabilità condizionata $p(x(k)|Y_1^k)$, $p(x(k+1)|Y_1^k)$ e $p(x(k+1)|Y_1^{k+1})$ sono Gaussiane; con le dinamiche non lineari (40), queste funzioni non sono Gaussiane. Per calcolare il primo ed il secondo momento, il filtro ottimale non lineare deve propagare la funzione di densità di probabilità completa, che normalmente rappresenta un peso notevole dal punto di vista computazionale.

Il **Filtro di Kalman Esteso (EKF)** da un'approssimazione della stima ottimale. La non linearità della dinamica del sistema è approssimata da una versione linearizzata del modello del sistema non lineare intorno all'ultima stima dello stato.

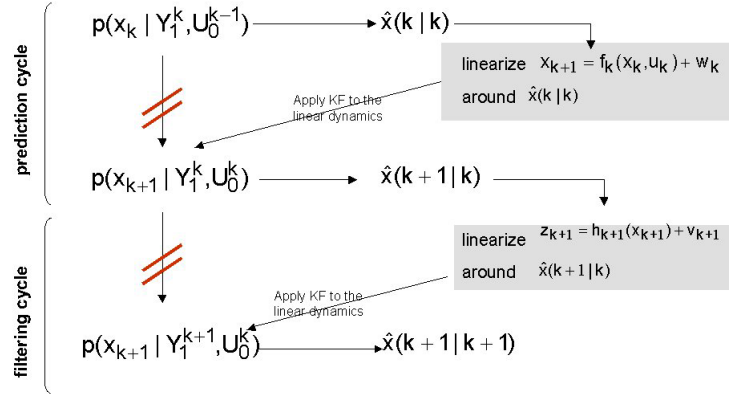


Figura 4: Dinamica del EKF.

Invece di propagare la funzione non Gaussiana, il filtro considera, ad ogni ciclo, una linearizzazione della dinamica non lineare intorno alle ultime stime (predetta e filtrata) consecutive dello stato e, per le dinamiche linearizzate, applica il Filtro di Kalman.

$$p(x_k | Y_1^k, U_0^{k-1}) \rightarrow p(x_{k+1} | Y_1^k, U_0^k) \rightarrow p(x_{k+1} | Y_1^{k+1}, U_0^k).$$

Un ciclo dell'EKF è composto dai seguenti step:

1. considera l'ultima stima dello stato filtrato $\hat{x}(k|k)$
2. linearizza la dinamica del sistema, $x_{k+1} = f(x_k) + w_k$ intorno a $\hat{x}(k|k)$

3. applica il Prediction step del Filtro di Kalman alla dinamica del sistema linearizzato appena ottenuto, produce $\hat{x}(k+1|k+1)$ e $P(k+1|k)$
4. linearizza le dinamiche delle misurazioni, $y_k = h(x_k + v_k)$ intorno a $\hat{x}(k+1|k)$
5. applica il filtering cycle del Filtro di Kalman alle dinamiche delle misurazioni linearizzate, produce $\hat{x}(k+1|k+1)$ e $P(k+1|k+1)$.

Siano $F(k)$ e $H(k)$ le matrici Jacobiane di $f(\cdot)$ e $h(\cdot)$

$$F(k) = \nabla f_k|_{\hat{x}(k|k)}$$

$$H(k+1) = \nabla h|_{\hat{x}(k+1|k)},$$

l'algoritmo del EKF è:

Ciclo di predizione

$$\hat{x}(k+1|k) = f_k(\hat{x}(k|k))$$

$$P(k+1|k) = F(k)P(k|k)F^T(k) + Q(k)$$

Ciclo di filtraggio

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)[y_{k+1} - h_{k+1}(\hat{x}(k+1|k))]$$

$$K(k+1) = P(k+1|k)H^T(k+1)[H(k+1)P(k+1|k)H^T(k+1) + R(k+1)]^{-1}$$

$$P(k+1|k+1) = [I - K(k+1)H(k+1)]P(k+1|k) \quad (43)$$

L'EKF non è un filtro ottimale, ma è basato su una serie di approssimazioni. Contrariamente al Filtro di Kalman, l'EKF potrebbe divergere se le continue linearizzazioni non sono buone approssimazioni del modello lineare.

3 Filtrare su gruppi di Lie

Essendo il problema in un gruppo di Lie matriciale, si introducono alcune nozioni sulla struttura di un generico gruppo di Lie per poter comprendere come il filtro di Kalman per sistemi lineari sia stato applicato ad uno spazio non Euclideo.

3.1 Gruppi di Lie

Un gruppo G [3][4] è definito come l'insieme di elementi g che soddisfano delle proprietà:

1. esiste una legge di composizione: presi $g_1, g_2 \in G$ allora $g_1 g_2 = g_3 \in G$,
2. esiste l'elemento identità: $\exists e \in G$ tale che $ge = eg = g$,
3. esiste l'elemento inverso: se $g \in G$ allora $\exists g^{-1} \in G \mid gg^{-1} = g^{-1}g = e$,
4. associativa: $(g_1 g_2)g_3 = g_1(g_2 g_3)$.

I gruppi di Lie sono quelli i cui elementi dipendono in modo continuo da alcuni parametri, ad esempio le rotazioni intorno all'asse z formano un gruppo di Lie i cui elementi sono parametrizzati da un angolo $\theta \in [0, 2\pi]$.

Un gruppo di Lie è un "smooth differential manifold", ovvero un manifold localmente simile ad uno spazio lineare e analitico ($\subseteq C^\infty$, differenziabile per ogni grado). Studiando le trasformazioni infinitesime generate dal gruppo, cioè trasformazioni poco differenti dall'identità, si ottiene l'**algebra di Lie** del gruppo, un'algebra che riassume le funzioni essenziali del gruppo, quindi una versione linearizzata di se stesso.

