

Filtro di Kalman

Maria S. Greco

Corso di Fondamenti di Radar

Ing. delle Telecomunicazioni

Novembre 2014

Modello del segnale

$$s(n) = a(n)s(n-1) + w(n)$$

Il filtro di Kalman funziona anche con processi non stazionari, cioè se il coefficiente $a(n)$ varia con il tempo. Per semplicità supponiamo che il segnale osservato sia stazionario, quindi $a(n)=a$

$$s(n) = as(n-1) + w(n)$$

$$E\{s(n)\} = 0 \quad \sigma_s^2 = \frac{\sigma_w^2}{1-a^2}$$

$w(n)$ rumore bianco di generazione. Se è Gaussiano, anche $s(n)$ risulta essere Gaussiano.

Modello dell'osservato

$$x(n) = cs(n) + v(n)$$

c è una costante che dipende dal sistema di osservazione, $v(n)$ rumore bianco di osservazione a valor medio nullo e varianza σ_v^2

Per implementare il filtro di Wiener si richiede che $x(n)$ e $s(n)$ siano congiuntamente stazionari. Tale ipotesi non è necessaria per il filtro di Kalman.

Se sia $v(n)$ che $w(n)$ sono Gaussiani, il filtro di Kalman lineare che otteniamo è il filtro ottimo (tra tutti), nel senso che minimizza l'errore quadratico medio

$$\varepsilon^2 = E\left\{(\hat{s}(n) - s(n))^2\right\}$$

altrimenti è solo il filtro ottimo tra i lineari.

Filtro di Kalman scalare

Dimostriamo che la stima lineare ottima in media quadratica è data da:

$$\hat{s}(n) = \alpha(n)\hat{s}(n-1) + k(n)x(n)$$

Ciò equivale a calcolare le espressioni dei coefficienti $a(n)$ e $k(n)$ per cui l'errore

$$\varepsilon(n) = s(n) - \hat{s}(n)$$

è ortogonale ai dati $x(j)$, $j \leq n$ (Principio di ortogonalità)

Filtro di Kalman scalare

Applichiamo la condizione di ortogonalità tra $x(n)$ e $\varepsilon(n)$ cioè:

$$E\{\varepsilon(n)x(n)\} = 0$$

Sostituiamo al posto di $x(n)$ la sua espressione, per cui

$$\begin{aligned} E\{\varepsilon(n)x(n)\} &= E\{\varepsilon(n)[cs(n) + v(n)]\} = E\{\varepsilon(n)[c\hat{s}(n) + c\varepsilon(n) + v(n)]\} \\ &= cP(n) + E\{\varepsilon(n)v(n)\} = 0 \end{aligned}$$

Dove abbiamo posto $P(n) = E\{\varepsilon^2(n)\}$, sostituito $s(n) = \hat{s}(n) + \varepsilon(n)$ e osservato che, poiché $\hat{s}(n)$ dipende linearmente dai dati, $E\{\varepsilon(n)\hat{s}(n)\} = 0$

Dall'espressione precedente otteniamo per ora che

$$P(n) = -\frac{E\{\varepsilon(n)v(n)\}}{c}$$

Riscriviamo $\varepsilon(n)$

$$\varepsilon(n) = as(n-1) + w(n) - \alpha(n)\hat{s}(n-1) - k(n)x(n)$$

e sostituiamolo nel calcolo del valor medio $E\{\varepsilon(n)v(n)\}$

I primi due termini sono indipendenti da $v(n)$ quindi il loro contributo è nullo. Anche

$$E\{\hat{s}(n-1)v(n)\} = 0$$

poiché $\hat{s}(n-1)$ dipende solo dagli osservati (e quindi dal rumore) fino al passo $n-1$. Quindi

$$P(n) = -\frac{E\{\varepsilon(n)v(n)\}}{c} = k(n) \frac{E\{[cs(n) + v(n)]v(n)\}}{c} = k(n) \frac{\sigma_v^2}{c}$$

da cui ricaviamo che

$$k(n) = \frac{cP(n)}{\sigma_v^2}$$

Consideriamo ora le altre condizioni di ortogonalità tra $x(j)$ e $\varepsilon(n)$

$$E\{\varepsilon(n)x(j)\} = 0 = E\{[s(n) - \hat{s}(n)]x(j)\} \text{ con } j < n$$

Con calcoli simili ai precedenti, possiamo scrivere

$$E\{[s(n) - \alpha(n)s(n-1) + \alpha(n)\varepsilon(n-1) - k(n)x(n)]x(j)\} = 0$$

$$R_{xs}(n-j) - \alpha(n)R_{xs}(n-1-j) = k(n)R_x(n-j) \text{ con } j < n$$

$$R_{xs}(n-j) = E\{s(n)x(j)\} = E\{s(n)[cs(j) + v(j)]\} = cR_s(n-j)$$

$$R_x(n-j) = E\{[cs(n) + v(n)][cs(j) + v(j)]\} = c^2R_s(n-j)$$

$$R_s(n-j) = \alpha(n)R_s(n-1-j) + ck(n)R_s(n-j)$$

Dal modello AR(1) si sa che

$$R_s(n-j) = aR_s(n-1-j)$$

Sostituiamo nell'equazione precedente, si ottiene

$$\alpha(n) = a[1 - c k(n)]$$

Sostituiamo questo risultato in $\hat{s}(n) = \alpha(n)\hat{s}(n-1) + k(n)x(n)$

$$\hat{s}(n) = a\hat{s}(n-1) + k(n)[x(n) - c\hat{s}(n|n-1)]$$

dove $\hat{s}(n|n-1) = a\hat{s}(n-1)$ (predizione ad un passo)

mentre $c\hat{s}(n|n-1) = \hat{x}(n|n-1)$ (predizione dell'osservato)

Torniamo indietro al calcolo di $P(n) = E\{\varepsilon^2(n)\}$

$$P(n) = E\left\{[s(n) - \hat{s}(n)]^2\right\} = E\left\{[s(n) - a\hat{s}(n-1) - k(n)[x(n) - c a \hat{s}(n-1)]]^2\right\}$$

Utilizzando il modello del segnale e dell'osservato

$$\begin{aligned} P(n) &= E\left\{[as(n-1) + w(n) - a\hat{s}(n-1) - k(n)[ca s(n-1) + v(n) + cw(n) - ca \hat{s}(n-1)]]^2\right\} \\ &= E\left\{[a\varepsilon(n-1) + w(n) - k(n)[ca\varepsilon(n-1) + v(n) + cw(n)]]^2\right\} \\ &= a^2 P(n-1) + \sigma_w^2 + k^2(n)c^2 a^2 P(n-1) + k^2(n)\sigma_v^2 + c^2 k^2(n)\sigma_w^2 \\ &\quad - 2a^2 k(n)c P(n-1) - 2k(n)c\sigma_w^2 \end{aligned}$$

