

23.1 Unscented Kalman filter

Si continua ora con l'unscented Kalman filter (UKF) visto nella lezione precedente. In particolare, il filtro UKF si basa su un metodo per l'approssimazione di densità di probabilità ottenute tramite l'applicazione di mappe non lineari a v.a. statiche. Questo metodo è poi applicato in maniera iterativa per ottenere un'approssimazione in sistemi dinamici.

23.1.1 Trasformazioni nonlineari statiche di variabili aleatorie

Il procedimento dell'UKF consta di tre passaggi. Nel primo la variabile casuale x viene approssimata da $2N + 1$ punti v_i pesati da dei pesi w_i scelti in modo tale che media e varianza coincidano con quelle originali di x . Nel secondo passo si mappano questi punti attraverso la funzione non lineare $g(x)$. Infine i punti mappati vengono approssimati con un'opportuna gaussiana.

Si ha:

$$p_x(\cdot) \sim \mathcal{N}(\bar{x}, P_x) \quad y = g(x)$$

e si approssima la p_y con una gaussiana ottenendo:

$$p_y(\cdot) \simeq \mathcal{N}(\bar{y}, P_y)$$

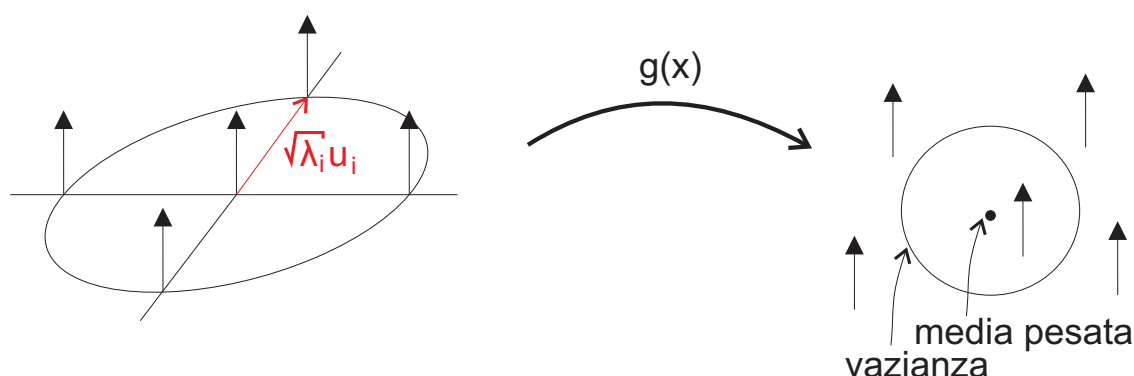


Figura 23.1. Procedimento dell'Unscented Kalman Filter

1. Da $p_x(\cdot) \sim \mathcal{N}(\bar{x}, P_x)$ si generano $2N + 1$ punti (dove $x \in \mathbb{R}^N$) nel modo seguente:

- $v_0 = \bar{x}$
- $v_1 = \bar{x} + \alpha\sigma_1$
- $v_2 = \bar{x} - \alpha\sigma_1$
- \vdots
- $v_{2N-1} = \bar{x} + \alpha\sigma_N$
- $v_{2N} = \bar{x} - \alpha\sigma_N$

dove α è un fattore di scala che viene imposto in modo che la media campionaria dei v_i risulti pari a \bar{x} e la varianza campionaria degli v_i pari a P_x . Il calcolo di α si comincia fattorizzando la varianza di x $P_x = UDU^T$, con $U = [u_1, \dots, u_N]$ una matrice di vettori u_i ortonormali e D la matrice diagonale

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix} \geq 0$$

Si può scrivere la P_x nel modo seguente:

$$\begin{aligned} P_x &= UDU^T \\ &= [u_1 \ \dots \ u_N] \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_N} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_N} \end{bmatrix} \begin{bmatrix} u_1^T \\ \vdots \\ u_N^T \end{bmatrix} \\ &= [\sqrt{\lambda_1}u_1 \ \dots \ \sqrt{\lambda_N}u_N] \begin{bmatrix} \sqrt{\lambda_1}u_1^T \\ \vdots \\ \sqrt{\lambda_N}u_N^T \end{bmatrix} \\ &= \lambda_1 u_1 u_1^T + \dots + \lambda_N u_N u_N^T. \end{aligned}$$