Un elemento $g(\alpha)$ di un gruppo di Lie G si può parametrizzare nel seguente modo:

$$g(\alpha) = e^{i\alpha_a T^a} \in G \quad a = 1, \dots, \dim G \quad (44)$$

dove i parametri α_a sono numeri reali che parametrizzano i vari elementi del gruppo e sono scelti in modo tale che per $\alpha_a = 0$ si ha l'identità $g = 1$, i è l'unità immaginaria, ovvero la soluzione dell'equazione $x^2 = -$, mentre gli operatori T^a sono i generatori del gruppo. Pensando il gruppo nella rappresentazione definita come gruppo di matrici $n \times n$, anche i generatori risultano essere matrici $n \times n$. Essi generano trasformazioni infinitesime quando $\alpha_a \ll 1$ (basta sviluppare in serie di Taylor la funzione esponenziale e tenere i termini di ordine basso)

$$g = 1 + i\alpha_a T^a + \dots \quad (45)$$

dove 1 indica l'elemento identità del gruppo. Studiando la relazione che cattura le proprietà di composizione del gruppo con trasformazioni infinitesime (che in generale sono non commutative) si ottiene l'algebra di Lie del gruppo G

$$[T^a, T^b] = i f_c^{ab} T^c \quad (46)$$

Le costanti f_c^{ab} sono costanti di struttura del gruppo e lo caratterizzano quasi completamente (gruppi diversi ma con stessa algebra differiscono per la loro topologia ma localmente sono simili).

3.1.1 Struttura di un generico gruppo di Lie

Un elemento del gruppo non abeliano G connesso all'identità può essere parametrizzato con le coordinate α_a collegate ai corrispondenti generatori T^a . Elenchiamo alcune definizioni e proprietà generali

- (i) $g = \exp(i\alpha_a T^a) \in G$
- (ii) $[T^a, T^b] = i f_c^{ab} T^c$
- (iii) $\text{tr}(T^a T^b) = \frac{1}{2} \gamma^{ab}$
- (iv) $[[T^a, T^b], T^c] + [[T^b, T^c], T^a] + [[T^c, T^a], T^b] = 0$
 $\implies f_d^{ab} f_e^{dc} f_d^{bc} f_e^{da} + f_d^{ca} f_e^{db} = 0$
- (v) $f^{abc} = f_d^{ab} \gamma^{dc}$

La (i) descrive la parametrizzazione esponenziale di un elemento arbitrario del gruppo che sia connesso all'identità. L'indice a assume tanti valori quante le dimensioni del gruppo. Un elemento del gruppo è quindi parametrizzato dai parametri α_a con $a = 1, \dots, \dim(G)$.

La (ii) corrisponde all'algebra di Lie descritta dai generatori infinitesimi T^a . Le costanti f_c^{ab} sono dette costanti di struttura e caratterizzano il gruppo G .

La (iii) identifica una metrica γ^{ab} detta "metrica di Killing". Tale metrica è definita positiva solo per gruppi di Lie compatti.

Le (iv) sono le cosiddette "identità di Jacobi" che possono essere sfruttate per costruire la rappresentazione aggiunta dell'algebra di Lie e del relativo gruppo. Infatti, denotando con $(T_{(A)}^a)_c^b$ gli elementi di matrice dei generatori della rappresentazione aggiunta $T_{(A)}^a$, si ha $(T_{(A)}^a)_c^b = -i f_c^{ab}$. Le identità di Jacobi permettono di provare che questa è una rappresentazione di dimensioni uguali alle dimensioni del gruppo, poiché $a, b, c = 1, 2, \dots, \dim G$.

In (v) si è usata la metrica di Killing per alzare un indice nelle costanti di struttura. Le f^{abc} sono completamente antisimmetriche in tutti gli indici: questa proprietà si può dedurre prendendo la traccia delle identità di Jacobi del punto (iv) ed usando la (ii) e la (iii). L'antisimmetria negli indici a e b è ovvia per la (ii).

A questo punto è utile citare la formula di Baker-Campbell-Hausdorff per il prodotto degli esponenziali di due operatori lineari A e B

$$e^A e^B = e^{A+B + \frac{1}{2}[A,B] + \frac{1}{12}[A,[A,B]] - \frac{1}{12}[B,[A,B]] + \dots} \quad (47)$$

Questa formula mostra che la conoscenza dell'algebra di Lie è sufficiente per ricostruire il prodotto (in generale non-commutativo) del corrispondente gruppo di Lie.

4 Impostazione del problema

Dopo aver introdotto qualche nozione teorica sulla quale si fonda l'esperimento, si introducono i modelli matematici su cui andare a lavorare.

4.1 Modello a tempo-continuo

Si considera una variabile di stato χ_t che prende valori in un Gruppo di Lie matriciale G con elemento neutrale I_d e il modello tempo-continuo con misurazioni discrete:

$$\begin{cases} \frac{d}{dt}\chi_t = (v_t + w_t)\chi_t + \chi_t\omega_t \\ Y_n = h(\chi_{t_n}, V_n) \end{cases} \quad (48)$$

dove v_t e ω_t sono input che prendono valori nell'algebra di Lie G , w_t è un rumore bianco Langeviano con matrice di diffusione Σ_t , $(Y_n)_{n \geq 0}$ sono le misurazioni tempo-discreto appartenenti ad uno spazio misurabile Υ e V_n è un rumore che prende valori in R^p , per un $p > 0$. Si fa inoltre la seguente ipotesi:

Ipotesi 1: equivarianza destra-sinistra La funzione di output h è equivariante destra-sinistra, ovvero esiste un'operazione sinistra di G in Υ tale che valga la seguente legge:

$$\forall g, \chi \in G, h(\chi g, V_n) = g^{-1}(\chi, V_n) \quad (49)$$

$$e h(\chi g, 0) = g^{-1}h(\chi, 0) \quad (50)$$

si può vedere χ_t come una matrice di rotazione e quindi sostituire l'ipotesi precedente con la seguente, più stretta, ipotesi:

Ipotesi 1 bis $\forall n \geq 0 \exists b$, vettore, tale che $h(\chi, V_n) = \chi^{-1}(b + V_n)$.

