

MA201 - Séance 5

Filtrage de Kalman

H. Piet-Lahanier - L. Meyer



Contenu du cours

- Rappels sur l'algorithme des moindres carrés récurrents
- Cadre du filtrage de Kalman
- Les équations du filtre
- Exemples d'applications
- Résultats de convergence et d'optimalité
- Extensions

Content

- 1 Rappels sur l'algorithme des moindres carrés récursifs**
- 2 Cadre du filtre de Kalman
- 3 Les équations du filtre
- 4 Propriétés du filtre de Kalman
- 5 Extension pour des systèmes non linéaires

Exemple introductif

Dans ce cours, nous introduisons un aspect dynamique de l'estimation.

Exemple

Un industriel reçoit un nouveau lot de pièces à tester toutes les heures. Il souhaite estimer un paramètre sur l'ensemble de ces pièces.

- Il peut attendre la fin de la journée et considérer l'ensemble total des pièces pour estimer le paramètre inconnu.
=> méthode globale. Exemple : inversion de l'équation normale.
- Il peut souhaiter obtenir rapidement une première estimation et l'affiner au fur et à mesure qu'il reçoit de nouveaux lots de pièces.
=> méthode récursive. Exemple : algorithme des moindres carrés récursifs.

Nous considérerons une succession d'instants t_0, \dots, t_N . Nous noterons pour tout signal x , $x_k = x(t = t_k)$, la valeur du signal x à l'instant t_k .

Moindres carrés récurrents

Cadre : Modèle Linéaire Gaussien (Rappel)

Considérons une succession de mesures z_k avec $k \geq 0$:

$$z_k = H_k \theta + v_k, \quad (1)$$

avec $\theta \in \mathbb{R}^{n_\theta}$ le vecteur de paramètres inconnus, $H_k \in \mathcal{M}_{n_z, n_\theta}(\mathbb{R})$ une matrice à coefficients connus, et $v_k \sim \mathcal{N}(0, V_k)$, avec $V_k \in \mathcal{M}_{n_z}(\mathbb{R})$ matrice de covariance connue.

Hypothèses (Rappel)

Nous supposons les v_k tous mutuellement indépendants. C'est donc aussi le cas des z_k .

Moindres carrés récursifs (cont.)

Algorithme des moindres carré récursifs (Rappel)

1 Initialisation

Choix de $\hat{\theta}_0$ et de P_0 la matrice de covariance de l'estimation à l'instant initial. Sans aucune connaissance *a priori* sur θ , on choisira $\hat{\theta}_0 = 0$ et $P_0 = \alpha I$ (avec α arbitrairement grand).

2 Itération

Passage de k à $k + 1$

$$\begin{cases} K_{k+1} &= P_k H_{k+1}^T (H_{k+1} P_k H_{k+1}^T + V_k)^{-1} \\ \hat{\theta}_{k+1} &= \hat{\theta}_k + K_{k+1} (z_{k+1} - H_{k+1} \hat{\theta}_k) \\ P_{k+1} &= (I - K_{k+1} H_{k+1}) P_k \end{cases}, \quad (2)$$

Propriété d'optimalité (Rappel)

Pour tout $k \geq 0$, l'estimateur sans biais à minimum de variance de θ , est l'estimateur du maximum de vraisemblance (sachant les mesures jusqu'à l'instant k) est $\hat{\theta}_k$, et suit une loi normale $\mathcal{N}(\theta, P_k)$, P_k donné par l'algorithme précédent.

Content

- 1 Rappels sur l'algorithme des moindres carrés récursifs
- 2 Cadre du filtre de Kalman**
- 3 Les équations du filtre
- 4 Propriétés du filtre de Kalaman
- 5 Extension pour des systèmes non linéaires

Exemple introductif

Supposons un satellite en orbite autour de la Terre. Nous souhaitons estimer sa position et sa vitesse à tout instant.

Première approche

- Idée : si l'on connaît sa position et sa vitesse à l'instant initial, et les équations de la mécanique qui régissent son comportement, on peut en déduire sa position et vitesse à tout instant.
=> Intéressant pour une estimation sur de courtes périodes.
- Limite : il existe des perturbations qui rendent approximatives les équations de la mécanique en question (valeurs du champ magnétique en un point de l'espace, flux solaire, etc.).
=> Dérive sur des temps importants.

