

Note sull'assimilazione dati

per gli studenti del corso di
Analisi dei Sistemi Ambientali

Indice

1	Assimilazione dati per il monitoraggio ambientale	1
1.1	Modello e sensori lineari: filtro di Kalman	1
1.2	Modello e/o sensori non lineari: filtro di Kalman esteso	6
1.3	Modello e misure a tempo-continuo: filtro di Kalman-Bucy	7
1.4	Modello a tempo-continuo con misure discrete: filtro di Kalman ibrido . . .	8
1.5	Modello a grandi dimensioni: il filtro di Kalman <i>ensemble</i>	9
2	Calibrazione dei parametri mediante assimilazione dati	11

Capitolo 1

Assimilazione dati per il monitoraggio ambientale

1.1 Modello e sensori lineari: filtro di Kalman

Si consideri un modello lineare a tempo-discreto

$$x_{k+1} = A_k x_k + b_k + w_k \quad (1.1.1)$$

dove: $x_k \in \mathbb{R}^n$ è lo stato che si desidera monitorare in tempo reale; $A_k \in \mathbb{R}^{n \times n}$ è una matrice di modello supposta nota; $b_k \in \mathbb{R}^n$ è una quantità nota; $w_k \in \mathbb{R}^n$ rappresenta un *errore di modello* incognito che tiene conto di diverse sorgenti di incertezza. Per monitorare lo stato x_k sono disponibili, all'istante di campionamento k , misure (dati, osservazioni) fornite da appropriati sensori

$$y_k = C_k x_k + v_k \quad (1.1.2)$$

dove: $y_k \in \mathbb{R}^p$ è il vettore di misura; p è il numero di sensori; $C_k \in \mathbb{R}^{p \times n}$ è una matrice nota; $v_k \in \mathbb{R}^p$ rappresenta un *errore di misura* incognito che tiene conto delle imprecisioni dei sensori utilizzati. L'obiettivo è quello di assimilare i dati (1.1.2), man mano che questi vengono acquisiti dai sensori, nel modello (1.1.1) al fine di stimare in tempo reale lo stato x_k dell'ecosistema. A tale proposito, un criterio ragionevole potrebbe essere quello di scegliere le stime degli stati $x_1, x_2, \dots, x_k, x_{k+1}$ che minimizzano un indice di prestazione della forma

$$\begin{aligned} J_{1:k}(x_1, \dots, x_k, x_{k+1}) = & (x_1 - \hat{x}_{1|0})^T P_{1|0}^{-1} (x_1 - \hat{x}_{1|0}) + \\ & \sum_{i=1}^k (y_i - C_i x_i)^T R_i^{-1} (y_i - C_i x_i) + \\ & \sum_{i=1}^k (x_{i+1} - A_i x_i - b_i)^T Q_i^{-1} (x_{i+1} - A_i x_i - b_i) \end{aligned} \quad (1.1.3)$$

che somma i quadrati - opportunamente ponderati con le matrici di peso $P_{1|0}^{-1}, R_i^{-1}, Q_i^{-1}$ - dell'errore di stima iniziale $\tilde{x}_{1|0} \triangleq x_1 - \hat{x}_{1|0}$, degli errori di misura $v_i = y_i - C_i x_i$ e degli errori di modello $w_i = x_{i+1} - A_i x_i - b_i$. Poiché risulta naturale dare a ciascuno degli errori sopra menzionati un'importanza relativa inversamente proporzionale alla corrispondente incertezza, le matrici di peso in (1.1.3) sono state indicate come inverse di opportune matrici $P_{1|0}, R_i, Q_i$ *simmetriche*¹ e *definite positive*² che quantificano l'incertezza rispettivamente della stima iniziale, della misura, del modello. Pertanto, il problema di assimilazione dati si riconduce all'ottimizzazione

$$\{\hat{x}_{1|k}, \dots, \hat{x}_{k|k}, \hat{x}_{k+1|k}\} = \arg \min_{x_1, \dots, x_k, x_{k+1}} J_{1:k}(x_1, \dots, x_k, x_{k+1}). \quad (1.1.4)$$