Questa ipotesi può sembrare restrittiva, ma è verificata in pratica in vari casi.

4.2 Stima dell'assetto sulla Terra piatta

Si prende per esempio un modello per la stima dell'assetto di un corpo rigido considerando la Terra come piatta.

$$\begin{cases} \frac{d}{dt}R_t = R_t(\omega_t + w_t^\omega)_\times \\ Y_n = (R_{t_n}^T b_1 + V_n^1, R_{t_n}^T b_2 + V_n^2) \end{cases} \quad (51)$$

$R_t \in SO(3)$ rappresenta la rotazione che serve per passare dal sistema di riferimento della terra a quello del corpo, ovvero l'angolazione che c'è tra i due sistemi di riferimento. $\omega_t \in R^3$ è il vettore di rotazione istantanea e $w_t^\omega \in R^3$ è un rumore bianco Gaussiano continuo che rappresenta il rumore del giroscopio. Sia $(x)_\times \in R^{3 \times 3}$ una funzione che trasforma un vettore 3×3 nella sua matrice antisimmetrica associata col prodotto vettoriale con un vettore:

$$(v)_\times w = v \times w = \begin{bmatrix} cy + bz \\ cx - az \\ -bx + ay \end{bmatrix} \implies (v)_\times = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}$$

con $v = [a \ b \ c]^T$ e $w = [x \ y \ z]^T$; ovvero se si moltiplica la matrice antisimmetrica $(v)_\times$ per un vettore si ha lo stesso risultato del prodotto vettoriale $v \times w$.

$(Y_n)_{n \geq 0}$ è una serie di misurazioni disturbate di due vettori b_1, b_2 del sistema di riferimento della Terra che verificano $b_1 \times b_2 \neq 0$, e V_n^1, V_n^2 sono sequenza di rumori bianchi Gaussiani isotropici indipendenti.

Il sistema (51) non coincide con il (48), che corrisponde ad un sistema di riferimento fissato sulla Terra con rumori definiti. Può essere risolto nel caso particolare in cui il rumore del giroscopio sia isotropico:

Definizione 1. Un rumore Langeviano w_t di G è detto isotropico se $\tilde{w}_t := gw_tg^{-1} = w_t \forall g \in G$.

Si nota che per un rumore che prende valori in $SO(3)$ questa definizione corrisponde intuitivamente al un rumore isotropico in R^3 . Quindi il sistema può essere riscritto come:

$$\begin{cases} \frac{d}{dt} R_t = (\tilde{w}_t^\varphi)_\times R_t + R_t(\omega_t)_\times, \\ Y_n = (R_{t_n}^T b_1 + V_n^1, R_{t_n}^T b_2 + V_n^2), \end{cases} \quad (52)$$

che corrisponde al (48).

4.3 Discretizzazione del modello considerato

Per trattare rigorosamente il problema di integrare misurazioni discrete, bisogna discretizzare il modello tempo-continuo con gli stessi time-step delle misurazioni. Al contrario del caso generale di stime non lineari, la discretizzazione esatta si può ottenere:

Teorema 1. Sia $\chi_n = \chi_{t_n}$, il sistema discreto (χ_n, Y_n) soddisfa le seguenti equazioni

$$\begin{cases} \chi_{n+1} = \Gamma_n W_n \chi_n \Omega_n \\ Y_n = h(\chi_n, V_n), \end{cases} \quad (53)$$

dove W_n è una variabile random con valori in G , la cui legge dipende dai valori che prende v_t per $t \in [t_n, t_{n+1}]$, Γ_n e Ω_n sono elementi di G che dipendono solo dai valori presi rispettivamente da v_t e ω_t per $t \in [t_n, t_{n+1}]$. La dimostrazione è riportata nel documento[1].

5 Filtri intrinseci tempo discreto

Si considera il caso particolare di un sistema lineare in R^n della forma $\frac{d}{dt}X_t = X_t + b_t + w_t$, con w_t un rumore bianco e b_t un input deterministico noto. Si assume di avere misurazioni discrete della forma $Y_n = HX_{t_n} + V_n \in R^p$, con V_n un rumore Gaussiano in R^p . Da una discretizzazione si ottiene:

$$\begin{cases} X_{n+1} = X_n + B_n + W_n \\ Y_n = HX_n + V_n \end{cases} \quad (54)$$

con $B_n = \int_{t_n}^{t_{n+1}} b_s ds$ e $W_n = \int_{t_n}^{t_{n+1}} w_s ds$ è il valore di un movimento Browniano in R^n . In questo caso il filtro ottimale è il filtro di Kalman lineare, le cui equazioni sono:

$$\begin{cases} \hat{X}'_{n+1} = \hat{X}_n + B_n \\ \hat{X}_{n+1} = \hat{X}'_{n+1} + K_{n+1}(Y_{n+1} - H\hat{X}'_{n+1}) \end{cases} \quad (55)$$

La prima equazione è il prediction step e la seconda è il correction step. Sia $e'_n = \hat{X}'_n - X_n$ e $e_n = \hat{X}_n - X_n$ rispettivamente l'errore di stima predetta e l'errore di stima corretta

$$\begin{cases} e'_{n+1} = e_n - W_n \\ e_{n+1} = e'_{n+1} - K_{n+1}(He'_{n+1} - V_{n+1}) \end{cases} \quad (56)$$

L'equazione di errore è autonoma e può essere ottimizzata indipendentemente dagli input B_n . Di conseguenza, il guadagno ottimale K_{n+1} può essere calcolato offline. È un risultato noto che gli errori convergono sotto le ipotesi di osservabilità e detectabilità, così i calcoli possono essere fatti offline e solo il primo valore può essere salvato; in questo caso solo l'equazione (55) deve essere calcolata online così da richiedere solo poca potenza di calcolo a bordo.