Compattando e sostituendo l'espressione di $k(n)$, si ottiene

$$P(n) = a^2 [1 - k(n)c]^2 P(n-1) + k^2(n)\sigma_v^2 + [1 - k(n)c]^2 \sigma_w^2$$



$$P(n) = [1 - k(n)c] \left[a^2 P(n-1) + \sigma_w^2 \right]$$

$$k(n) = \frac{c \left[a^2 P(n-1) + \sigma_w^2 \right]}{\sigma_v^2 + c^2 \left[a^2 P(n-1) + \sigma_w^2 \right]}$$

Per completare la derivazione dobbiamo calcolare l'espressione dell'errore quadratico di predizione ad un passo

$$\begin{aligned} P(n|n-1) &= E\left\{ [s(n) - \hat{s}(n|n-1)]^2 \right\} = E\left\{ [s(n) - a\hat{s}(n-1)]^2 \right\} \\ &= E\left\{ [as(n-1) + w(n) - a\hat{s}(n-1)]^2 \right\} = \left[a^2 P(n-1) + \sigma_w^2 \right] \end{aligned}$$

Da cui le nuove espressioni di $k(n)$ e $P(n)$

$$k(n) = \frac{cP(n|n-1)}{\sigma_v^2 + c^2 P(n|n-1)} \quad P(n) = [1 - k(n)c] P(n|n-1)$$

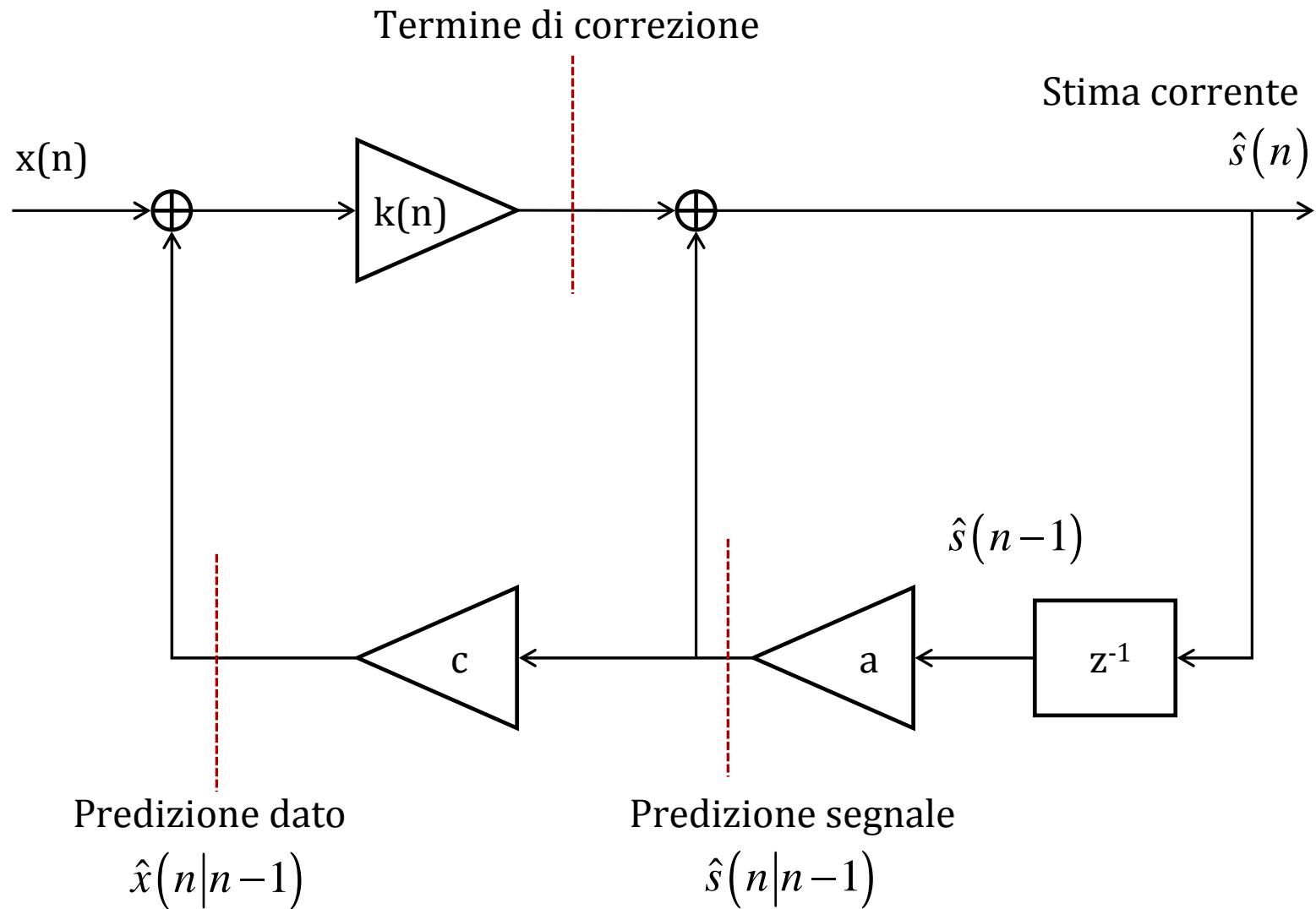
Riassumiamo:

Stimatore: $\hat{s}(n) = a\hat{s}(n-1) + k(n)[x(n) - c\hat{s}(n-1)]$

Guadagno filtro: $k(n) = \frac{cP(n|n-1)}{\sigma_v^2 + c^2 P(n|n-1)}$

Errore quadratico medio: $P(n) = [1 - k(n)c]P(n|n-1)$

con $P(n|n-1) = a^2 P(n-1) + \sigma_w^2$



Il filtraggio si ottiene in modo ricorsivo. Si suppongono noti $\hat{s}(n-1)$ e $P(n-1)$

Step 1: $\hat{s}(n|n-1) = a\hat{s}(n-1)$ Predizione ad un passo

Step 2: $\hat{x}(n|n-1) = c\hat{s}(n|n-1)$ Predizione del dato

Step 3: $\Delta(n) = x(n) - \hat{x}(n|n-1)$ Scarto al passo n

Step 4: $P(n|n-1) = a^2 P(n-1) + \sigma_w^2$ $k(n) = \frac{cP(n|n-1)}{\sigma_v^2 + c^2 P(n|n-1)}$

Step 5: $\hat{s}(n) = \hat{s}(n|n-1) + k(n)\Delta(n)$ Nuova stima

Step 6: $P(n) = [1 - k(n)c]P(n|n-1)$ Errore quadratico medio

Inizializzazione del filtro

$$\hat{s}(-1) = E\{s(-1)\} = 0 \quad \text{Stima ottima in assenza di osservazioni}$$

$$P(-1) = E\left\{\left[s(-1) - \hat{s}(-1)\right]^2\right\} = \sigma_s^2$$

$$\text{Step 1: } \hat{s}(0|-1) = a\hat{s}(-1) = 0$$

$$\text{Step 2: } \hat{x}(0|-1) = c\hat{s}(0|-1) = 0$$

$$\text{Step 3: } \Delta(0) = x(0) - \hat{x}(0|-1) = x(0)$$

$$\text{Step 4: } P(0|-1) = a^2 P(-1) + \sigma_w^2 = a^2 \sigma_s^2 + \sigma_w^2 \quad k(n) = \frac{cP(0|-1)}{\sigma_v^2 + c^2 P(0|-1)}$$

$$\text{Step 5: } \hat{s}(0) = k(0)x(0)$$

$$\text{Step 6: } P(0) = [1 - k(0)c]P(0|-1)$$

Osservazioni:

- Se, come supposto nella derivazione, $x(n)$ e $s(n)$ sono congiuntamente stazionari, il filtro di Kalman a regime si riduce a quello di Wiener, ma è di più semplice implementazione perché ricorsivo.
- Se il filtro è stazionario, i pesi $k(n)$, gli errori $P(n|n-1)$ e $P(n)$ possono essere precalcolati.

