École Doctorale des Sciences de l'Environnement d'Ile-de-France
Année 2021-2022

Cours
Modélisation Numérique de l'Écoulement Atmosphérique
et Assimilation d'Observations
(Olivier Talagrand, mars-mai 2022)

Formulaire Récapitulatif

En règle générale, les caractères gras ($\boldsymbol{x}$, $\boldsymbol{y}$, …) désignent un vecteur, les caractères simples majuscules (P, H, …) une matrice. La notation $E(.)$ désigne une moyenne statistique. Pour toute quantité aléatoire u (scalaire ou vecteur), u' désigne la quantité aléatoire centrée correspondante : $u' \equiv u - E(u)$. L'indice supérieur ^{T} désigne la transposition (d'un vecteur ou d'une matrice). La matrice de covariance $E(\boldsymbol{x}'\boldsymbol{y}'^T)$ de deux vecteurs aléatoires $\boldsymbol{x}$ et $\boldsymbol{y}$ est notée C_{xy}.

1. *Interpolation Optimale*. On cherche à évaluer un vecteur aléatoire $\boldsymbol{x}$ à partir de la valeur connue d'un vecteur aléatoire $\boldsymbol{y}$. La fonction affine $\boldsymbol{x}^a$ de $\boldsymbol{y}$ qui minimise la moyenne statistique $E[(\boldsymbol{x}^a - \boldsymbol{x})^T (\boldsymbol{x}^a - \boldsymbol{x})]$ de l'erreur quadratique d'estimation est égale à :

$$\boldsymbol{x}^a = E(\boldsymbol{x}) + C_{xy} [C_{yy}]^{-1} \boldsymbol{y}' \qquad (1a)$$

L'erreur d'estimation correspondante $\boldsymbol{x}^a - \boldsymbol{x}$ est nulle en moyenne statistique. Sa matrice de covariance est égale à :

$$P^a \equiv E[(\boldsymbol{x}^a - \boldsymbol{x}) (\boldsymbol{x}^a - \boldsymbol{x})^T] = C_{xx} - C_{xy} [C_{yy}]^{-1} C_{yx} \qquad (1b)$$

Quand la distribution de probabilité du couple ($\boldsymbol{x}$, $\boldsymbol{y}$) est gaussienne, la distribution de probabilité de $\boldsymbol{x}$ connaissant $\boldsymbol{y}$, soit $P(\boldsymbol{x} | \boldsymbol{y})$, est la distribution gaussienne de moyenne $\boldsymbol{x}^a$ et de matrice de covariance P^a, notée $\mathcal{N}[\boldsymbol{x}^a, P^a]$.

2. *Forme générale de l'Estimation Statistique Linéaire*. *Best Linear Unbiased Estimator* (*BLUE*).

On considère ici le cas général où l'on cherche à estimer un vecteur $\underline{x}$, appartenant à l'*espace des états*, de dimension n, à partir d'un *vecteur de données* $\boldsymbol{z}$ de dimension m ($m \geq n$, $m = n + p$). Le vecteur de données est de la forme :

$$z = \Gamma \boldsymbol{x} + \boldsymbol{\zeta} \tag{2}$$

où Γ est un opérateur connu de l'espace des états dans l'espace des données, supposé de rang n, et où $\boldsymbol{\zeta}$ est une 'erreur' aléatoire centrée, de matrice de covariance connue S.

On cherche un estimé $\boldsymbol{x}^a$, fonction affine de z, vérifiant les deux conditions suivantes :

a) $\boldsymbol{x}^a$ est invariant dans un changement d'origine dans l'espace des états.

b) $\boldsymbol{x}^a$ minimise la moyenne statistique $E[(\boldsymbol{x}^a - \boldsymbol{x})^{\mathrm{T}} (\boldsymbol{x}^a - \boldsymbol{x})]$ de l'erreur quadratique d'estimation correspondante.

L'estimé ainsi défini est égal à :

$$\boldsymbol{x}^a = (\Gamma^{\mathrm{T}} S^{-1} \Gamma)^{-1} \Gamma^{\mathrm{T}} S^{-1} z \tag{3a}$$,

la matrice de covariance de l'erreur d'estimation correspondante étant égale à :

$$P^a \equiv E[(\boldsymbol{x}^a - \boldsymbol{x}) (\boldsymbol{x}^a - \boldsymbol{x})^{\mathrm{T}}] = (\Gamma^{\mathrm{T}} S^{-1} \Gamma)^{-1} \tag{3b}$$

Forme variationnelle. L'estimé $\underline{x}^a$ peut aussi être obtenu par minimisation de la fonction scalaire suivante (*fonction objective*), définie sur l'espace des états :

$$\boldsymbol{\xi} \rightarrow \mathcal{J}(\boldsymbol{\xi}) = (1/2) (\Gamma\boldsymbol{\xi} - z)^{\mathrm{T}} S^{-1} (\Gamma\boldsymbol{\xi} - z) \tag{4a}$$