Riprendendo la forma del sistema (55) nel gruppo di Lie G sostituendo la somma con il gruppo moltiplicativo si ottengono risultati simili. Sostituendo perciò nel sistema (55) la somma col gruppo moltiplicativo si ottiene una classe di filtri data da:

$$\begin{cases} \hat{\chi}'_{n+1} = \Gamma_n \hat{\chi}_n \Omega_n \\ \hat{\chi}_{n+1} = K_{n+1}(\hat{\chi}'_{n+1} Y_{n+1}) \hat{\chi}'_{n+1} \end{cases} \quad (57)$$

dove K_{n+1} può essere una qualsiasi funzione $Y \implies G$ tale che $K(h(I_d)) = I_d$. Ci sono due modi per capire la connessione tra le equazioni (55) e (57). Per prima cosa definendo un errore di stima nel gruppo G dato da $\chi_{n+1} \hat{\chi}_{n+1}^{-1}$ e usando l'ipotesi dell'equivalenza destra-sinistra, che permette di interpretare $\hat{\chi}'_{n+1} = (\chi_{n+1} \hat{\chi}'_{n+1}^{-1}, V_{n+1})$ come un gruppo equivalente a $H(X'_{n+1} - \hat{X}_{n+1}) + V_{n+1}$. Per seconda prendendo il sistema nel caso lineare (54) come un caso specifico dell'approccio proposto tramite la seguente proposizione

Proposizione: Esiste una rappresentazione isomorfica di R^N come un gruppo matriciale di Lie tale che (54) prende la forma canonica (48), e la (55) diventa il filtro invariante (57).

Si definiscono ora gli errori di output invarianti che sono una trasposizione degli errori lineari al gruppo moltiplicativo

$$\eta_n = \chi_n \hat{\chi}_n^{-1}, \quad \eta'_n = \chi_n \hat{\eta}'_n^{-1}. \quad (58)$$

Si ha anche il seguente teorema:

Teorema: Le variabili errore η_n e η'_n sono processi di Markov, sono quindi indipendenti dagli input Ω_n . (dimostrazione in bibliografia)

La conseguenza più importante è che se gli input sono conosciuti in anticipo, o sono fissati, come nel modello della terra piatta 4.2, la funzione di guadagno K_n può essere ottimizzata offline. SI discute nel prossimo paragrafo di come settare il guadagno.

6 Fissare il guadagno dei filtri

In alcuni casi si può costruire un osservatore quasi globalmente convergente per un sistema deterministico con disturbi nulli, usando una famiglia di funzioni di guadagno costanti $K_n(\cdot) = K(\cdot)$. Se i rumori del filtro sono a zero ha proprietà di dimenticarsi le sue condizioni iniziali, la convergenza ad un singolo punto è impossibile da ottenere a causa dell'imprevedibilità del rumore che "guida" il sistema, ma la convergenza della distribuzione può essere prevista, assumendo che, come nel caso lineare, Γ_n è indipendente da n . Si dimostra però in questo capitolo che, quando il rumore è presente, l'errore dimentica le sue condizioni iniziali sotto alcune condizioni.

6.1 Convergenza

Per ipotesi gli input destri Γ_n sono presunti fissati, il guadagno K_n costante, perciò la sequenza di errori η_n è una catena di Markov omogenea. Sotto determinate condizioni, la catena ha una distribuzione stazionaria unica e la sequenza converge in questa distribuzione invariante. Sia d una distanza destra-invariante nel gruppo G , si propongono le seguenti assunzioni:

1. **Confinamento dell'errore:** esiste un insieme compatto C tale che $\forall n \in N, \eta_n \in C, \forall \eta_0 \in C$.
2. **Diffusibilità:** la sequenza di rumori ha una parte continua (misura di Haar), con densità positiva e uniforme circoscritta in una palla di raggio $\alpha > 0$ intorno a I_d .
3. **Rumore di misurazione:** $\forall g \in G, P[gK(h(g, V_n))^{-1} \in B_0(gK(h(g, 0)))^{-1}, \frac{\alpha}{2}] > \epsilon'$ per qualche $\epsilon' > 0$.

La seconda assunzione implica una più generale assunzione tecnica: $\exists \epsilon > 0$ tale che $\forall U$ sottoinsieme della palla $B_0(I_d, \alpha)$ si ha $P(W_n \subset U) > \epsilon \mu(U) \forall n \geq 0$ dove μ denota la misura di Haar. Le altre due assunzioni garantiscono che il processo degli errori è ben approssimato dalla sua dinamica senza rumori, seguita da un piccola diffusione. Nella teoria della catena di Harris, lo step di diffusione è un elemento chiave per permettere alle leggi di probabilità di mischiarsi ad ogni step e eventualmente dimenticare la loro distribuzione iniziale.

Teorema: Per input destri costanti $\Gamma_n = \Gamma$, si considera il filtro:

$$\begin{cases} \hat{\chi}'_{n+1} = \Gamma \hat{\chi}_n \Omega \\ \hat{\chi}_{n+1} = K(\hat{\chi}'_{n+1} Y_{n+1}) \hat{\chi}'_{n+1} \end{cases} \quad (59)$$

Si suppone che le assunzioni (1)(2)(3) sono verificate, dove l'insieme compatto soddisfa $C = cl(C^0)$. Quando gli errori sono nulli, l'equazione di errore è:

$$\begin{cases} \gamma'_{n+1} = \Gamma \gamma_n \Gamma \\ \gamma_{n+1} = \gamma'_{n+1} K_{n+1}(h(\gamma'_{n+1}, 0))^{-1} \end{cases} \quad (60)$$

si suppone che $\forall \gamma_0 \in C$, eccetto per un insieme di misure di Haar nullo, γ_n converge a I_d .