Chiamando $\sigma_i = \sqrt{\lambda_i}u_i$ (vedi vettore rosso in Figura 23.1) si ha che $P_x = \sum_{i=1}^N \sigma_i \sigma_i^T$. La densità di probabilità di x può essere approssimata nel modo seguente:

$$p_x(\cdot) \simeq \sum_{i=0}^{2N} w_i \delta(x - v_i) = \hat{p}_x(\cdot)$$

con i pesi

$$\begin{aligned} w_1 &= w_2 = \dots = w_{2N} = \frac{1}{2(N + \kappa)} \\ w_0 &= 1 - \sum_{i=1}^{2N} w_i = 1 - \frac{N}{N + \kappa} = \frac{\kappa}{N + \kappa}. \end{aligned}$$

Vengono ora calcolate la media e la varianza di x relative alla densità di probabilità approssimata:

$$\begin{aligned}\mathbb{E}_{\hat{p}_x}[x] &= \int_x \sum_{i=0}^{2N} x w_i \delta(x - v_i) dx = \sum_{i=0}^{2N} w_i v_i = \bar{x} \sum_{i=0}^{2N} w_i = \bar{x} \\ \text{Var}_{\hat{p}_x}(x) &= \mathbb{E}_{\hat{p}_x}[(x - \bar{x})(x - \bar{x})^T] = \int_x \sum_{i=0}^{2N} (x - \bar{x})(x - \bar{x})^T w_i \delta(x - v_i) dx, \\ &= \int_{\xi} \sum_{i=0}^{2N} \xi \xi^T w_i \delta(\xi + \bar{x} - v_i) d\xi = \sum_{i=1}^N \frac{1}{2(N + \kappa)} 2\alpha^2 \sigma_i \sigma_i^T \\ &= \frac{\alpha^2}{N + \kappa} \sum_{i=1}^N \sigma_i \sigma_i^T = \frac{\alpha^2}{N + \kappa} P_x\end{aligned}$$

avendo effettuato il cambio di variabili $\xi = x - \bar{x}$. Poiché si vuole imporre che $\text{Var}_{\hat{p}_x}(x) = P_x$ si determina $\alpha = \sqrt{N + \kappa}$.

2. Al secondo passo si devono mappare le v_i usando la funzione non lineare $g(\cdot)$ ottenendo così $y_i = g(v_i)$ con $i = 0, \dots, 2N$. La densità di probabilità approssimata di y risulta essere

$$p_y(\cdot) \simeq \sum_{i=0}^{2N} w_i \delta(y - y_i) = \hat{p}_y(\cdot)$$

3. Al terzo passo si effettuano le seguenti approssimazioni:

$$\hat{p}_y(\cdot) = \mathcal{N}(\bar{y}, P_y) \cong \hat{p}_y(\cdot), \quad \bar{y} := \mathbb{E}_{\hat{p}_y}[y] = \sum_{i=0}^{2N} w_i y_i$$

$$P_y := \text{Var}_{\hat{p}_y}(y) = \mathbb{E}_{\hat{p}_y}[(y - \bar{y})(y - \bar{y})^T] = \sum_{i=0}^N w_i (y_i - \bar{y})(y_i - \bar{y})^T \geq 0$$

Tutti i tre passaggi precedenti dovrebbero garantire, almeno sotto opportune ipotesi, che

$$\hat{\hat{p}}_y(\cdot) \cong p_y(\cdot).$$

23.1.2 Filtro di Kalman Unscented

Fino ad ora si sono mappate densità di probabilità in densità di probabilità e le abbiamo approssimate con delle gaussiane.