Exemple introductif (cont.)

Deuxième approche

- Idée : si l'on arrive à prendre des mesures (via radar ou télescope) de la position du satellite à intervalles réguliers, on peut utiliser l'algorithme des moindres carrés récursifs à chaque pas de temps.
=> Pas de dérive possible.
- Limite : intermittence des mesures, précision limitée par les caractéristiques des instruments de mesures, et non utilisation d'une connaissance *gratuite* (les équations de la mécanique céleste).

Intuition

Le filtre de Kalman fusionne les mesures et le modèle physique du système étudié.

Cadre du filtre de Kalman

Plaçons nous dans le cas où le paramètre θ n'est plus constant, mais qu'il évolue au cours du temps. On l'appellera *état* du système étudié et on le notera x_k .

Cadre

Considérons le système d'équations suivant :

$$\begin{cases} x_{k+1} = F_k x_k + w_k & \text{Equation de la dynamique} \\ z_k = H_k x_k + v_k & \text{Equation de mesures} \end{cases}, \quad (3)$$

avec $x_k \in \mathbb{R}^{n_x}$ le vecteur *inconnu* d'état du système au pas de temps k , z_k une mesure disponible au pas de temps k , w_k le bruit de modèle (ou d'état) suivant la loi normale $\mathcal{N}(0, W_k)$ avec W_k matrice connue, et v_k le bruit de mesures suivant la loi normale $\mathcal{N}(0, V_k)$ avec V_k matrice connue. Enfin, $F_k \in \mathcal{M}_{n_x, n_x}(\mathbb{R})$ et $H_k \in \mathcal{M}_{n_z, n_x}(\mathbb{R})$ sont deux matrices connues.

Hypothèses

- Pour tous $k, l, m, n \geq 0$, $k \neq l$, $m \neq n$, les bruits w_k , w_l , v_m et v_n sont mutuellement indépendants.
- La matrice de covariance V_k est inversible pour tout $k \geq 0$.

Cadre du filtre de Kalman (cont.)

Exemples

Un grand nombre de phénomènes sont représentés par de telles équations

- Aérospatial : équations de Képler pour la dynamique d'un satellite, associées à des mesures radar.
- Finance : dynamique de cours financiers, associée à des mesures de volatilité d'actifs.
- Météorologie : dynamique des courants atmosphériques, associée à des relevés météo.

Objectif

Obtenir à chaque pas de temps $k \geq 0$ une estimation \hat{x}_k non biaisée et à minimum de variance de x_k , compte-tenu de la connaissance des équations (3) et des mesures z_1, \dots, z_k .

Content

- 1 Rappels sur l'algorithme des moindres carrés récursifs
- 2 Cadre du filtre de Kalman
- 3 Les équations du filtre**
- 4 Propriétés du filtre de Kalman
- 5 Extension pour des systèmes non linéaires

Filtre de Kalman

Algorithme

Le filtre de Kalman est un algorithme répondant à la problématique précédente. Outre l'initialisation, chaque itération se décompose en deux étapes : une étape de prédiction, et une étape de correction

1 Initialisation

On suppose que $x_0 \sim \mathcal{N}(\hat{x}_0, P_0)$. Dans la pratique, sans connaissance *a priori* sur x_0 , on prend une estimation \hat{x}_0 de x_0 arbitraire parmi les valeurs possibles de x_0 , et une matrice de covariance arbitrairement grande (par exemple $P_0 = \alpha I$ avec α arbitrairement grand).

2 Itération

■ Prédiction

Utilisation de l'estimation \hat{x}_k , et de l'équation d'état pour obtenir une première estimation $\hat{x}_{k+1|k}$ de x_{k+1} , et la matrice de covariance associée $P_{k+1|k}$.

■ Correction

Utilisation de la nouvelle mesure disponible z_{k+1} , et de l'équation de mesure pour affiner l'estimation $\hat{x}_{k+1|k}$ de x_{k+1} , et obtenir ainsi l'estimation \hat{x}_{k+1} , et la matrice de covariance associée P_{k+1} .