Esiste una distribuzione stazionaria π in G tale che per ogni legge precedente μ_0 dell'errore η_0 in C , la legge $(\mu_n)_{n>0}$ di $(\eta_n)_{n>0}$ soddisfa la proprietà di convergenza della norma di totale variazione:

$$\lim_n \|\mu_n - \pi\|_{T.V.} \rightarrow 0 \quad (61)$$

Corollario: Quando il gruppo G è compatto, la convergenza del teorema precedente è globale, ovvero senza la restrizione dell'assunzione (1).

Teorema: Sotto l'ipotesi del teorema precedente, assumendo solo $h(\gamma_n, 0) \rightarrow 0$ invece della convergenza globale di $(\eta_n)_{n>0}$ (G è compatto), l'insieme $K = \{g \in G, h(g, 0) = h(I_d, 0)\}$ connesso e $h(\Gamma, 0) = h(I_d, 0)$, i risultati del teorema sono ancora validi. Inoltre, se W_n è isotropico, si ha $\pi(\tilde{\Gamma}V) = \pi(V) \forall \tilde{\Gamma} \in K$ che commuta con $\Gamma(\tilde{\Gamma}\Gamma = \tilde{\Gamma})$.

6.2 Applicazione alla stima dell'assetto

Si considera come modello l'esempio del paragrafo 4.2. A tempo-continuo sono numerosi gli osservatori che convergono quasi globalmente, per applicare queste teorie in questo esempio bisogna considerare due aspetti: prima cosa l'osservatore deterministico deve essere adattato a tempo discreto e dimostrato essere quasi globalmente convergente, poi il filtro deve soddisfare tutti i teoremi e proprietà precedenti in presenza di rumori.

A tempo-discreto il sistema è:

$$\begin{cases} R_{n+1} = W_n R_n \Omega_n \\ Y_n = (R_n^T b^1 + V_n^1, R_n^T b^2 + V_n^2) \end{cases} \quad (62)$$

Si propone il filtro nel gruppo $SO(3)$:

$$\begin{cases} \hat{R}'_{n+1} = \hat{R}_n \Omega_n \\ \hat{R}_{n+1} = K(\hat{R}'_{n+1} Y_{n+1}) \hat{R}'_{n+1} \end{cases} \quad (63)$$

con $K(y_1, y_2) = \exp(k_1(y_1 \times b_1) + k_2(y_2 \times b_2))$, $k_1 > 0$, $k_2 > 0$, $k_1 + k_2 \leq 1$.

Proposizione: Senza la presenza di rumori, l'osservatore invariante discreto (63) è quasi globalmente convergente, quindi l'errore converge a I_d per ogni condizione iniziale eccetto una.

Proposizione: Si considera un gruppo di Lie G con una metrica $\langle \cdot, \cdot \rangle$, e una equazione deterministica discreta sinistra invariante in G della forma

$$\gamma_{n+1} = \gamma_n \exp(-k(\gamma_n)),$$

per ipotesi esiste una funzione $E \in C^2$ tale che $E : G \rightarrow R_{\geq 0}$, un minimo globale I_d , e una funzione continua strettamente positiva $u : G \rightarrow R_{> 0}$ tale che: $\forall x \in G, k(x) = u(x)[x^{-1} \cdot \text{grad}_E(x)]$. Se la condizione $\forall x \in G, \left| \frac{\partial k}{\partial x} \right| \leq 1$ è verificata, $\forall \gamma_0$ condizione iniziale tale che I_d è il solo punto critico di E nel sottoinsieme $x \in G \mid E(x) \leq E(\gamma_0)$ si ha:

$$\gamma_n \xrightarrow[n \rightarrow \infty]{} I_d$$

Usando il teorema e la preposizione precedente si ottiene:

Teorema: La distribuzione della variabile d'errore del filtro invariante (63) converge per la norma di totale variazione ad una distribuzione asintotica, che non dipende dalle sue condizioni iniziali.

6.3 Applicazione per definire un orizzonte artificiale

Come dimostrato precedentemente, sotto appropriate condizioni, la variabile d'errore è una catena di Markov convergente, la cui legge asintotica dipende dalla funzione di guadagno ma non dalla traiettoria del sistema, quindi il guadagno può essere calcolato offline.

Un problema classico è l'impostazione di un orizzonte artificiale usando l'unità di misura inerziale IMU. Un'approssimazione dell'asse verticale è data dall'accelerometro e dall'ipotesi che il volo sia stazionario, quindi la sua velocità lineare è costante. Il problema è che nelle fasi dinamiche questa approssimazione non funziona, che è proprio dove è più utile. Il problema è così impostato:

$$\begin{cases} \frac{d}{dt}R_t = R_t(\omega_t + w_t)_\times \\ Y_n = R_{t_n}^T g + V_n + N_n \end{cases} \quad (64)$$

dove R_t è l'assetto del velivolo, ovvero la rotazione tra il sistema di riferimento del velivolo stesso e quello inerziale, ω_t è l'incremento continuo del giroscopio e Y_n è la misurazione dell'accelerometro. I rumori dei sensori w_t e V_n possono essere considerati Gaussiani e N_n rappresentano fluttuazioni date dall'accelerazione del velivolo che si presuppone essere nulle, ma se così non fosse potrebbero prendere valori molto alti.