Si noti come tale ottimizzazione abbia complessità crescente nel tempo, ovvero all'aumentare di k . Per assimilare dati in tempo reale risulta, quindi, necessario procedere in modo ricorsivo. A tale scopo, si assuma di disporre all'istante k , per effetto delle elaborazioni effettuate agli istanti precedenti, di una stima $\hat{x}_{k|k-1}$ di x_k che ha assimilato in modo ottimale i dati $y_{1:k-1} \triangleq \{y_1, \dots, y_{k-1}\}$ e della relativa matrice di incertezza $P_{k|k-1}$. L'obiettivo è quello di determinare, a partire da $\hat{x}_{k|k-1}$ e $P_{k|k-1}$, le stime $\hat{x}_{k|k}$ e $\hat{x}_{k+1|k}$ aggiornate con l'assimilazione dell'ultimo dato y_k , risolvendo il seguente problema di ottimizzazione

$$\{\hat{x}_{k|k}, \hat{x}_{k+1|k}\} = \arg \min_{x_k, x_{k+1}} \underbrace{\begin{aligned} & (x_k - \hat{x}_{k|k-1})^T P_{k|k-1}^{-1} (x_k - \hat{x}_{k|k-1}) + \\ & (y_k - C_k x_k)^T R_k^{-1} (y_k - C_k x_k) + \\ & (x_{k+1} - A_k x_k - b_k)^T Q_k^{-1} (x_{k+1} - A_k x_k - b_k) \end{aligned}}_{J_{k:k+1}(x_k, x_{k+1})} \quad (1.1.5)$$

Sviluppando i prodotti nel costo (1.1.5) e raccogliendo i termini simili, si ottiene

$$\begin{aligned} J_{k:k+1}(x_k, x_{k+1}) &= x_k^T \underbrace{\left(P_{k|k-1}^{-1} + C_k^T R_k^{-1} C_k + A_k^T Q_k^{-1} A_k \right)}_{M_k} x_k \\ &\quad - 2x_k^T \underbrace{\left(P_{k|k-1}^{-1} \hat{x}_{k|k-1} + C_k^T R_k^{-1} y_k + A_k^T Q_k^{-1} (x_{k+1} - b_k) \right)}_{\nu_k} \\ &\quad + x_{k+1}^T Q_k^{-1} x_{k+1} - 2x_{k+1}^T Q_k^{-1} b_k \\ &\quad + \hat{x}_{k|k-1}^T P_{k|k-1}^{-1} \hat{x}_{k|k-1} + y_k^T R_k^{-1} y_k + b_k^T Q_k^{-1} b_k \end{aligned} \quad (1.1.6)$$

Il minimo globale della forma quadratica $J_{k:k+1}(x_k, x_{k+1})$ si ottiene imponendo l'annullamento del gradiente, ovvero

$$\frac{\partial J_{k:k+1}}{\partial x_k} = 0 \quad \text{e} \quad \frac{\partial J_{k:k+1}}{\partial x_{k+1}} = 0$$

¹Una matrice M dicesi simmetrica se $M = M^T$.

²Una matrice quadrata M dicesi definita positiva (si scrive $M \succ 0$) se, per ogni vettore non nullo v , la forma quadratica associata è positiva, i.e. $v^T M v > 0$.

In particolare, ricordando che $\frac{\partial}{\partial x} (x^T M x) = 2Mx$ e $\frac{\partial}{\partial x} (v^T x) = \frac{\partial}{\partial x} (x^T v) = v$, si ha

$$\frac{\partial J_{k:k+1}}{\partial x_k} = 0 \Rightarrow 2(M_k x_k - \nu_k) = 0 \Rightarrow M_k x_k = \nu_k \quad (1.1.7)$$

In modo analogo,

$$\frac{\partial J_{k:k+1}}{\partial x_{k+1}} = 0 \Rightarrow Q_k^{-1} (x_{k+1} - A_k x_k - b_k) = 0 \Rightarrow x_{k+1} = A_k x_k + b_k \quad (1.1.8)$$