Filtro di Kalman vettoriale

- Il filtro di Kalman vettoriale è necessario per estendere i risultati a modelli di segnale ARMA e alla stima di segnali multidimensionali come nel caso radar.
- Il filtro vettoriale viene usato nel radar per il tracking, cioè per stimare la traiettoria ed i parametri cinematici del target (posizione, velocità e possibile accelerazione) a partire da misure di distanza, angolo di azimuth ed, eventualmente, anche elevazione
- La posizione e la velocità stimate vengono usate per la predizione della misura successiva e, nel caso di phased array radar, per orientare il fascio dell'antenna.

Modello dinamico o equazione di stato

$$\mathbf{s}_n = \mathbf{F}_n \mathbf{s}_{n-1} + \mathbf{G}_n \mathbf{w}_n$$

\mathbf{F}_n è la matrice che definisce il modello cinematico del target tra l'istante t_n e t_{n+1} .

\mathbf{G}_n è la matrice che lega gli errori del sistema allo stato del target all'istante n.

\mathbf{w}_n è il rumore di generazione. Si suppone Gaussiano, $\mathbf{w}_n \in N(\mathbf{0}, \mathbf{Q}_n)$

Modello dell'osservato

$$\mathbf{z}_n = \mathbf{H}_n \mathbf{s}_n + \mathbf{v}_n$$

\mathbf{H}_n è la matrice che lega l'osservato allo stato del target

\mathbf{v}_n è il rumore di osservazione, $\mathbf{v}_n \in N(\mathbf{0}, \mathbf{R}_n)$

Riscriviamo tutti gli step nel caso vettoriale stazionario

Step 1: $\hat{\mathbf{s}}_{n|n-1} = \mathbf{F}\hat{\mathbf{s}}_{n-1}$ Predizione ad un passo

Step 2: $\hat{\mathbf{z}}_{n|n-1} = \mathbf{H}\hat{\mathbf{s}}_{n|n-1}$ Predizione del dato

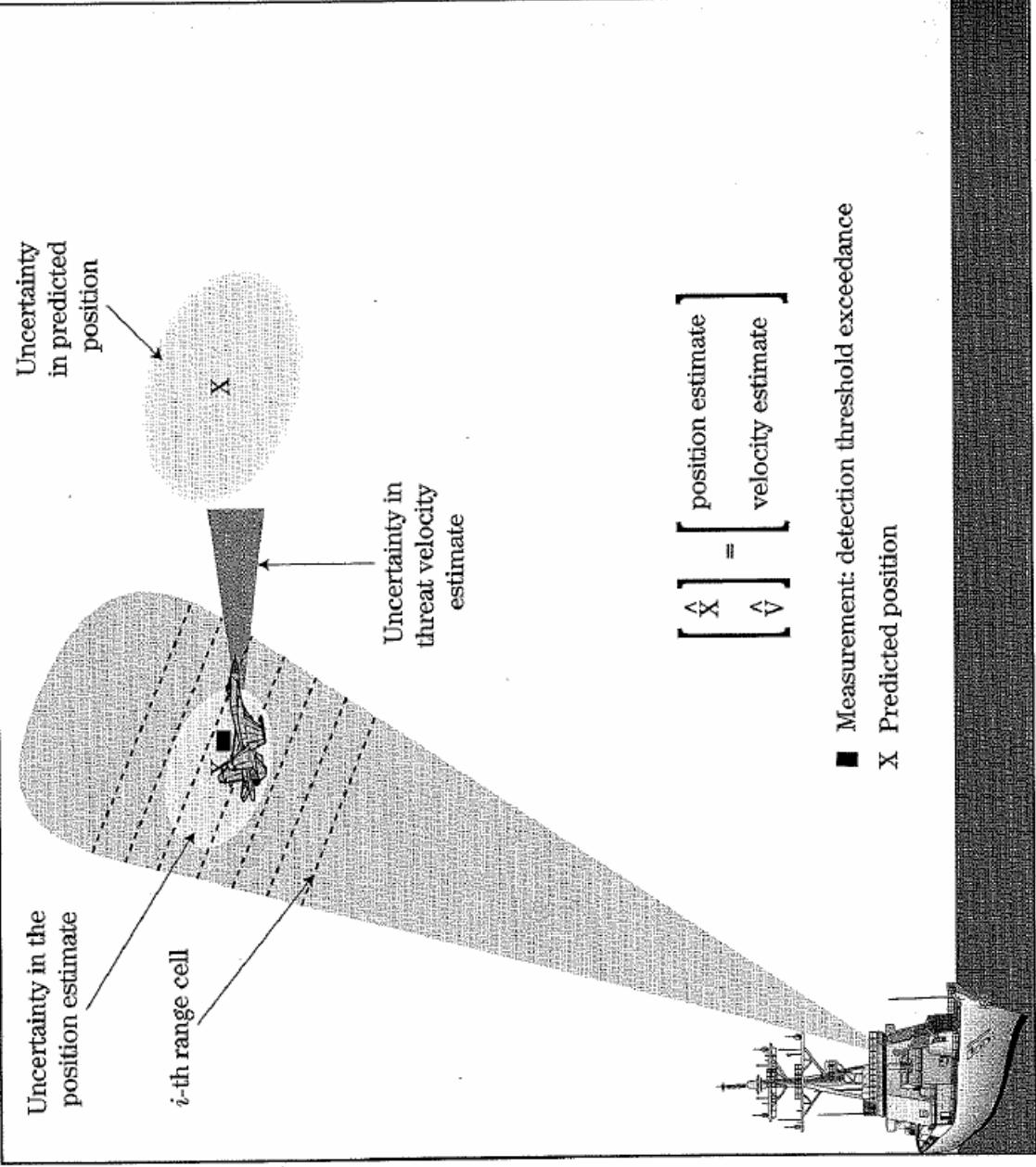
Step 3: $\Delta_n = \mathbf{z}_n - \hat{\mathbf{z}}_{n|n-1}$ Scarto al passo n