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} P_x & P_{xy} \\ P_{xy}^T & P_y \end{bmatrix}\right)$$

con

$$P_{xy} = \text{cov}(x, y) \cong \sum_{i=1}^N w_i (v_i - \bar{x})(y_i - \bar{y})^T$$

$$x|y \sim \mathcal{N}(\mu_x + P_{xy}P_y^{-1}(y - \mu_y), P_x - P_{xy}P_y^{-1}P_{xy}^T)$$

Si cerca ora di applicare quanto visto al modello dinamico non lineare:

$$\begin{cases} x_{k+1} = f(x_k, u_k, w_k) \\ y_k = h(x_k, v_k) \end{cases}$$

Non essendo necessariamente w_k e v_k gaussiani questi vengono approssimati nel modo seguente:

$$w_k \sim p_w(\cdot) \cong \mathcal{N}(\bar{w}, P_w)$$

$$v_k \sim p_v(\cdot) \cong \mathcal{N}(\bar{v}, P_v)$$

$$x_k \cong \mathcal{N}(\hat{x}_{k|k}, P_{k|k})$$

$$x_k \perp v_k \perp w_k$$

e dunque si può scrivere $y = g(\xi)$ con

$$\xi = \begin{bmatrix} x_k \\ w_k \\ v_k \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \hat{x}_{k|k} \\ \bar{w} \\ \bar{v} \end{bmatrix}, \begin{bmatrix} P_{k|k} & 0 & 0 \\ 0 & P_w & 0 \\ 0 & 0 & P_v \end{bmatrix} \right), \quad \begin{cases} x \in \mathbb{R}^n \\ y \in \mathbb{R}^m \\ w \in \mathbb{R}^p \\ v \in \mathbb{R}^q \\ \xi \in \mathbb{R}^{n+p+q} = \mathbb{R}^N \end{cases}$$

Si dovranno dunque generare $2N + 1$ punti $v_i = 0, 1, \dots, 2N$ con i pesi

$$w_i = \frac{1}{2(N + \kappa)} \quad i \neq 0$$

$$w_0 = \frac{\kappa}{N + \kappa}.$$

E ora per calcolare $\hat{x}_{k+1|k+1}$ ho bisogno di $\hat{x}_{k+1|k}$, $\hat{y}_{k+1|k}$, $P_{k+1|k}^x$, $P_{k+1|k}^y$, $P_{k+1|k}^{xy}$, infatti:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} - P_{k+1|k}^{xy} (P_{k+1|k}^y)^{-1} (y_{k+1} - \hat{y}_{k+1|k}) \quad (23.1)$$

e calcolando:

$$\begin{aligned}\hat{x}_{k+1|k} &= \sum_{i=0}^{2N} w_i f(v_i^x, u_k^x, w_i^w) \\ P_{k+1|k}^x &= \sum_{i=0}^{2N} w_i (f(v_i^x, u_k, w_i^w) - \hat{x}_{k+1|k})(f(v_i^x, u_k, w_i^w) - \hat{x}_{k+1|k})^T \\ \hat{y}_{k+1|k} &= \sum_{i=0}^{2N} w_i h(\hat{x}_{k+1|k}^i, v_i^v) \\ P_{k+1|k}^y &= \sum_{i=0}^{2N} w_i (\hat{y}_{k+1|k}^i - \hat{y}_{k+1|k})(\hat{y}_{k+1|k}^i - \hat{y}_{k+1|k})^T \\ P_{k+1|k}^{xy} &= \sum_{i=0}^{2N} w_i (\hat{x}_{k+1|k}^i - \hat{x}_{k+1|k})(\hat{y}_{k+1|k}^i - \hat{y}_{k+1|k})^T\end{aligned}$$

si hanno tutti i termini necessari per implementare la (23.1).

Si noti come in questo caso non è necessario linearizzare ad ogni passo la matrice come nel caso del filtro esteso e si approssimano distribuzioni piuttosto che non linearità. Inoltre con questo metodo si ottiene una precisione del secondo ordine senza calcolare hessiani o jacobiani. I pesi possono poi essere scelti in modo tale da catturare statistiche di ordine più elevato.

Tuttavia nell'implementazione UKF si rende necessario diagonalizzare la matrice di varianza di ξ al variare di $P_{k|k}$. Per rendere meno oneroso questo calcolo si può usare la Squared UKF (che utilizza la fattorizzazione di Cholesky), chiamando Q fattorizzazione di Cholesky di P si può scrivere:

$$P = QQ^T = [v_1 \ \dots \ v_N] \begin{bmatrix} v_1^T \\ \vdots \\ v_N^T \end{bmatrix} = \sum_{i=0}^N v_i v_i^T$$

dove Q è triangolare e risulta numericamente più facile da calcolare rispetto alla diagonalizzazione di P tramite la decomposizione ai valori singolari (SVD).