Combinando (1.1.7), (1.1.8) e le definizioni di M_k, ν_k in (1.1.6), si ha

$$\left(P_{k|k-1}^{-1} + C_k^T R_k^{-1} C_k + A_k^T Q_k^{-1} A_k \right) x_k = P_{k|k-1}^{-1} \hat{x}_{k|k-1} + C_k^T R_k^{-1} y_k + A_k^T Q_k^{-1} A_k x_k \quad (1.1.9)$$

da cui si deduce la stima ottima

$$\begin{aligned} x_k = \hat{x}_{k|k} &= \left(P_{k|k-1}^{-1} + C_k^T R_k^{-1} C_k \right)^{-1} \left(P_{k|k-1}^{-1} \hat{x}_{k|k-1} + C_k^T R_k^{-1} y_k \right) \\ &= P_{k|k} \left(P_{k|k-1}^{-1} \hat{x}_{k|k-1} + C_k^T R_k^{-1} y_k \right) \end{aligned} \quad (1.1.10)$$

dove si è posto

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + C_k^T R_k^{-1} C_k \quad (1.1.11)$$

Sfruttando (1.1.11), l'aggiornamento della stima (1.1.10) può essere espresso equivalentemente nella seguente forma

$$\begin{aligned} \hat{x}_{k|k} &= P_{k|k} \left(P_{k|k}^{-1} - C_k^T R_k^{-1} C_k \right) \hat{x}_{k|k-1} + P_{k|k} C_k^T R_k^{-1} y_k \\ &= \hat{x}_{k|k-1} + \underbrace{P_{k|k} C_k^T R_k^{-1}}_{L_k} \underbrace{(y_k - C_k \hat{x}_{k|k-1})}_{e_k}. \end{aligned} \quad (1.1.12)$$

Dalla precedente forma si evince come la stima nuova $\hat{x}_{k|k}$, quella che ha assimilato anche l'ultimo dato y_k , si ottenga dalla stima vecchia $\hat{x}_{k|k-1}$, che aveva assimilato solo i dati precedenti, aggiungendole un termine correttivo dato dal prodotto di un guadagno L_k per un fattore di innovazione e_k che coincide con l'errore di predizione della nuova misura. Negli sviluppi che seguono, si farà ampio ricorso alla seguente formula di inversione (di immediata verifica)

$$(A + BCB^T)^{-1} = A^{-1} - A^{-1}B(C^{-1} + B^T A^{-1}B)^{-1} B^T A^{-1} \quad (1.1.13)$$

valida quando le matrici A e C sono invertibili. Si farà inoltre uso della seguente formula, diretta conseguenza di (1.1.13),

$$(A + BCB^T)^{-1} BC = A^{-1}B(C^{-1} + B^T A^{-1}B)^{-1} \quad (1.1.14)$$

Applicando (1.1.13) con $A = P_{k|k-1}^{-1}$, $B = C_k^T$, $C = R_k^{-1}$ a (1.1.11), si ottiene la seguente formula di aggiornamento della matrice di incertezza

$$P_{k|k} = P_{k|k-1} - P_{k|k-1} C_k^T (R_k + C_k P_{k|k-1} C_k^T)^{-1} C_k P_{k|k-1} \quad (1.1.15)$$

Inoltre, applicando (1.1.14) con le stesse matrici, si ha

$$L_k = P_{k|k} C_k^T R_k^{-1} = \left(P_{k|k-1}^{-1} + C_k^T R_k^{-1} C_k \right)^{-1} C_k^T R_k^{-1} = P_{k|k-1} C_k^T (R_k + C_k P_{k|k-1} C_k^T)^{-1} \quad (1.1.16)$$

Dalla stima ottima $\hat{x}_{k|k}$, tramite (1.1.8), si ottiene la predizione ottima ad un passo dello stato

$$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k} + b_k \quad (1.1.17)$$

A questo punto la soluzione del problema (1.1.5) risulta determinata da (1.1.12) e (1.1.17) con guadagno L_k fornito da (1.1.16). Per procedere ricorsivamente al passo successivo, occorre esprimere il costo residuo $J_{k+1}(x_{k+1}) \triangleq J_{k:k+1}(M_k^{-1}\nu_k, x_{k+1})$ nella forma

$$J_{k+1}(x_{k+1}) = (x_{k+1} - \hat{x}_{k+1|k})^T P_{k+1|k}^{-1} (x_{k+1} - \hat{x}_{k+1|k}) + \alpha_k \quad (1.1.18)$$

dove α_k è un termine aggiuntivo che non dipende da x_{k+1} e $P_{k+1|k}$ una opportuna matrice di incertezza da determinare. Sostituendo $x_k = M_k^{-1}\nu_k$ in (1.1.6), si ha