Step 4: $\mathbf{P}_{n|n-1} = \mathbf{F}\mathbf{P}_{n-1}\mathbf{F}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T$ Errore di predizione

$$\mathbf{k}_n = \mathbf{P}_{n|n-1} \mathbf{H}^T \mathbf{M}_n^{-1} \quad \mathbf{M}_n = \mathbf{H}\mathbf{P}_{n|n-1}\mathbf{H}^T + \mathbf{R}$$

Step 5: $\hat{\mathbf{s}}_n = \hat{\mathbf{s}}_{n|n-1} + \mathbf{k}_n \Delta_n$ Nuova stima

Step 6: $\mathbf{P}_n = [\mathbf{I} - \mathbf{k}_n \mathbf{H}] \mathbf{P}_{n|n-1}$ Errore quadratico medio



Modello a velocità quasi-costante

Consideriamo il caso semplicissimo a singola coordinata x. **Le equazioni del moto** sono:

$$\begin{cases} x_{n+1} = x_n + T \dot{x}_n + \frac{T^2}{2} w_n \\ \dot{x}_{n+1} = \dot{x}_n + T w_n \end{cases}$$

quindi:

$$\mathbf{s}_n = \begin{bmatrix} x_n \\ \dot{x}_n \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} T^2/2 \\ T \end{bmatrix} \quad \mathbf{Q} = \sigma_w^2$$

Il moto è detto a velocità quasi-costante perché il termine di rumore w_n introduce una piccola accelerazione casuale.

Il **modello delle misure** è $z_n = x_n + v_n$

quindi: $\mathbf{H} = [1 \ 0]$ $\mathbf{R} = \sigma_v^2$

A regime, cioè alla fine del transitorio dovuto all'inizializzazione, se, come supposto, il rumore di generazione e di osservazione sono stazionari, i pesi \mathbf{k}_n del filtro di Kalman diventano costanti, e cioè:

$$\mathbf{k}_n = [\alpha \quad \beta/T]^T$$

Il filtro prende appunto il nome **alfa-beta** ed è il più semplice dei filtri di Kalman. Il filtro alfa-beta è caratterizzato a regime dalle equazioni:

$$\begin{aligned} \hat{x}_{n|n-1} &= \hat{x}_{n-1} + T \hat{\dot{x}}_{n-1} \\ \hat{\dot{x}}_{n|n-1} &= \hat{\dot{x}}_{n-1} \end{aligned} \quad \text{Predizione} \quad \begin{aligned} \hat{x}_n &= \hat{x}_{n|n-1} + \alpha \left[z_n - \hat{x}_{n|n-1} \right] \\ \hat{\dot{x}}_n &= \hat{\dot{x}}_{n|n-1} + \frac{\beta}{T} \left[z_n - \hat{x}_{n|n-1} \right] \end{aligned} \quad \text{Stima}$$

$$\mathbf{P}_n = \mathbf{P} = \begin{bmatrix} \alpha & \beta/T \\ \beta/T & \frac{\beta(2\alpha - \beta)}{2(1-\alpha)T^2} \end{bmatrix}$$

Errore
quadratico
medio

Filtro alfa-beta

Il filtro alfa-beta funziona molto bene a regime. Se invece viene applicato durante il transitorio, gli errori possono essere anche piuttosto elevati.

Per calcolare i valori di α e β si usano le seguenti relazioni

$$\Gamma = \frac{\sigma_w T^2}{\sigma_v} = \frac{\beta}{\sqrt{1-\alpha}} \quad \beta = 2(2-\alpha) - 4\sqrt{1-\alpha}$$



Modello ad accelerazione quasi-costante

Le equazioni del moto sono:

$$\begin{cases} x_{n+1} = x_n + T \dot{x}_n + \frac{T^2}{2}(\ddot{x}_n + w_n) \\ \dot{x}_{n+1} = \dot{x}_n + T(\ddot{x}_n + w_n) \\ \ddot{x}_{n+1} = \ddot{x}_n + w_n \end{cases}$$

quindi:

$$\mathbf{s}_n = \begin{bmatrix} x_n \\ \dot{x}_n \\ \ddot{x}_n \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} 1 & T & T^2/2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} T^2/2 \\ T \\ 1 \end{bmatrix} \quad \mathbf{Q} = \sigma_w^2$$

Il moto è detto ad accelerazione quasi-costante perché il termine di rumore w_n introduce una piccola accelerazione casuale.

Il **modello delle misure** è $z_n = x_n + v_n$

quindi: $\mathbf{H} = [1 \ 0 \ 0] \quad \mathbf{R} = \sigma_v^2$

A regime, cioè alla fine del transitorio dovuto all'inizializzazione, se, come supposto, il rumore di generazione e di osservazione sono stazionari, i pesi \mathbf{k}_n del filtro di Kalman diventano costanti, e cioè:

$$\mathbf{k}_n = \begin{bmatrix} \alpha & \beta/T & \gamma/(2T^2) \end{bmatrix}^T$$

Il filtro prende il nome **alfa-beta-gamma**. Il filtro alfa-beta è caratterizzato a regime dalle equazioni:

$$\begin{cases} \hat{x}_{n|n-1} = \hat{x}_{n-1} + T \hat{x}_{n-1} + \frac{T^2}{2} \hat{\ddot{x}}_{n-1} \\ \hat{\dot{x}}_{n|n-1} = \hat{\dot{x}}_{n-1} + T \hat{\ddot{x}}_{n-1} \\ \hat{\ddot{x}}_{n|n-1} = \hat{\ddot{x}}_{n-1} \end{cases}$$

Predizione

$$\begin{cases} \hat{x}_n = \hat{x}_{n|n-1} + \alpha [z_n - \hat{x}_{n|n-1}] \\ \hat{\dot{x}}_n = \hat{\dot{x}}_{n|n-1} + \frac{\beta}{T} [z_n - \hat{x}_{n|n-1}] \\ \hat{\ddot{x}}_n = \hat{\ddot{x}}_{n|n-1} + \frac{\gamma}{2T^2} [z_n - \hat{x}_{n|n-1}] \end{cases}$$