Algorithme

Algorithme

1 Initialisation

\hat{x}_0, P_0 . Par exemple : $\hat{x}_0 = 0$, et $P_0 = \alpha I$, avec α arbitrairement grand.

2 Pour $k \geq 0$:

■ Prédiction

$$\begin{cases} \hat{x}_{k+1|k} &= F_k \hat{x}_k \\ P_{k+1|k} &= F_k P_k F_k^T + W_k \end{cases} \quad (4)$$

■ Correction

Calcul du gain de Kalman :

$$K_{k+1} = P_{k+1|k} H_{k+1}^T (H_{k+1} P_{k+1|k} H_{k+1}^T + V_{k+1})^{-1} \quad (5)$$

Mise à jour :

$$\begin{cases} \hat{x}_{k+1} &= \hat{x}_{k+1|k} + K_{k+1} (z_{k+1} - H_{k+1} \hat{x}_{k+1|k}) \\ P_{k+1} &= (I - K_{k+1} H_{k+1}) P_{k+1|k} \end{cases} \quad (6)$$

Algorithme (cont.)

Remarque

Dans les équations du filtre de Kalman, la suite de matrices de covariance P_k ne dépend ni des \hat{x}_k , ni des mesures.

Exemple

Exemple

Considérons un mobile en rotation sur un plan.

Hypothèses :

- Absence de mouvement de translation
- Vitesse de rotation constante

On s'intéresse à l'estimation de son angle de rotation α_k et sa vitesse de rotation ω_k à différents instants $k \geq 0$ séparés d'un pas de temps T .

Ainsi, on a les équations de récursions suivantes :

$$\begin{cases} \omega_{k+1} &= \omega_k \\ \alpha_{k+1} &= T\omega_k + \alpha_k \end{cases} \quad (7)$$

Exemple (cont.)

Instruments de mesures

On dispose des instruments de mesures suivant :

- Un gyromètre donnant une mesure bruitée de la vitesse de rotation :

$$g_k = \omega_k + v_{g_k} \quad (8)$$

Un accéléromètre donnant (via une procédure spécifique) une mesure bruitée de l'angle de rotation :

$$a_k = \alpha_k + v_{a_k} \quad (9)$$

Exemple (cont.)

Mise en équations

Le problème peut donc se mettre en équations.

L'équation de la dynamique s'écrit :

$$\begin{bmatrix} \omega_{k+1} \\ \alpha_k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ T & 1 \end{bmatrix} \begin{bmatrix} \omega_k \\ \alpha_k \end{bmatrix} \quad (10)$$

L'équation de mesures s'écrit :

$$\begin{cases} \begin{bmatrix} g_k \\ a_k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_k \\ \alpha_k \end{bmatrix} + \begin{bmatrix} v_{gk} \\ v_{ak} \end{bmatrix} \end{cases} \quad (11)$$

=> Le filtre de Kalman peut donc être appliqué sur ce modèle.

Autre exemple

Prise en compte du biais du gyromètre

Considérons l'exemple précédent, avec la différence suivante : au fil du temps, un biais de mesure b peut apparaître au niveau du gyromètre. L'équation de mesure associée devient alors :

$$g_k = \omega_k + b_k + v_{g_k}. \quad (12)$$

On va donc intégrer ce biais au vecteur d'état du système.

Ce biais sera supposé évoluer très lentement du cours du temps (évolution quasi-statique). Il pourra donc être modélisé par :

$$b_{k+1} = b_k + w_k, \quad (13)$$

où w_k est un bruit suivant une loi normale $\mathcal{N}(0, \sigma^2)$, avec σ assez faible (donné par l'expérience).

Autre exemple (cont.)

Mise en équations

Ainsi, l'équation de la dynamique peut se réécrire :

$$\begin{bmatrix} \omega_{k+1} \\ \alpha_{k+1} \\ b_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ T & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_k \\ \alpha_k \\ b_k \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ w_k \end{bmatrix} \quad (14)$$

et l'équation de mesures s'écrit :

$$\begin{Bmatrix} g_k \\ a_k \end{Bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \omega_k \\ \alpha_k \\ b_k \end{bmatrix} + \begin{bmatrix} v_{gk} \\ v_{ak} \end{bmatrix} \quad (15)$$

=> Ainsi, sans utiliser d'instruments de mesures supplémentaires, le filtre de Kalman peut être appliqué sur ce modèle et permettre d'estimer outre les variables α , ω , le biais du gyromètre b .