$$\begin{aligned} J_{k+1}(x_{k+1}) &= J_{k:k+1}(M_k^{-1}\nu_k, x_{k+1}) \\ &= -\nu_k^T M_k^{-1} \nu_k \\ &\quad + x_{k+1}^T Q_k^{-1} x_{k+1} - 2x_{k+1}^T Q_k^{-1} b_k + \\ &\quad + \hat{x}_{k|k-1}^T P_{k|k-1}^{-1} \hat{x}_{k|k-1} + y_k^T R_k^{-1} y_k + b_k^T Q_k^{-1} b_k \\ &= - \left(P_{k|k}^{-1} \hat{x}_{k|k} + A_k^T Q_k^{-1} (x_{k+1} - b_k) \right)^T \left(P_{k|k}^{-1} + A_k^T Q_k^{-1} A_k \right)^{-1} (\dots) \\ &\quad + x_{k+1}^T Q_k^{-1} x_{k+1} - 2x_{k+1}^T Q_k^{-1} b_k + \\ &\quad + \hat{x}_{k|k-1}^T P_{k|k-1}^{-1} \hat{x}_{k|k-1} + y_k^T R_k^{-1} y_k + b_k^T Q_k^{-1} b_k \\ &= x_{k+1}^T \left[Q_k^{-1} - Q_k^{-1} A_k \left(P_{k|k}^{-1} + A_k^T Q_k^{-1} A_k \right)^{-1} A_k^T Q_k^{-1} \right] x_{k+1} \\ &\quad - 2x_{k+1}^T \left\{ \left[Q_k^{-1} A_k \left(P_{k|k}^{-1} + A_k^T Q_k^{-1} A_k \right)^{-1} \left(P_{k|k}^{-1} \hat{x}_{k|k} - A_k^T Q_k^{-1} b_k \right) \right] + Q_k^{-1} b_k \right\} \\ &\quad + \text{somma di termini indipendenti da } x_{k+1} \\ &= x_{k+1}^T P_{k+1|k}^{-1} x_{k+1} - 2x_{k+1}^T q_{k+1} + \dots \end{aligned} \quad (1.1.19)$$

dove

$$P_{k+1|k} = \left[Q_k^{-1} - Q_k^{-1} A_k \left(P_{k|k}^{-1} + A_k^T Q_k^{-1} A_k \right)^{-1} A_k^T Q_k^{-1} \right]^{-1} \quad (1.1.20)$$

$$\begin{aligned} q_{k+1} &= Q_k^{-1} A_k \left(P_{k|k}^{-1} + A_k^T Q_k^{-1} A_k \right)^{-1} P_{k|k}^{-1} \hat{x}_{k|k} \\ &\quad + \left[Q_k^{-1} - Q_k^{-1} A_k \left(P_{k|k}^{-1} + A_k^T Q_k^{-1} A_k \right)^{-1} A_k^T Q_k^{-1} \right]^{-1} b_k \end{aligned} \quad (1.1.21)$$

e i termini indicati con \dots non dipendono dagli stati x_{k+1}, x_{k+2}, \dots e risultano quindi ininfluenti ai fini della stima (ottimizzazione) negli istanti successivi all'istante k . Applicando la formula di inversione (1.1.13) con $A = Q_k^{-1}$, $B = A_k$, $C = P_{k|k}$ a (1.1.20), si ottiene

$$P_{k+1|k} = A_k P_{k|k} A_k^T + Q_k \quad (1.1.22)$$

Applicando la stessa formula, si ha

$$\begin{aligned} &Q_k^{-1} A_k \left(P_{k|k}^{-1} + A_k^T Q_k^{-1} A_k \right)^{-1} P_{k|k}^{-1} = \\ &Q_k^{-1} A_k \left[P_{k|k} - P_{k|k} A_k^T \left(Q_k + A_k P_{k|k} A_k^T \right)^{-1} A_k P_{k|k} \right] P_{k|k}^{-1} = \\ &Q_k^{-1} \left[I - A_k P_{k|k} A_k^T \left(Q_k + A_k P_{k|k} A_k^T \right)^{-1} \right] A_k P_{k|k} P_{k|k}^{-1} = \\ &Q_k^{-1} \left(Q_k + A_k P_{k|k} A_k^T - A_k P_{k|k} A_k^T \right) \left(Q_k + A_k P_{k|k} A_k^T \right)^{-1} A_k = \\ &\left(Q_k + A_k P_{k|k} A_k^T \right)^{-1} A_k = \\ &P_{k+1|k}^{-1} A_k \end{aligned} \quad (1.1.23)$$

