

Filtrage de Kalman

Gilles Chardon

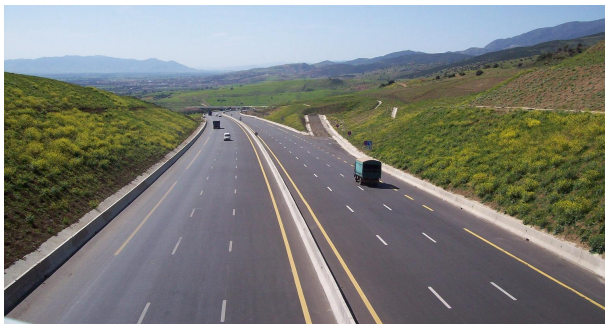
CentraleSupélec

Plan

- 1 Introduction
- 2 Modèle de signaux
- 3 Filtre de Kalman
 - Cas Gaussien
 - Cas général
- 4 Exemples
 - Constante
 - Suivi de pendule
- 5 Cas non linéaire
 - Extended Kalman filter
 - Filtrage particulière
- 6 Conclusion

Une voiture sur l'autoroute

Comment estimer sa position ?

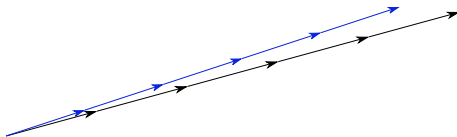


Méthode 1 : modèle physique

On suppose la position initiale et la vitesse connues (navigation inertielle) :

$$x(t) = x(0) + \int_0^t v(\tau) d\tau$$

- estimation valide à court terme
- mais les erreurs s'accroissent, en particulier le biais

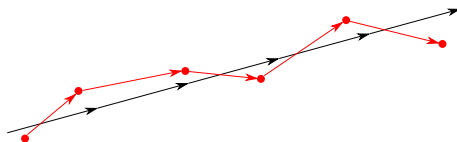


Méthode 2 : méthode de localisation

Ex: GPS

À chaque instant t_n , une estimation \hat{x}_n de la position $x(t_n)$ est disponible.

- estimation toujours fautive, mais pertinente “en moyenne”
- estimation de la vitesse très imprécise (problème mal conditionné)



Filtrage de Kalman

Le filtrage de Kalman fusionne les mesures et le modèle physique pour construire une meilleure estimation de l'état du système.

Deux manières de l'interpréter :

- recalage du modèle physique à partir des mesures
- ou décomposition des mesures en bruit + composante obéissant au modèle

Remarque : filtrage de Kalman \neq filtres linéaires

Cadre du filtrage de Kalman

En résumé :

- Estimation de l'état d'un système au cours du temps,
- à partir d'observations incomplètes et bruitées,
- connaissant un **modèle d'évolution** du système.

On modélise la connaissance sur l'état du système par une densité de probabilité.

On s'intéresse particulier à l'espérance de l'état ainsi qu'à sa covariance (i.e. les incertitudes).

Application historique

Années 1960, suivi de la trajectoire des modules Apollo

- estimation de la position à partir de mesures terrestres
- connaissance de l'accélération du module (centrale inertielle)

Fusion des deux types de données par le filtrage de Kalman.

Modèle physique = déplacement d'un point dans un champ de gravité et soumis à des forces (moteur).

Plan

- 1 Introduction
- 2 Modèle de signaux**
- 3 Filtre de Kalman
 - Cas Gaussien
 - Cas général
- 4 Exemples
 - Constante
 - Suivi de pendule
- 5 Cas non linéaire
 - Extended Kalman filter
 - Filtrage particulière
- 6 Conclusion

Modèle d'évolution

L'état x_k à estimer est un vecteur à n composantes.

Connaissance d'un modèle **linéaire** d'évolution de l'état

$$x_{k+1} = Fx_k + u_k$$

F matrice $n \times n$ qui modélise l'évolution du système.

u_k bruit centré de covariance connue Q de taille n , qui modélise l'incertitude sur le modèle.

F et Q peuvent varier au cours du temps.

La forme de l'état est un **choix**. Il n'est pas limité aux grandeurs d'intérêt.