Stima

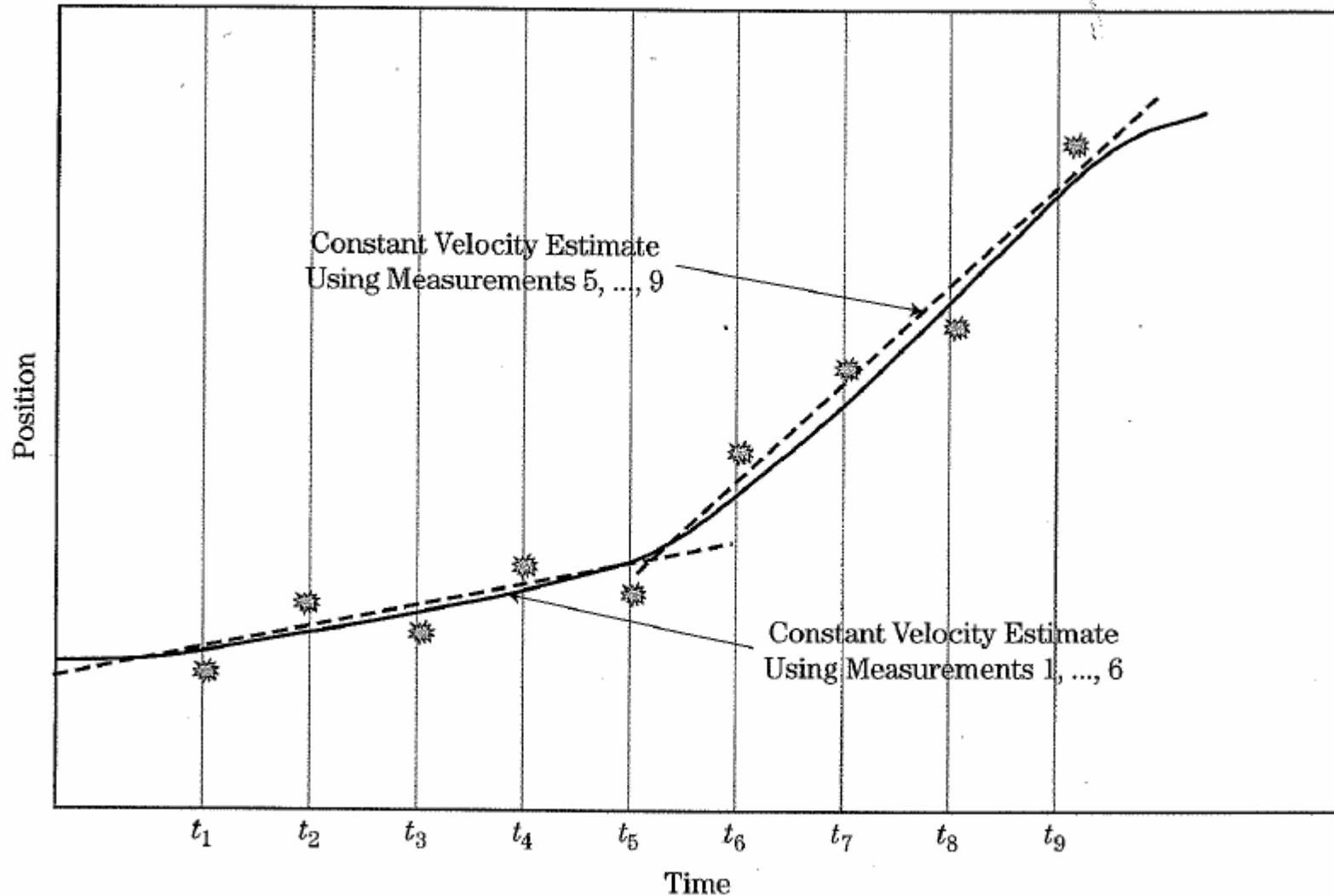
La matrice dell'errore quadratico medio a regime è dato da:

$$\mathbf{P}_n = \mathbf{P} = \begin{bmatrix} \alpha & \beta/T & \frac{\gamma}{2T^2} \\ \beta/T & \frac{8\alpha\beta + \gamma(\beta - 2\alpha - 4)}{8(1-\alpha)T^2} & \frac{\beta(2\beta - \gamma)}{4(1-\alpha)T^3} \\ \frac{\gamma}{2T^2} & \frac{\beta(2\beta - \gamma)}{4(1-\alpha)T^3} & \frac{\gamma(2\beta - \gamma)}{4(1-\alpha)T^4} \end{bmatrix}$$

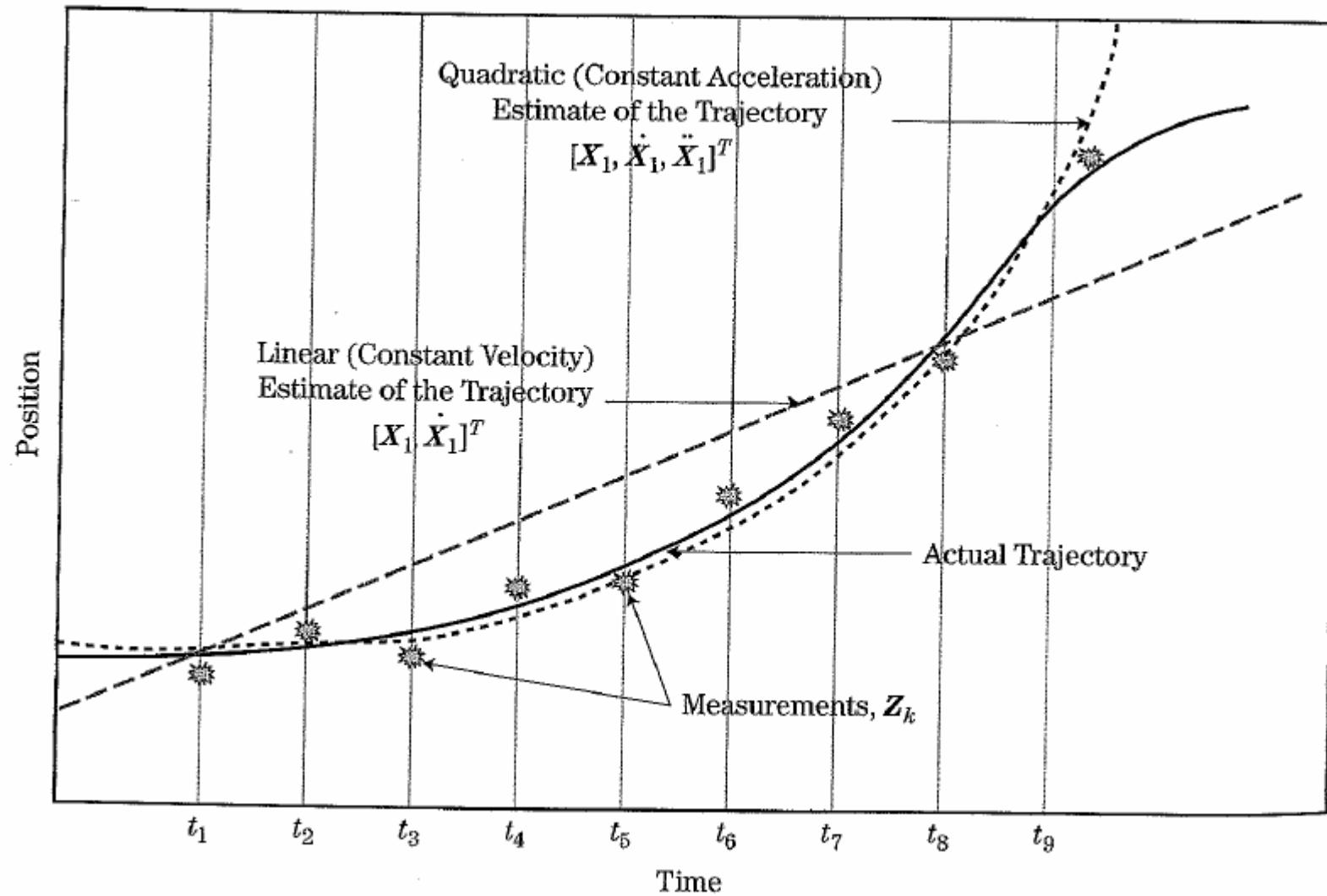
I 3 parametri sono legati tra loro dalle relazioni

$$\Gamma = \frac{\gamma^2}{4(1-\alpha)} \quad \beta = 2(2-\alpha) - 4\sqrt{1-\alpha} \quad \gamma = \frac{\beta^2}{\alpha}$$