Sostituendo (1.1.20) e (1.1.23) in (1.1.21), si ottiene

$$q_{k+1} = P_{k+1|k}^{-1} \left(A_k \hat{x}_{k|k} + b_k \right) = P_{k+1|k}^{-1} \hat{x}_{k+1|k} \quad (1.1.24)$$

Pertanto, (1.1.19) assume effettivamente la forma (1.1.18) con $\hat{x}_{k+1|k}$ e $P_{k+1|k}$ dati da (1.1.17) e, rispettivamente, (1.1.22). Riassumendo i precedenti sviluppi algebrici, si ottiene l'algoritmo ricorsivo di assimilazione dati, noto in letteratura come *filtro di Kalman*, di seguito riportato.

Filtro di Kalman

Inizializzazione: $\hat{x}_{1|0}, P_{1|0}$

Per $k = 1, 2, \dots$

% *Correzione*

$$S_k = R_k + C_k P_{k|k-1} C_k^T$$

$$L_k = P_{k|k-1} C_k^T S_k^{-1}$$

$$e_k = y_k - C_k \hat{x}_{k|k-1}$$

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + L_k e_k \\ P_{k|k} &= P_{k|k-1} - L_k S_k L_k^T \\ \% \text{ Predizione} \\ \hat{x}_{k+1|k} &= A_k \hat{x}_{k|k} + b_k \\ P_{k+1|k} &= A_k P_{k|k} A_k^T + Q_k\end{aligned}$$

1.2 Modello e/o sensori non lineari: filtro di Kalman esteso

Si consideri adesso la situazione più generale in cui il modello dell'ecosistema e/o i sensori possano essere non lineari, i.e.

$$\begin{cases} x_{k+1} = f_k(x_k) + w_k \\ y_k = h_k(x_k) + v_k \end{cases} \quad (1.2.1)$$

In modo analogo al caso lineare, si può formulare in modo *batch* (non ricorsivo) l'assimilazione dati tramite il seguente problema di ottimizzazione

$$\begin{aligned} \{ \hat{x}_{1|k}, \dots, \hat{x}_{k|k}, \hat{x}_{k+1|k} \} = \arg \min_{x_1, \dots, x_k, x_{k+1}} & \left\{ (x_1 - \hat{x}_{1|0})^T P_{1|0}^{-1} (x_1 - \hat{x}_{1|0}) + \right. \\ & \sum_{i=1}^k (y_i - h_i(x_i))^T R_i^{-1} (y_i - h_i(x_i)) + \\ & \left. \sum_{i=1}^k (x_{i+1} - f_i(x_i))^T Q_i^{-1} (x_{i+1} - f_i(x_i)) \right\} \end{aligned} \quad (1.2.2)$$

Diversamente dal caso lineare, tuttavia, il problema (1.2.2), a causa delle non linearità delle funzioni $f_i(\cdot)$ e $h_i(\cdot)$, non risulta risolvibile analiticamente per cui si deve cercare una soluzione numerica. In particolare, la soluzione numerica di (1.2.2) può essere determinata ricorsivamente con il seguente approccio ad orizzonte mobile

$$\begin{aligned} \{ \hat{x}_{k-N|k}, \dots, \hat{x}_{k|k}, \hat{x}_{k+1|k} \} = \arg \min_{x_{k-N}, \dots, x_k, x_{k+1}} & \left\{ (x_{k-N} - \hat{x}_{k-N|k-1})^T P^{-1} (x_{k-N} - \hat{x}_{k-N|k-1}) + \right. \\ & \sum_{i=k-N}^k (y_i - h_i(x_i))^T R_i^{-1} (y_i - h_i(x_i)) + \\ & \left. \sum_{i=k-N}^k (x_{i+1} - f_i(x_i))^T Q_i^{-1} (x_{i+1} - f_i(x_i)) \right\} \end{aligned} \quad (1.2.3)$$

dove $N \geq 1$ è l'ampiezza della finestra mobile di assimilazione dati. Un approccio computazionalmente meno oneroso consiste nel linearizzare il sistema (1.2.1) nell'intorno della stima corrente e applicare al sistema linearizzato il filtro di Kalman precedentemente introdotto. Sviluppando $h_k(x_k)$ nell'intorno di $\hat{x}_{k|k-1}$, si ha