23.2 Confronto tra UKF e EKF: esempio scalare

Si consideri la seguente mappa lineare, che mappa una v.a. gaussiana in un'altra v.a.:

$$y = x^2, \quad x \in \mathbb{R}, \quad x \sim \mathcal{N}(\bar{x}, \sigma_x^2)$$

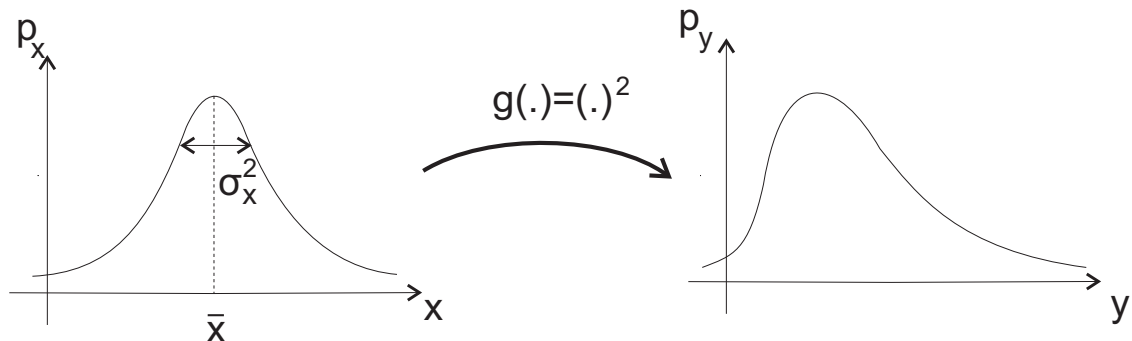


Figura 23.2. Procedimento del filtro UKF

Rappresentando x come $x := \bar{x} + \delta x$ con $\delta x \sim \mathcal{N}(0, \sigma_x^2)$ si calcolano la media e la varianza esatta di y :

$$\begin{aligned} \bar{y}_T &= \mathbb{E}_y[y] = \mathbb{E}_x[x^2] = \mathbb{E}_x[(\bar{x} + \delta x)^2] \\ &= \mathbb{E}[\bar{x}^2 + 2\bar{x}\delta x + (\delta x)^2] \\ &= \bar{x}^2 + \sigma_x^2 \end{aligned}$$

$$\begin{aligned} (\sigma_y^2)_T &= \text{Var}(y) = \mathbb{E}[(x^2 - (\bar{x}^2 + \sigma_x^2))^2] \\ &= \mathbb{E}[4\bar{x}^2(\delta x) + (\delta x)^4 + \sigma_x^4 + 4\bar{x}(\delta x)^3 - 4\bar{x}\sigma_x^2\delta x - 2(\delta x)^2\sigma_x^2] \\ &= 4\bar{x}^2\sigma_x^2 + \mathbb{E}[(\delta x)^4] + \sigma_x^4 - 2\sigma_x^4 \\ &= 2\sigma_x^4 + 4\bar{x}^2\sigma_x^2 \end{aligned}$$

dove nell'ultimo passaggio si è sfruttato il fatto che $\mathbb{E}[(\delta x)^4] = 3\sigma_x^4$. L'esempio prosegue nella prossima lezione.

Bibliografia

- [1] Giorgio Picci. *Fitraggio Statistico (Wiener, Levinson, Kalman) e Applicazioni*. Libreria Progetto, 2006.
- [2] J. Uhlmann S. Julier. A general method for approximating nonlinear transformations of probability distributions. Technical report, University of Oxford, 1996.
- [3] E. Wan R. van der Merwe. Sigma-point kalman filters for integrated navigation. *Proceedings of the 60th Annual of the Institute of Navigation*, 2004.
- [4] G. Oriolo P. Peliti T. Fiorenzani, C. Manes. Comparative study of unscented kalman filter and extended kalman filter for position/attitude estimation in unmanned aerial vehicles. Technical report, IASI CNR, 2002.