Content

- 1 Rappels sur l'algorithme des moindres carrés récursifs
- 2 Cadre du filtre de Kalman
- 3 Les équations du filtre
- 4 Propriétés du filtre de Kalman**
- 5 Extension pour des systèmes non linéaires

Résultats

Malgré son caractère récursif, le filtre de Kalman présente une propriété d'optimalité globale.

Théorème d'optimalité

Si, outre les hypothèses précédentes (indépendance des bruits et inversibilité de V_k), l'on suppose que $x_0 \sim \mathcal{N}(\hat{x}_0, P_0)$, alors :

- A l'instant k , on a $x_k \sim \mathcal{N}(\hat{x}_k, P_k)$.
- De plus, \hat{x}_k est l'estimateur sans biais à minimum de variance sachant les mesures z_1, \dots, z_k :
 - il est sans biais :

$$E[\hat{x}_k - x_k] = 0, \quad (16)$$

- il est à minimum de variance (parmi tous les estimateurs linéaires sans biais) :

$$\hat{x}_k = \arg \min_{\tilde{x}_k \text{ est lin.} | E[\tilde{x}_k - x_k] = 0} \{E[(\tilde{x}_k - x_k)^T (\tilde{x}_k - x_k) | z_1, \dots, z_k]\}, \quad (17)$$

(espérance sur la loi *a posteriori*, i.e. la loi de densité $f(\cdot | z_1, \dots, z_k)$).

Preuve.

Résultats (cont.)

- Montrons par récurrence que, pour tout $k \geq 0$, $x_k \sim \mathcal{N}(\hat{x}_k, P_k)$.
 - L'initialisation de la récurrence est donnée par l'hypothèse $x_0 \sim \mathcal{N}(\hat{x}_0, P_0)$ de l'énoncé.
 - Soit un certain $k \geq 0$. Supposons que $x_k \sim \mathcal{N}(\hat{x}_k, P_k)$, et montrons que $x_{k+1} \sim \mathcal{N}(\hat{x}_{k+1}, P_{k+1})$.
 Notons $e_k = x_k - \hat{x}_k$. Nous avons donc $e_k \sim \mathcal{N}(0, P_k)$, et nous devons montrer que $e_{k+1} \sim \mathcal{N}(0, P_{k+1})$.
 La dynamique de e_k s'écrit :

$$\begin{aligned}
 e_{k+1} &= x_{k+1} - \hat{x}_{k+1} \\
 &= F_k x_k + w_k - (F_k \hat{x}_k + K_{k+1}(z_{k+1} - H_k F_k \hat{x}_k)) \\
 &= F_k x_k + w_k - (F_k \hat{x}_k + K_{k+1}(H_{k+1} x_{k+1} + v_{k+1} - H_k F_k \hat{x}_k)) \\
 &= F_k x_k + w_k - (F_k \hat{x}_k + K_{k+1}(H_{k+1}(F_k x_k + w_k) + v_{k+1} - H_k F_k \hat{x}_k)) \\
 &= (F_k - K_{k+1} H_{k+1} F_k) e_k + (K_{k+1} H_{k+1} - I) w_k + K_{k+1} v_{k+1}.
 \end{aligned} \tag{18}$$

Ainsi, puisqu'on a $E[w_k] = 0$ et $E[v_k] = 0$, l'espérance de e_{k+1} peut s'écrire en fonction de celle de e_k :

$$E[e_{k+1}] = (F_k - K_{k+1} H_{k+1} F_k) E[e_k] \tag{19}$$

Résultats (cont.)