$$h_k(x_k) = h_k(\hat{x}_{k|k-1}) + \underbrace{\frac{\partial h_k}{\partial x_k}(\hat{x}_{k|k-1})}_{C_k} \underbrace{(x_k - \hat{x}_{k|k-1})}_{\tilde{x}_{k|k-1}} + o(\|\tilde{x}_{k|k-1}\|^2) \approx h_k(\hat{x}_{k|k-1}) + C_k \tilde{x}_{k|k-1}$$

e analogamente per $f_k(x_k)$ nell'intorno di $\hat{x}_{k|k}$:

$$f_k(x_k) = f_k(\hat{x}_{k|k}) + \underbrace{\frac{\partial f_k}{\partial x_k}(\hat{x}_{k|k})}_{A_k} \underbrace{(x_k - \hat{x}_{k|k})}_{\tilde{x}_{k|k}} + o(\|\tilde{x}_{k|k}\|^2) \approx f_k(\hat{x}_{k|k}) + A_k \tilde{x}_{k|k}$$

Grazie alle precedenti approssimazioni, si può estendere il filtro di Kalman al sistema non-lineare (1.2.1) ottenendo il seguente algoritmo di assimilazione dati, noto in letteratura come *filtro di Kalman esteso*.

Filtro di Kalman esteso

Inizializzazione: $\hat{x}_{1|0}, P_{1|0}$

Per $k = 1, 2, \dots$

% Correzione

$$C_k = \frac{\partial h_k}{\partial x_k}(\hat{x}_{k|k-1})$$

$$S_k = R_k + C_k P_{k|k-1} C_k^T$$

$$L_k = P_{k|k-1} C_k^T S_k^{-1}$$

$$e_k = y_k - h_k(\hat{x}_{k|k-1})$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k e_k$$

$$P_{k|k} = P_{k|k-1} - L_k S_k L_k^T$$

% Predizione

$$A_k = \frac{\partial f_k}{\partial x_k}(\hat{x}_{k|k})$$

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

$$P_{k+1|k} = A_k P_{k|k} A_k^T + Q_k$$

1.3 Modello e misure a tempo-continuo: filtro di Kalman-Bucy

Si considera adesso un modello a tempo-continuo, come è la prassi nel monitoraggio di ecosistemi, e sensori che acquisiscono misure a tempo-continuo, sebbene questo sia poco

realistico, i.e.

$$\begin{cases} \dot{x}(t) = f(t, x(t)) + w(t) \\ y(t) = h(t, x(t)) + v(t) \end{cases} \quad (1.3.1)$$

In modo analogo al caso tempo-discreto, si cerca l'assimilatore dati che ottimizza il seguente costo

$$\begin{aligned} \{\hat{x}(\sigma), 0 \leq \sigma \leq t\} = \arg \min_{x(\sigma), 0 \leq \sigma \leq t} & \left\{ [x(0) - \hat{x}(0)]^T P^{-1}(0) [x(0) - \hat{x}(0)] \right. \\ & + \int_0^t [y(\sigma) - h(\sigma, x(\sigma))]^T R^{-1}(\sigma) [y(\sigma) - h(\sigma, x(\sigma))] d\sigma \\ & \left. + \int_0^t [\dot{x}(\sigma) - f(\sigma, x(\sigma))]^T Q^{-1}(\sigma) [\dot{x}(\sigma) - f(\sigma, x(\sigma))] d\sigma \right\} \end{aligned} \quad (1.3.2)$$

Nel caso lineare, i.e. $f(t, x) = A(t)x + b(t)$ e $h(t, x) = C(t)x$, la soluzione di (1.3.2) è fornita dal seguente *filtro di Kalman-Bucy*.

Filtro di Kalman-Bucy

Si risolvono le equazioni differenziali ordinarie

$$\begin{aligned} \dot{\hat{x}}(t) &= A(t)\hat{x}(t) + b(t) + P(t)C^T(t)R^{-1}(t)[y(t) - C(t)\hat{x}(t)] \\ \dot{P}(t) &= A(t)P(t) + P(t)A^T(t) + Q(t) - P(t)C^T(t)R^{-1}(t)C(t)P(t) \end{aligned} \quad (1.3.3)$$

con condizioni iniziali $\hat{x}(0)$ e $P(0)$ assegnate.