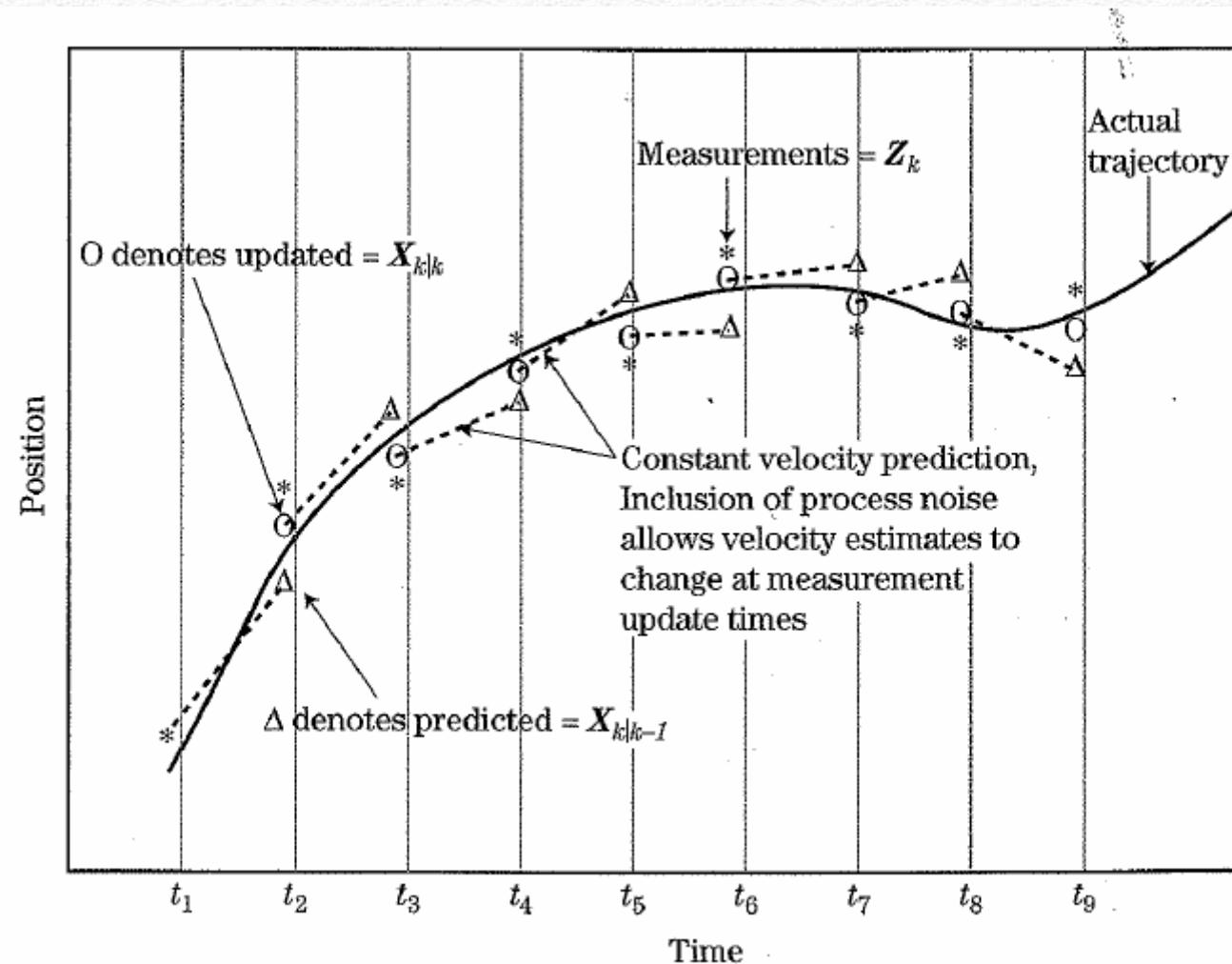
Alcuni esempi



Alcuni esempi



Alcuni esempi



Misure in coordinate non cartesiane e EKF

Sfortunatamente nei casi reali i modelli di moto di target sono in coordinate cartesiane, le misure del radar invece in coordinate polari o sferiche.

L'equazione delle misure non è più lineare nei parametri di stato del target ed in genere si ha:

$$\mathbf{z}_n = \mathbf{H}_n \mathbf{s}_n + \mathbf{v}_n$$

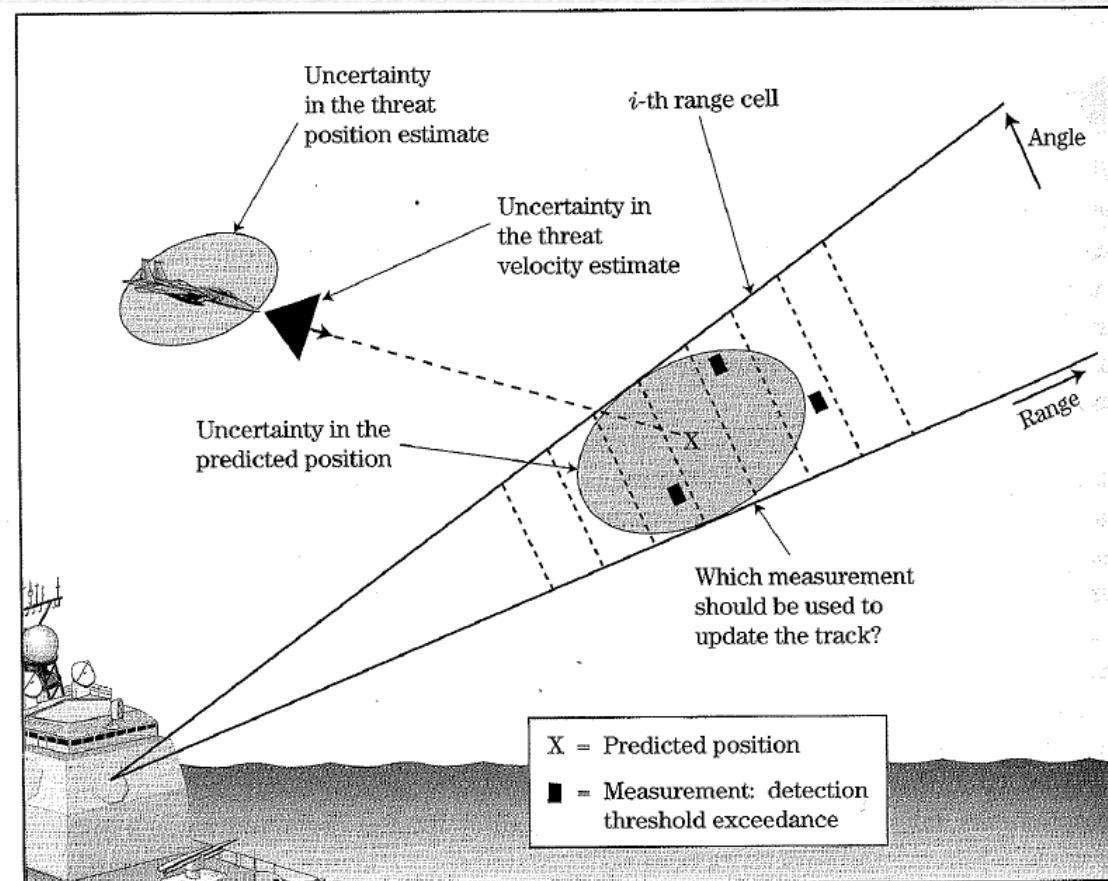
Al posto del filtro di Kalman si usa il filtro di Kalman esteso (EKF) che si ottiene linearizzando il modello delle osservazioni così:

Il filtro EKF può avere anche grossi errori a seconda della bontà dell'approx lineare ni vari pt della traiettoria del target.

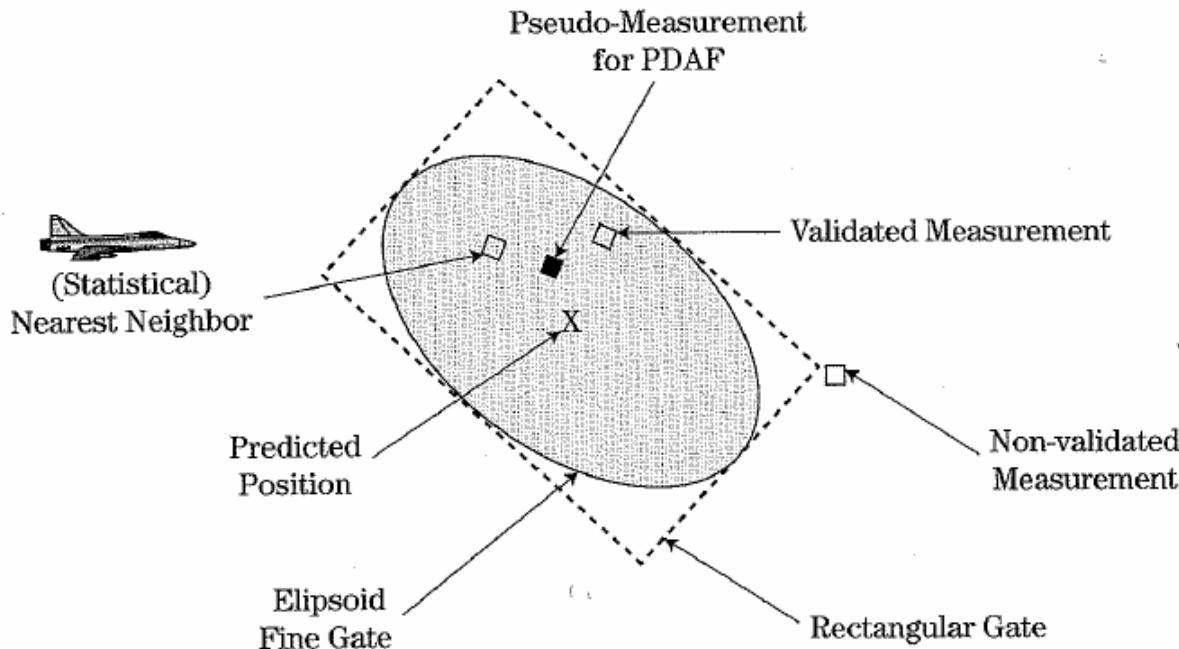
Inizializzazione delle tracce

Supponiamo che a 2 scan successive ci siano 2 rivelazioni vicine, compatibili con la massima velocità possibile del target. La traccia viene inizializzata come **traccia di tentativo**. Se M misurazioni di N scan consecutive sono associate alla traccia di tentativo, la **traccia viene confermata** altrimenti abortita.

Validazione delle misure o Gating



Validazione delle misure o Gating



Si fa individuando una zona intorno alla posizione predetta. Di solito si sceglie un gating rettangolare e poi uno più fine ellissoidale. Le dimensioni dell'ellissoide sono legate alla velocità del target e all'errore del filtro di Kalman. Le misure all'interno vengono validate, quelle all'esterno escluse

Validazione delle misure o Gating

Tra le misure validate, quale si associa alla traccia?

Nearest Neighbor (NN): si calcolano tutte le distanze di Mahalanobis e si sceglie la misura a distanza minima

Strongest Neighbor (SN): Il filtro seleziona la misura a rapporto segnale-rumore massimo

Probabilistic Data Association (PDA): Tutte le misure validate vengono considerate nel costruire la traccia e il termine di correzione introdotto sulla predizione dalla nuova misura è ottenuto pesando tutte le misure con la rispettiva probabilità che ciascuna misura sia originata dal target.