De même, la matrice de covariance $Cov(e_{k+1}) = E[e_{k+1}e_{k+1}^T]$ de e_{k+1} peut s'écrire en fonction de la matrice de covariance $Cov(e_k) = E[e_k e_k^T]$ de e_k :

$$Cov(e_{k+1}) = (F_k - K_{k+1}H_{k+1}F_k)Cov(e_k)(F_k - K_{k+1}H_{k+1}F_k)^T + (K_{k+1}H_{k+1} - I)W_k(K_{k+1}H_{k+1} - I)^T + K_{k+1}V_kK_{k+1}^T. \quad (20)$$

En utilisant l'hypothèse de récurrence ($E[e_k] = 0$ et $Cov(e_k) = P_k$), on en déduit que :

$$E[e_{k+1}] = 0, \quad (21)$$

et que

$$Cov(e_{k+1}) = (F_k - K_{k+1}H_{k+1}F_k)P_k(F_k - K_{k+1}H_{k+1}F_k)^T + (K_{k+1}H_{k+1} - I)W_k(K_{k+1}H_{k+1} - I)^T + K_{k+1}V_kK_{k+1}^T = P_{k+1}. \quad (22)$$

Ce qui achève la preuve de la première assertion du théorème.

Résultat de convergence

Définition

Le cas LTI (Linear Time Invariant) correspond au cas où les matrices F_k , H_k , W_k et V_k ne dépendent pas de k (et sont donc constantes). Elles sont alors notées respectivement F , H , W , et V .

Théorème

Dans le cas LTI, si l'on suppose de plus (condition d'observabilité) :

$$\text{rang} \begin{pmatrix} H \\ HF \\ \vdots \\ HF^{n_x-1} \end{pmatrix} = n_x, \quad (23)$$

(rappel n_x : taille du vecteur x_k), alors la suite P_k converge vers la matrice P solution de :

$$P = FPF^T - FPH^T(HPH^T + V)^{-1}HPF^T + W. \quad (24)$$

Content

- 1 Rappels sur l'algorithme des moindres carrés récursifs
- 2 Cadre du filtre de Kalman
- 3 Les équations du filtre
- 4 Propriétés du filtre de Kalman
- 5 Extension pour des systèmes non linéaires**

Extension aux systèmes non linéaires : le Filtre de Kalman Etendu (EKF)

Cadre

Considérons le système d'équations suivantes :

$$\begin{cases} x_{k+1} &= f(x_k) + w_k \\ z_k &= h(x_k) + v_k \end{cases} \quad (25)$$

avec f et h deux fonctions différentiables connues, et pour tout $k \geq 0$, w_k et v_k des bruits d'état et de mesures, suivant respectivement les lois normales $\mathcal{N}(0, W_k)$ et $\mathcal{N}(0, V_k)$.

Hypothèses

- Les fonctions f et h sont C^1 sur l'ensemble des valeurs possibles pour x_k .
- Pour tous $k, l, m, n \geq 0$, $k \neq l$, $m \neq n$, les bruits w_k , w_l , v_m et v_n sont mutuellement indépendants.
- La matrice V est inversible.

Extension aux systèmes non linéaires : le Filtre de Kalman Étendu (EKF) (cont.)

Filtre de Kalman étendu

1 Initialisation de \hat{x}_0 et P_0 .

2 Pour $k \geq 0$:

■ Prédiction

$$\begin{cases} \hat{x}_{k+1|k} &= f(\hat{x}_k) \\ P_{k+1|k} &= F_k P_k F_k^T + W_k \end{cases}, \quad (26)$$

avec $F_k = \frac{\partial f(x)}{\partial x} \big|_{x=\hat{x}_k}$ (matrice jacobienne de f prise en \hat{x}_k).

■ Correction

$$\begin{cases} K_{k+1} &= P_{k+1|k} H_{k+1}^T (H_{k+1} P_{k+1|k} H_{k+1}^T + V_k)^{-1} \\ \hat{x}_{k+1} &= \hat{x}_{k+1|k} + K_{k+1} (z_{k+1} - h(\hat{x}_{k+1|k})) \\ P_{k+1} &= (I - K_{k+1} H_{k+1}) P_{k+1|k} \end{cases}, \quad (27)$$

avec $H_k = \frac{\partial h(x)}{\partial x} \big|_{x=\hat{x}_{k+1|k}}$ (matrice jacobienne de h prise en $\hat{x}_{k+1|k}$).

Extension aux systèmes non linéaires : le Filtre de Kalman Étendu (EKF) (cont.)

Pas de résultat d'optimalité, ni de convergence *en général* pour le filtre de Kalman étendu.