6.3.1 Risultati

Si considera la seguente classe di filtri:

$$\begin{cases} \hat{R}'_{n+1} = \hat{R}_n \Omega_n \\ \hat{R}_{n+1} = K(\hat{R}'_{n+1} Y_{n+1}) \hat{R}'_{n+1} \end{cases} \quad (65)$$

con $K(y) = \exp(f_{k,\lambda}(y))$ dove $f_{k,\lambda} = k \cdot \min(\text{angle}(x, g), \lambda) \frac{x \times g}{\|x \times g\|}$ se $x \times g \neq 0$ e $f_{k,\lambda}(x) = 0$ altrimenti.

Se l'accelerometro misura un valore y si considera la più piccola rotazione dando a g la stessa direzione di y . Con gli stessi assi, l'angolo di questa rotazione è limitato per dare meno peso ai outlier. Si sceglie poi come funzione di guadagno una rotazione di una frazione dell'angolo ottenuto. Si hanno i seguenti risultati:

Lemma: $\forall k, 0 < k \leq 1$ e $0 < \lambda \leq \pi$ l'errore di output $\|Y_n - \hat{R}_n^T g\|$ associato all'osservatore (65) senza rumori converge a 0.

La seguente proposizione è il risultato del lemma e del teorema precedente:

Proposizione: La variabile di errore associata al filtro tramite (65) converge ad una distribuzione stazionaria per la norma di totale variazione, che non dipende dalla sua distribuzione iniziale.

7 Filtri Gaussiani

I due prossimi filtri hanno il merito di dare un'approssimazione dell'incertezza portata dalle stime. Per primo si introduce il filtro di Kalman esteso invariante tempo-discreto IEKF, che è un EKF ma basato sulla linearizzazione dell'errore invariante dello stato. Poi si introduce il IEnKF che calcola la matrice delle covarianze empirica dell'errore usando un frammento della densità di errore. Siccome tutti i calcoli sono fatti offline, è di facile implementazione finché è possibile salvare i guadagni, e è mostrato come dia una migliore approssimazione della dispersione dell'errore per rumori grandi.

7.1 Filtro di Kalman Esteso Invariante tempo-discreto (IEKF)

Come nella teoria standard dell'EKF, l'idea è di linearizzare l'equazione d'errore assumendo i rumori e l'errore dello stato piccoli abbastanza, usando le equazioni di Kalman per registrare il guadagno in questo sistema lineare e implementando il guadagno in quello non lineare. Il guadagno e la matrice delle covarianze dell'errore convergono ad un valore finito, ciò fa sì che si sostituisca il guadagno con la costante fissata portando ad una versione matematicamente leggera e asintotica del IEKF.

7.1.1 Linearizzazione delle equazioni del IEKF

Si assume che errori e rumori siano piccoli e si introduce la loro proiezione nello spazio tangente, ovvero nell'algebra di Lie tramite la mappa esponenziale matriciale $\exp(\cdot)$. Siccome tale spazio vettoriale matriciale può essere identificato con $R^{\dim(g)}$ tramite la mappa lineare $(\cdot)_m$ si assume $\exp : R^{\dim(g)} \mapsto G$. Si definiscono ora le seguenti quantità in $R^{\dim(g)}$:

$$\xi_n = \exp^{-1}(\eta_n), \quad \xi'_n = \exp^{-1}(\eta'_n), \quad w_n = \exp^{-1}(W_n) \quad (66)$$

La via più semplice per impostare il guadagno è di usare una funzione che è lineare in $R^{\dim g}$ e poi mappala a G tramite l'esponenziale:

$$K_n : y \mapsto \exp[L_n(y - h(I_d, 0))] \quad (67)$$

Le equazioni di errore quindi mappate a $R^{\dim g}$ diventano:

$$\begin{cases} \xi'_{n+1} = \exp^{-1}(\exp(Ad_{\Gamma_n} w_n) \exp(Ad_{\Gamma_n} \xi_n)) \\ \xi_{n+1} = \exp^{-1}(\exp(\xi'_{n+1}) \exp[-L_n(h(\exp(\xi'_{n+1}), V_{n+1}) - h(I_d, 0))]) \end{cases} \quad (68)$$

dove si introduce la matrice aggiunta, ovvero la trasposta della matrice dei co-fattori della matrice stessa:

Definizione: La seguente proprietà può essere usata come definizione:

$$\forall g \in G, \forall u \in R^{\dim(\mathfrak{g})}, \exp(Ad_g u) = g \exp(u) g^{-1} \quad (69)$$

il differenziale di $\xi \mapsto Ad_{\exp(\xi)}$ a zero è indicato $\xi \mapsto ad_\xi$.

Per settare il guadagno L_n tramite la teoria di Kalman bisogna valutare ad

ogni istante di tempo, usando $P_n = Var(\xi_n) \in R^{dim\ g \times dim\ g}$, le seguenti quantità:

$$P_{n+1|n} = Var(\xi'_{n+1}), \quad S_{n+1|n} = Var(Y_{n+1}) \quad (70)$$

L'approccio consiste nell'usare uno sviluppo al primo ordine delle funzioni di propagazione e misurazione. Per farlo si usa la formula di Baker-Campbell-Hausdorff e tenere solo i termini di primo ordine di ξ'_{n+1} , w_n e V_{n+1} :

$$\begin{aligned} \xi'_{n+1} &= Ad_{\Gamma_{n_e}} w_n + Ad_{\Gamma_n} \xi_n \\ h(exp(\xi'_{n+1}), V_{n+1}) &= h(I_d, 0) + H_\xi \xi'_{n+1} + H_V V_{n+1} \\ \xi_{n+1} &= \xi'_{n+1} - L_{n+1}(H_\xi \xi'_{n+1} + H_V V_{n+1}) \end{aligned}$$

L'algoritmo del filtro è descritto da Algoritmo 1 a fine capitolo.