Si noti come la soluzione (1.3.3) possa essere utilizzata in modo approssimato anche per modello e/o sensori non lineari, sostituendo $A(t)\hat{x}(t) + b(t)$ con $f(t, \hat{x}(t))$, $C(t)\hat{x}(t)$ con $h(t, \hat{x}(t))$ in (1.3.3), e ponendo

$$A(t) \triangleq \frac{\partial f}{\partial x}(t, \hat{x}(t)) \quad \text{e} \quad C(t) \triangleq \frac{\partial h}{\partial x}(t, \hat{x}(t)) \quad (1.3.4)$$

1.4 Modello a tempo-continuo con misure discrete: filtro di Kalman ibrido

L'acquisizione e l'elaborazione di misure a tempo-continuo risultano irrealistiche. Viceversa, avere a che fare con un modello a tempo-continuo e misure campionate $y_k = y(t_k)$ ad istanti discreti $t_1 < t_2 < \dots < t_k < t_{k+1} < \dots$ è la situazione pratica più ricorrente nelle applicazioni di monitoraggio ambientale. Questa situazione ibrida (modello

tempo-continuo e misure tempo-discreto) è descritta come segue:

$$\begin{cases} \dot{x}(t) = f(t, x(t)) + w(t) \\ y(t) = h(t, x(t)) + v(t) \\ y_k = y(t_k) \end{cases} \quad (1.4.1)$$

Combinando la correzione del filtro di Kalman agli istanti discreti t_k con la predizione (simulazione) a tempo-continuo negli intervalli $[t_k, t_{k+1}]$, si ottiene il seguente assimilatore dati ibrido.

Filtro di Kalman esteso ibrido

Inizializzazione: $\hat{x}_{1|0}, P_{1|0}$

Per $k = 1, 2, \dots$

% Correzione

$$C_k = \frac{\partial h}{\partial x}(t_k, \hat{x}_{k|k-1})$$

$$R_k = R(t_k)$$

$$S_k = R_k + C_k P_{k|k-1} C_k^T$$

$$L_k = P_{k|k-1} C_k^T S_k^{-1}$$

$$e_k = y_k - h(t_k, \hat{x}_{k|k-1})$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k e_k$$

$$P_{k|k} = P_{k|k-1} - L_k S_k L_k^T$$

% Predizione

Si risolve, nell'intervallo $[t_k, t_{k+1}]$, l'equazione differenziale vettoriale

$$\dot{x}(t) = f(t, x(t)) \quad (1.4.2)$$

con condizione iniziale $x(t_k) = \hat{x}_{k|k}$ e si pone $\hat{x}_{k+1|k} = x(t_{k+1})$.

Si risolve, nell'intervallo $[t_k, t_{k+1}]$, l'equazione differenziale matriciale

$$\dot{P}(t) = A(t)P(t) + P(t)A^T(t) + Q(t) \quad (1.4.3)$$

con

$$A(t) \triangleq \frac{\partial f}{\partial x}(t, x(t))$$

e condizione iniziale $P(t_k) = P_{k|k}$; si pone $P_{k+1|k} = P(t_{k+1})$.

1.5 Modello a grandi dimensioni: il filtro di Kalman ensemble

Questa parte è stata solo accennata, non svolta in dettaglio, nel corso dell'A.A. 2016/17 per motivi di tempo. Gli algoritmi di assimilazione dati finora esplorati hanno il loro

“collo di bottiglia”, dal punto di vista computazionale, nella necessità di dover calcolare e memorizzare la matrice, $n \times n$, P di incertezza dello stato stimato. Questo comporta un onere di calcolo dell'ordine di n^3 , i.e. $O(n^3)$, ed un onere di memorizzazione $O(n^2)$ che risultano proibitivi per applicazioni ambientali su larga scala che coinvolgono dell'ordine di migliaia, milioni, miliardi, . . . di variabili di stato. Per problemi di queste dimensioni, è stato concepito un metodo *ad-hoc* di assimilazione dati, noto in letteratura come *filtro di Kalman ensemble* [Evensen, 2009], con la peculiarità di evitare il calcolo esplicito e la memorizzazione della matrice di incertezza $P \in \mathbb{R}^{n \times n}$. A differenza di EKF (filtro di Kalman esteso), EnKF (filtro di Kalman ensemble) rappresenta l'incertezza con un *ensemble* di $N \ll n$ realizzazioni del modello e calcola il guadagno di correzione L_k richiesto per assimilare la misura y_k attraverso medie campionarie sull'*ensemble* di realizzazioni. Solitamente valori della dimensione N dell'*ensemble* dell'ordine delle decine-centinaia sono sufficienti per un'assimilazione dati accurata con modelli dell'ordine di milioni di variabili di stato. In questo modo, la complessità di calcolo $O(n^2N)$ e di memorizzazione $O(nN)$ di EnKF si riduce di un ordine di grandezza rispetto a EKF. I dettagli dell'algoritmo EnKF sono reperibili, ad esempio, in [Evensen, 2007] e [Evensen, 2009].