Modèle d'évolution : exemple de la voiture

On veut mesurer la position du véhicule :

$$x(t') = x(t) + \int_t^{t'} v(\tau) d\tau$$

Pour écrire le modèle d'évolution, on discrétise et on rajoute la vitesse dans l'état.

Le modèle d'évolution est alors donné par

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \\ v_{k+1}^x \\ v_{k+1}^y \end{pmatrix} = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ v_k^x \\ v_k^y \end{pmatrix} + u_k$$

Modélisation des mesures

On a, à chaque instant, m mesures qui sont obtenues par une transformation linéaire de l'état :

$$y_k = Hx_k + v_k$$

H matrice connue de taille $m \times n$, qui modélise le processus de mesure.

Remarque : m peut être plus petit que n , i.e. moins de mesures que la dimension de l'état.

v_k bruit centré de covariance connue R de taille m .

H et R peuvent varier au cours du temps.

Modélisation des mesures : exemple de la voiture

Le GPS donne une mesure de la position :

$$\begin{pmatrix} \tilde{x}_k \\ \tilde{y}_k \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ v_k^x \\ v_k^y \end{pmatrix} + v_k$$

Bruit de modèle et bruit de mesure

$$\begin{cases} x_{k+1} = Fx_k + u_k \\ y_k = Hx_k + v_k \end{cases}$$

Les bruits u_k et v_k modélisent la confiance dans le modèle et les mesures.

- petit bruit de modèle : on voudra faire confiance au modèle et peu aux mesures,
- petit bruit de mesure : on voudra donner une influence plus importante aux mesures.

Le filtre de Kalman va automatiquement donner un poids plus ou moins fort au modèle ou aux mesures pour l'estimation de l'état.

Il donne à la fois une estimation de l'état, et la covariance de l'erreur d'estimation.

Plan

- 1 Introduction
- 2 Modèle de signaux
- 3 Filtre de Kalman**
 - Cas Gaussien
 - Cas général
- 4 Exemples
 - Constante
 - Suivi de pendule
- 5 Cas non linéaire
 - Extended Kalman filter
 - Filtrage particulière
- 6 Conclusion

Idée générale du filtre de Kalman

Estimer l'état à l'instant $k + 1$ à partir d'une estimation à l'instant k et des mesures à l'instant $k + 1$.

Estimation en deux étapes :

- prédiction de l'état à partir du modèle d'évolution
- correction de la prédiction à partir des mesures

Deux hypothèses possibles :

- les bruits sont gaussiens
- cas général, en se limitant à un algorithme linéaire et à l'estimation de la moyenne et la covariance de l'état

Les deux hypothèses aboutissent au même résultat

Notations

x_k état à l'instant k .

y_k mesures à l'instant k .

$\bar{x}_{k,k}$ estimation de l'état k à partir des données jusque l'instant k

covariance $P_{k,k} = E((\bar{x}_{k,k} - x_k)(\bar{x}_{k,k} - x_k)^*)$

$\bar{x}_{k+1,k}$ estimation de l'état $k + 1$ à partir des données jusque l'instant k

covariance $P_{k+1,k} = E((\bar{x}_{k+1,k} - x_k)(\bar{x}_{k+1,k} - x_k)^*)$

$Q = E(uu^*)$ covariance du bruit de modèle

$R = E(vv^*)$ covariance du bruit de mesure

Cas gaussien

On cherche à déterminer à chaque instant :

- la distribution de probabilité de l'état au temps k connaissant les mesures jusqu'au temps k :

$$p(x_k | y_{1:k})$$

- la distribution de probabilité de l'état au temps $k + 1$ connaissant les mesures jusqu'au temps k (prédiction) :

$$p(x_{k+1} | y_{1:k})$$

On va montrer par récurrence que si la distribution de x_0 est gaussienne, les $x_k | y_{1:k}$ et $x_{k+1} | y_{1:k}$ sont également gaussiennes