7.1.2 Convergenza del guadagno

Sotto determinate condizioni la matrice delle covarianze P_n e il guadagno K_n convergono a valori costanti. Le conseguenze pratiche di ciò sono principalmente due: per prima cosa la covarianza dell'errore converge a un valore fissato, perciò è più semplice da interpretare piuttosto che una matrice che varia in continuazione, poi dal fatto che a bordo del veicolo ci sono limitazioni computazionali, infatti l'approssimazione della matrice delle covarianze al suo valore asintotico porta ad una versione asintotica del IEKF che è numericamente poco onerosa.

Teorema: Se le matrici di errore $Q_n^w = Var(w_n)$, $Q_N^V = Var(V_n)$ e gli input sinistri Γ_n sono fissati, la coppia (Ad_{Γ_n}, H_ξ) è osservabile e Q_n^w h rango pieno, P_n e K_n convergono ad un valore fissato P_∞ e K_∞ come nel caso lineare temo-invariante.

7.2 EnKF Invariante tempo discreto

Quando l'equazione di errore è indipendente dai sui ingressi, la densità di probabilità esatta della variabile d'errore può essere prevista offline. Questo permette di migliorare drasticamente la precisione delle quantità coinvolte nel calcolo del guadagno di Kalman della sezione 6.1.1. Si propone il IEnKF, filtro di Kalman "Completo" Invariante; l'idea è di calcolare ricorsivamente tramite la simulazione di Monte-Carlo un esempio della densità di probabilità dell'errore e di usarla, invece che nella linearizzazione, per valutare precisamente l'innovazione e la matrice delle covarianze dell'errore usate per calcolare il guadagno del filtro. L'algoritmo del filtro è descritto da Algoritmo 2 a fine capitolo.

7.3 Simulazione

Ora si applicano i risultati degli Algoritmi 1 e 2 per il problema della stima dell'assetto descritta nella sezione 4.2:

$$\begin{cases} \frac{d}{dt} R_t = R_t(\omega_t + w_t^w) \\ Y_n = (R_n^T b_1 + V_n^1, R_n^T b_2 + V_n^2) \end{cases} \quad (71)$$

dove $R_t \in SO(3)$ rappresenta la rotazione che fa passare dal sistema di riferimento del velivolo a quello della terra, ω_t è il vettore di rotazione angolare

istantanea che è misurato dal giroscopio, $(Y_n)_{n \geq 0}$ è una sequenza di misurazioni disturbate dei vettori $b_1 b_2$, vettori campo gravitazionale e campo magnetico, $V_n^1 V_n^2$ sono sequenze di rumori bianchi isotropici indipendenti. Per eseguire la simulazione si usano gli Algoritmi 1 e 2, confrontando l'evoluzione degli errori e dei guadagni dei rispettivi filtri tramite dei grafici.

Per produrre un grafico si usano delle matrici di grandi dimensioni, che idealmente vengono divise in tante parti quante sono gli istanti in cui si vuole monitorare un parametro, per poi plottare per ogni istante una parte di quella matrice, così da avere il grafico che rappresenta l'andamento del parametro all'evoluzione del filtro.

Per eseguire l'esperimento si prendono come valori $b_1 = (-\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$, $b_2 = (0, 0, 1)$ ed altri descritti della Tabella sottostante.

Steps	Particelle	Simulazioni	P_0	R_1	R_2	Q
50	1000	200	$(\frac{30*\pi}{180})^2 * I_3$	$(\frac{10*\pi}{180})^2 * I_3$	$(\frac{1*\pi}{180})^2 * I_3$	$(\frac{0.1*\pi}{180})^2 * I_3$

Alle matrici elencate nella tabella, va applicata una fattorizzazione di Cholesky[5], ovvero si utilizza solamente la parte triangolare superiore della matrice, dove la parte triangolare inferiore è considerata il suo complesso coniugato trasposto. Quindi se la matrice è definita positiva:

$$R = CHOL(A) \implies R' \times R = A \quad (72)$$

La funzione $exp(\cdot)$ nei due algoritmi, è la mappa esponenziale che fa passare dal gruppo $SO(3)$ al suo spazio tangente, ovvero algebra di Lie, tramite la formula:

$$exp(x) = I_3 + \frac{\sin(\|x\|)}{\|x\|}(x)_\times + \frac{1}{\|x\|^2}(1 - \cos \|x\|)(x)_\times^2. \quad (73)$$

Algorithm 1 IEKF: Filtro di Kalman Estesio Invariante tempo-discreto in SO(3)

$$H_\xi = \begin{pmatrix} (b_1)_\times \\ (b_2)_\times \end{pmatrix}$$

$$Q_V = \begin{bmatrix} \text{Var}(V^1) & 0 \\ 0 & \text{Var}(V^2) \end{bmatrix}$$

$$Q_w = \int_{t_n}^{t_{n+1}} \text{Var}(w_s^\omega) ds$$

$$n = 0$$

loop

Sia Ω_n la soluzione a t_{n+1} del sistema $T_{t_n} = I_3$, $\frac{d}{dt}T_t = (\omega_t)_\times$

$$P_{n+1|n} = P_n + Q_w$$

$$S_{n+1|n} = H_V Q_V H_V^T + H_\xi P_{n+1|n} H_\xi^T$$

$$L_{n+1} = P_{n+1|n} H_\xi^T S_{n+1}^{-1}$$

$$P_{n+1} = (I - L_{n+1} H_\xi) P_{n+1|n}$$

$$\hat{R}'_{n+1} = \hat{R}_n \Omega_n$$

$$\nabla_\chi = L_{n+1} \left[\begin{pmatrix} \hat{R}'_{n+1} Y_{n+1}^1 \\ \hat{R}'_{n+1} Y_{n+1}^2 \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right]$$

$$\hat{R}_{n+1} = \exp(\nabla_\chi) \hat{R}'_{n+1}$$

$$n = n + 1$$

end loop

Algorithm 2 IEnKF: Filtro di Kalman Completo Invariante tempo-discreto

for $n = 0 : N - 1$ **do**

for $i = 1 : M$ **do**

$\eta_{n+1}^i = W_n^i \Gamma \eta_n^i \Gamma^{-1}$

$y_{n+1}^i = h(\eta_{n+1}^i, V_{n+1}^i)$

end for

$P_{n+1|n} = \frac{1}{M} \sum : i = 1^M \exp^{-1}(\eta_{n+1}^i) \exp^{-1}(\eta_{n+1}^i)^T$