Tracking in angolo

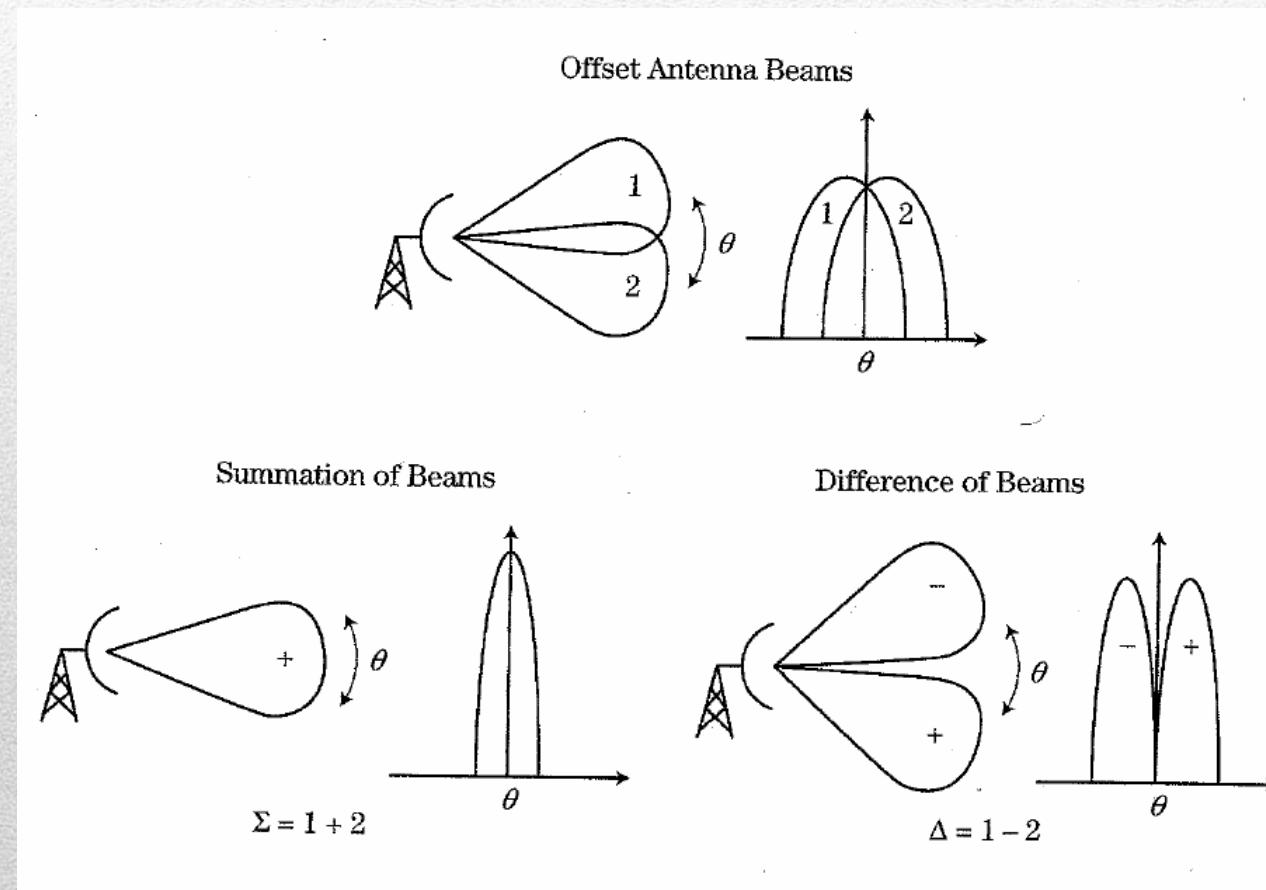
Quando si ha a che fare con radar di scoperta, determinare la posizione angolare di un target con la risoluzione del fascio a 3dB può essere sufficiente. Se invece si ha a che fare con radar di tracking, in genere è richiesta un'accuratezza maggiore.

Esistono varie tecniche di tracking d'angolo. Una delle più comuni è il cosiddetto **MONOPULSE**.

Come dice la parola stessa il Monopulse permette di stimare la direzione di arrivo (DOA) di un target con un solo impulso, utilizzando per ogni coordinata angolare due fasci d'antenna.

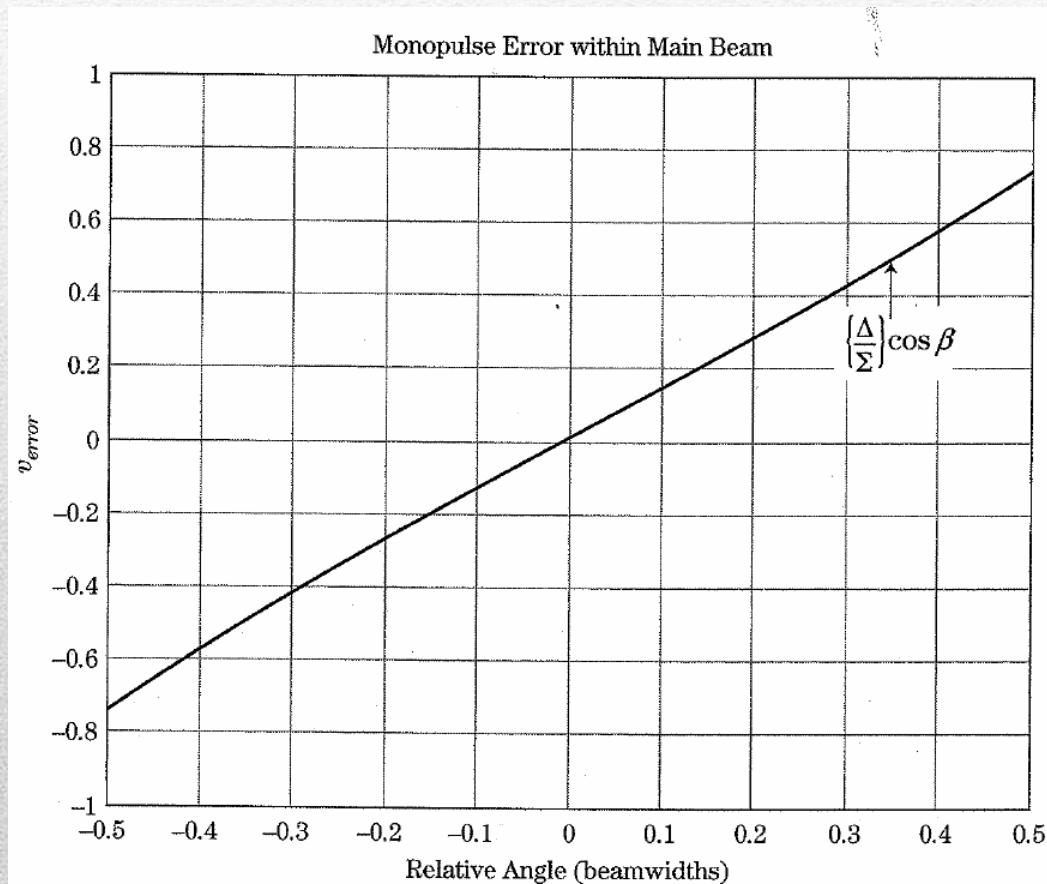
Tracking in angolo

I due fasci guardano in direzioni leggermente diverse, sono cioè “squinted”.



Monopulse di ampiezza

Il fascio somma è usato in trasmissione, mentre entrambi sono usati in ricezione. Indichiamo con Σ il fascio somma e con Δ il fascio differenza. Il loro rapporto è approssimativamente lineare intorno al boresight, e prende il nome di segnale errore



Monopulse di ampiezza

I segnali ricevuti sul fascio somma e differenza sono dati da:

Confrontando il valore di R calcolato dai segnali ricevuti con ve è possibile stimare la posizione del target e spostare di conseguenza il fascio d'antenna in modo da seguire il target.

Vantaggi: può funzionare con un singolo impulso e non è sensibile alle fluttuazioni dell'ampiezza del target anche se si dovessero usare più impulsi.

Criticità: se ci sono più target nel fascio a 3dB il monopulse non è in grado di distinguere tra di loro e fornisce una posizione che è il baricentro di tutti i target.

Sequential lobing

La prima misura è presa con il boresight dell'antenna che punta leggermente da un lato rispetto alla posizione prevista del target, e l'altro impulso nell'altra direzione.

Come stima della posizione vera si prende quella che ha fornito un'eco a potenza maggiore rispetto alle 2 confrontate.

Vantaggi: è un procedimento più semplice del monopulse perché richiede un solo fascio.

Criticità: è molto più sensibile del monopulse alle fluttuazioni in ampiezza del target.