Capitolo 2

Calibrazione dei parametri mediante assimilazione dati

Nei modelli matematici di ecosistemi compaiono spesso parametri non noti e di difficile determinazione sperimentale diretta. La calibrazione (stima) di tali parametri con dati sperimentali risulta indispensabile per poter utilizzare tali modelli ai fini dell'analisi e/o del monitoraggio dell'ecosistema. In modo del tutto generale, il modello dell'ecosistema assume la forma

$$\begin{cases} \dot{x}(t) &= f(t, x(t), \theta) \\ y(t) &= h(t, x(t)) \end{cases} \quad (2.0.1)$$

dove $\theta \in \mathbb{R}^{n_\theta}$ è il vettore dei parametri da calibrare. Per la calibrazione, si assume di disporre di N dati sperimentali $y_k = y(t_k)$ per $k = 1, 2, \dots, N$. Fra i numerosi approcci esistenti, uno dei più efficaci - noto in letteratura come *minimizzazione dell'errore di predizione* - si propone di minimizzare il seguente costo

$$J(\theta, y_{1:N}) \triangleq \sum_{k=1}^N (y_k - \hat{y}_{k|k-1}(\theta))^T W_k (y_k - \hat{y}_{k|k-1}(\theta)) = \sum_{k=1}^N e_k(\theta)^T W_k e_k(\theta) \quad (2.0.2)$$

per opportune matrici di peso (simmetriche e definite positive) W_k degli errori di predizione $e_k = y_k - \hat{y}_{k|k-1}(\theta)$. Per valutare le predizioni ad un passo $\{\hat{y}_{k|k-1}(\theta)\}_{k=1}^N$ e, conseguentemente, il costo $J(\theta, y_{1:N})$ da minimizzare, si fa uso di un assimilatore dati che, fissati i parametri θ , assimila i dati y_1, y_2, \dots, y_N per produrre in uscita le innovazioni (errori di predizione) $e_1(\theta), e_2(\theta), \dots, e_N(\theta)$ dove

$$e_k(\theta) = y_k - \hat{y}_{k|k-1}(\theta) = y_k - h(t_k, \hat{x}_{k|k-1}(\theta)) \quad (2.0.3)$$

Ricordando che l'assimilatore dati genera anche le matrici $S_k(\theta)$ che quantificano l'incertezza delle innovazioni $e_k(\theta)$, una possibile scelta naturale delle matrici di peso è $W_k = S_k^{-1}(\theta)$. Pertanto, secondo il metodo di minimizzazione dell'errore di predizione

CAPITOLO 2. CALIBRAZIONE DEI PARAMETRI MEDIANTE ASSIMILAZIONE DATI 12

(PEM = Prediction Error Minimization) il vettore dei parametri stimati viene determinato tramite

$$\hat{\theta} = \arg \min_{\theta \in \mathbb{R}^{n_\theta}} J(\theta, y_{1:N}) \quad (2.0.4)$$

utilizzando una routine di ottimizzazione globale del costo (2.0.2) basata esclusivamente sulla valutazione del costo in quanto, nella maggior parte dei casi, il gradiente e l'Hessiano rispetto a θ del costo $J(\theta, y_{1:N})$ sono di difficile, se non impossibile, determinazione.

Bibliografia

- [1] G. Evensen: *Data assimilation: the ensemble Kalman filter*, Springer, New York, 2007.
- [2] G. Evensen: “The ensemble Kalman filter for combined state and parameter estimation”, *IEEE Control Systems Magazine*, vol. 29, no. 3, pp. 83-104, 2009.