$S_{n+1} = \frac{1}{M} \sum : i = 1^M y_{n+1}^i (y_{n+1}^i)^T$

$L_{n+1|n} = P_{n+1|n} H^T S_{n+1}^{-1}$

for $i = 1 : M$ **do**

$\eta_{n+1}^i = \eta_{n+1}^i \exp(-L_{n+1|n} [y_{n+1}^i - h(I_d, 0)])$

end for

 Salvare la matrice L_{n+1}

end for

for $n = 0 : N - 1$ **do**

$\hat{\chi}'_{n+1} = \Gamma \hat{\chi}_n \Omega_n$

$\hat{\chi}_{n+1} = \exp(L_{n+1} [\hat{\chi}'_{n+1} Y_{n+1} - h(I_d, 0)]) \hat{\chi}'_{n+1}$

end for

8 Conclusioni

In questo documento si è quindi introdotto un approccio probabilistico per il problema di filtrare un sistema con rumori in uno spazio matriciale di un gruppo di Lie. Si è proposta una classe di filtri tempo-discreto con dei metodi per fissare il guadagno. Quando il guadagno è fissato, l'errore si è visto avere proprietà di convergenza nella sua distribuzione.

Inoltre, si è anche fatta una variante in cui ci si concentra sulla fase transitoria, dove si cercava di minimizzare la dispersione della stima dell'errore ad ogni step tramite la linearizzazione e l'approssimazione dei rumori Gaussiani. Questo ha portato ad una versione Invariante del filtro di Kalman Esteso (IEKF), la cui matrice delle covarianze dell'errore converge asintoticamente, ed ad una versione Invariante del EnKF, che è stato mostrato convergere ad una accurata indicazione del livello di incertezza portata dalla stima.

L'esperimento ha prodotto i risultati riportati nei grafici, i quali riportano il valore dell'errore, uno per ogni asse del sistema di riferimento del velivolo, e il valore del guadagno per ogni istante di tempo. Quindi le proprietà di entrambi i filtri proposti, IEKF e IEnKF, sono state dimostrate anche nella pratica tramite l'esperimento svolto poiché sia l'errore sia il guadagno convergono ad un valore stazionario o allo zero.

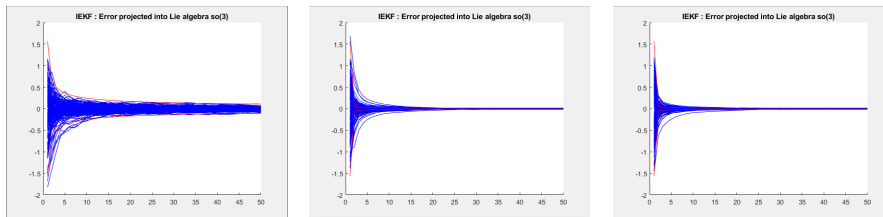


Figura 5: Grafici degli errori per ogni asse del sistema di riferimento del velivolo per il IEKF.

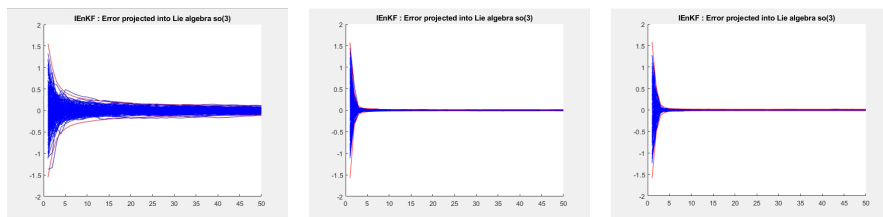


Figura 6: Grafici degli errori per ogni asse del sistema di riferimento del velivolo per il IEnKF.

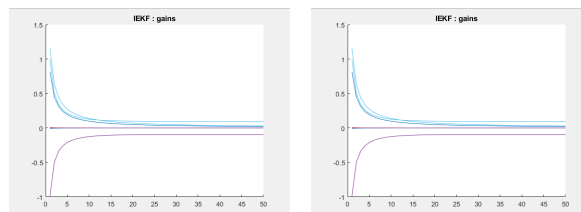


Figura 7: Grafici dei coefficienti della matrice dei guadagni dei due filtri (rispettivamente IEKF e IEnKF)

References

- [1] A. Barrau and S. Bonnabel, “Intrinsic filtering on lie groups with applications to attitude estimation,” *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 436–449, 2015.
- [2] M. I. Ribeiro, “Kalman and extended kalman filters: Concept, derivation and properties,” *Institute for Systems and Robotics*, vol. 43, 2004.
- [3] W. Ziller, “Lie groups. representation theory and symmetric spaces,” 2010.
- [4] A. W. Knap, *Lie groups beyond an introduction*. Springer Science & Business Media, 2013, vol. 140.
- [5] S. J. Julier and J. K. Uhlmann, “New extension of the kalman filter to non-linear systems,” in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068. International Society for Optics and Photonics, 1997, pp. 182–193.