De plus, la matrice P^a est l'inverse du hessien (matrice des dérivées secondes) de $\mathcal{J}$ par rapport à $\boldsymbol{\xi}$

$$P^a = [\partial^2\mathcal{J} / \partial\boldsymbol{\xi}^2]^{-1} \tag{4b}$$

L'estimé $\boldsymbol{x}^a$ est invariant dans tout changement de coordonnées linéaire inversible dans l'espace des états ou dans l'espace des données. De façon équivalente, il est indépendant du choix d'une norme, ou d'un produit scalaire, dans l'un ou l'autre de ces espaces. En particulier, il minimise l'erreur quadratique d'estimation $E[(x_i^a - x_i)^2]$ sur toute composante x_i de $\boldsymbol{x}$, et plus généralement l'erreur $E[(l\boldsymbol{x}^a - l\boldsymbol{x})^2]$ sur toute forme linéaire l définie sur l'espace des états. Il minimise aussi l'erreur $E[(\boldsymbol{x}^a - \boldsymbol{x})^{\mathrm{T}} A (\boldsymbol{x}^a - \boldsymbol{x})]$ pour toute matrice A symétrique non-négative.

L'estimé $\boldsymbol{x}^a$ est appelé en anglais le *Best Linear Unbiased Estimator* (*BLUE*) de $\boldsymbol{x}$ à partir de z. Dans le cas particulier où l'erreur $\boldsymbol{\zeta}$ est gaussienne, la distribution de probabilité conditionnelle de $\boldsymbol{x}$ connaissant z est la distribution gaussienne $\mathcal{N}[\boldsymbol{x}^a, P^a]$ de moyenne $\boldsymbol{x}^a$ et de matrice de covariance P^a.

Du fait que $rang\,\Gamma = n$, le vecteur de données peut toujours être transformé en:

$$\boldsymbol{x}^b = \boldsymbol{x} + \boldsymbol{\zeta}^b \tag{5a}$$
$$\boldsymbol{y} = H\boldsymbol{x} + \boldsymbol{\varepsilon} \tag{5b}$$

avec : $E(\boldsymbol{\zeta}^b \boldsymbol{\varepsilon}^{\mathrm{T}}) = 0$ (6)

Dans les expressions ci-dessus, le vecteur $\boldsymbol{x}^b$, de dimension n, est l'*ébauche* du vecteur d'état à estimer, à laquelle s'ajoute le vecteur de données complémentaires $\boldsymbol{y}$, de dimension p.

Posant :

$$E(\boldsymbol{\zeta}^b \boldsymbol{\zeta}^{b\mathrm{T}}) = P^b \tag{7a}$$
$$E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}) = R \tag{7b}$$,

les équations (3) s'écrivent sous les deux formes équivalentes :

$$\boldsymbol{x}^a = \boldsymbol{x}^b + P^a H^{\mathrm{T}} R^{-1} (\boldsymbol{y} - H\boldsymbol{x}^b) \tag{8a}$$
$$(P^a)^{-1} = (P^b)^{-1} + H^{\mathrm{T}} R^{-1} H \tag{8b}$$

et :

$$\boldsymbol{x}^a = \boldsymbol{x}^b + P^b H^{\mathrm{T}} (H P^b H^{\mathrm{T}} + R)^{-1} (\boldsymbol{y} - H\boldsymbol{x}^b) \tag{9a}$$
$$P^a = P^b - P^b H^{\mathrm{T}} (H P^b H^{\mathrm{T}} + R)^{-1} H P^b \tag{9b}$$

Dans les équations (8a) et (9a), l'estimé $\boldsymbol{x}^a$ est décomposé en la somme de l'ébauche $\boldsymbol{x}^b$ et d'une correction proportionnelle au *vecteur innovation* $\boldsymbol{y} - H\boldsymbol{x}^b$.

Dans la décomposition (5), la fonction objective (4a) devient :

$$\boldsymbol{\xi} \rightarrow \mathcal{J}(\boldsymbol{\xi}) = (1/2)\,(\boldsymbol{\xi} - \boldsymbol{x}^b)^{\mathrm{T}} (P^b)^{-1} (\boldsymbol{\xi} - \boldsymbol{x}^b) + (1/2)\,(H\boldsymbol{\xi} - \boldsymbol{y})^{\mathrm{T}} R^{-1} (H\boldsymbol{\xi} - \boldsymbol{y}) \tag{10}$$

3. *Système évolutif dans le temps. Filtre de Kalman.*

On va maintenant mettre en œuvre ce qui précède dans les circonstances où le système observé évolue au cours du temps, suivant la loi :

$$\boldsymbol{x}_{k+1} = M_k \boldsymbol{x}_k + \boldsymbol{\eta}_k \tag{11}$$

où $\boldsymbol{x}_k$ est l'état du système à l'instant k ($\dim \boldsymbol{x}_k = n$), M_k est une matrice connue, et $\boldsymbol{\eta}_k$ est l'*erreur modèle*, aléatoire, décorrélée dans le temps de covariance connue Q_k à l'instant k.

On dispose aux instants $k = 0, 1, 2, \ldots$ d'observations de la forme

$$\boldsymbol{y}_k = H_k \boldsymbol{x}_k + \boldsymbol{\varepsilon}_k \tag{12}$$

où H_k est une matrice connue, et $\boldsymbol{\varepsilon}_k$ une erreur d'observation aléatoire, décorrélée dans le temps, de covariance connue R_k à l'instant k. On dispose en outre d'une ébauche à l'instant k = 0, de la forme

$$\boldsymbol{x}_0^{\;b} \;=\; \boldsymbol{x}_0 + \boldsymbol{\zeta}_0^{\;b} \tag{13}$$

où l'erreur $\zeta_0^{\;b}$ est de covariance $P_0^{\;b}$. On suppose enfin que les erreurs $\boldsymbol{\eta}_k$, $\boldsymbol{\varepsilon}_k$ et $\boldsymbol{\zeta}_0^{\;b}$ sont mutuellement décorrélées.

Le *filtre de Kalman* est défini par l'algorithme récursif suivant, initialisé à partir de $(\boldsymbol{x}_0^b, P_0^b)$:

Analyse à l'instant k :

$$\boldsymbol{x}_k^{\;a} \;=\; \boldsymbol{x}_k^{\;b} + P_k^{\;b}\, H_k^{\;\mathrm{T}}\, (H_k\, P_k^{\;b}\, H_k^{\;\mathrm{T}} + R_k)^{-1}\, (\boldsymbol{y}_k - H_k \boldsymbol{x}_k^{\;b}) \tag{14a}$$
$$P_k^{\;a} \;= P_k^{\;b} - P_k^{\;b}\, H_k^{\;\mathrm{T}} (H_k\, P_k^{\;b}\, H_k^{\;\mathrm{T}} + R_k)^{-1}\, H_k\, P_k^{\;b} \tag{14b}$$

Évolution temporelle entre les instants k et k+1 :

$$\boldsymbol{x}_{k+1}^{\;\;b} \;=\; M_k \boldsymbol{x}_k^{\;a} \tag{15a}$$
$$P_{k+1}^{\;\;b} \;=\; M_k P_k^{\;a} M_k^{\;\mathrm{T}} + Q_k \tag{15b}$$

L'estimé $\boldsymbol{x}_k^{\;b}$ (resp. $\boldsymbol{x}_k^{\;a}$) ainsi obtenu est le *BLUE* de l'état $\boldsymbol{x}_k$ à l'instant k à partir de $\boldsymbol{x}_0^{\;b}$ et de toutes les observations jusqu'à l'instant k-1 (resp. jusqu'à l'instant k). La matrice $P_k^{\;b}$ (resp. $P_k^{\;a}$) est la matrice de covariance d'erreur correspondante. Dans le cas où les erreurs ($\boldsymbol{\eta}_k$, $\boldsymbol{\varepsilon}_k$ et $\boldsymbol{\zeta}_0^{\;b}$) sont globalement gaussiennes, le filtre de Kalman est bayésien, en ce sens que $\mathrm{P}(\boldsymbol{x}_k \,|\, \boldsymbol{x}_0^{\;b}, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_{k-1}) = \mathcal{N}[\boldsymbol{x}_k^{\;b}, P_k^{\;b}]$ et que $\mathrm{P}(\boldsymbol{x}_k \,|\, \boldsymbol{x}_0^{\;b}, \boldsymbol{y}_0, \ldots, \boldsymbol{y}_{k-1}, \boldsymbol{y}_k) = \mathcal{N}[\boldsymbol{x}_k^{\;a}, P_k^{\;a}]$.

Forme variationnelle. Nous limitant maintenant à la suite des K+1 instants k = 0, 1, …, K, la minimisation de la fonction objective :

$$\begin{aligned}
&(\boldsymbol{\xi}_0, \boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_K) \rightarrow \\
&\mathcal{J}(\boldsymbol{\xi}_0, \boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_K) \;=\; (1/2)\, (\boldsymbol{\xi}_0 - \boldsymbol{x}_0^{\;b})^{\mathrm{T}}\, [P_0^{\;b}]^{-1}\, (\boldsymbol{\xi}_0 - \boldsymbol{x}_0^{\;b}) + \; (1/2)\, \Sigma_{k=0,K} (H_k \boldsymbol{\xi}_k - \boldsymbol{y}_k)^{\mathrm{T}} R_k^{\;-1}\, (H_k \boldsymbol{\xi}_k - \boldsymbol{y}_k) \\
&+\; (1/2)\, \Sigma_{k=0,\ldots,K-1}\, (\boldsymbol{\xi}_{k+1} - M_k \boldsymbol{\xi}_k)^{\mathrm{T}}\, Q_k^{\;-1}\, (\boldsymbol{\xi}_{k+1} - M_k \boldsymbol{\xi}_k)
\end{aligned} \tag{16}$$

où les $\boldsymbol{\xi}_k$ (k=0, …, K) appartiennent à l'espace des états du système, définit la suite des *BLUE* des $\boldsymbol{x}_k$ à partir de toutes les données (11-13). En particulier, l'état ainsi obtenu à l'instant final K est identique à l'estimé $\boldsymbol{x}_N^{\;a}$ produit par le filtre de Kalman (14-15). Dans le cas où les erreurs sont globalement gaussiennes, la minimisation de (16) produit l'estimation bayésienne correspondante.