Cas gaussien - prédiction

La distribution de x_k sachant les mesures jusqu'à l'instant k est normale (hypothèse de récurrence) :

$$x_k | y_{1:k} \sim \mathcal{N}(\bar{x}_{k,k}, P_{k,k})$$

Avec $x_{k+1} = Fx_k + u_k$, on obtient directement que $x_{k+1} | y_{1:k}$ suit une loi normale comme somme de deux variables normales :

$$x_{k+1} | y_{1:k} \approx \mathcal{N}(\bar{x}_{k+1,k}, P_{k+1,k})$$

avec

$$\bar{x}_{k+1,k} = E(Fx_k + u_k | y_{1:k}) = F\bar{x}_{k,k}$$

et

$$\begin{aligned} P_{k+1,k} &= E[(x_{k+1} - \bar{x}_{k+1,k})(x_{k+1,k} - \bar{x}_{k+1,k})^* | y_{1:k}] \\ &= FP_{k,k}F^* + Q \end{aligned}$$

Cas gaussien - Correction de la prédiction

On a à disposition une estimation a priori de l'état :

$$x_{k+1}|y_{1:k} \sim \mathcal{N}(\bar{x}_{k+1,k}, P_{k+1,k})$$

et des mesures

$$y_{k+1} = Hx_{k+1} + v_k$$

$$y_{k+1}|x_{k+1} \sim \mathcal{N}(Hx_{k+1}, R)$$

On veut trouver la distribution de

$$x_{k+1}|y_{1:k+1}$$

c'est à dire de l'état connaissant les mesures jusqu'à l'instant $k + 1$.

Théorème de Bayes

On applique le théorème de Bayes :

$$p(a|b) = \frac{p(b|a)p(a)}{p(b)}$$

$$p(x_{k+1}|y_{1:k}, y_{k+1}) \propto p(y_{k+1}|x_{k+1}, y_{1:k})p(x_{k+1}|y_{1:k})$$

On montre après quelques lignes de calcul que

$$x_{k+1}|y_{1:k+1} \sim \mathcal{N}(\bar{x}_{k+1,k+1}, P_{k+1,k+1})$$

avec

- $K_{k+1} = P_{k+1,k} H^* (H P_{k+1,k} H^* + R)^{-1}$
- $\bar{x}_{k+1,k+1} = \bar{x}_{k+1,k} + K_{k+1} (y_{k+1} - H \bar{x}_{k+1,k})$
- $P_{k+1,k+1} = (I - K_{k+1} H) P_{k+1,k}$

Cas non-gaussien

Si les bruits de modèles et de mesures ne sont pas gaussiens, les calculs précédents ne sont pas valables.

On pourrait considérer des distributions de probabilités plus générales voire quelconques.

On se limitera ici à estimer linéairement le vecteur d'état et à évaluer la variance de cette estimation.

Prédiction

Rappel : $x_{k+1} = Fx_k + u_k$

L'estimation a priori est donnée par

$$\hat{x}_{k+1,k} = E(Fx_k + u_k) = F\hat{x}_{k,k}$$

La covariance $P_{k+1,k}$ est donnée par :

$$\begin{aligned} P_{k+1,k} &= E[(x_{k+1} - \bar{x}_{k+1,k})(x_{k+1} - \hat{x}_{k+1,k}|y_{1:k})^*] \\ &= FP_{k,k}F^* + Q \end{aligned}$$

L'estimation a priori $x_{k+1,k}$ est généralement plus incertaine que $x_{k,k}$.

Correction

On définit l'innovation comme la différence entre les mesures observées et les mesures attendues :

$$z_{k+1} = y_{k+1} - H\hat{x}_{k+1,k}$$

La correction apportée à $\hat{x}_{k+1,k}$ est **choisie** sous la forme linéaire $K_{k+1}z_{k+1}$:

$$\hat{x}_{k+1,k+1} = \hat{x}_{k+1,k} + K_{k+1}(y_{k+1} - H\hat{x}_{k+1,k})$$

K_k = gain de Kalman

La covariance $P_{k+1,k+1}$ est donnée par

$$\begin{aligned} P_{k+1,k+1} &= (I - K_{k+1}H)P_{k+1,k}(I - K_{k+1}H)^* + K_{k+1}RK_{k+1}^* \\ &= P_{k+1,k} - K_{k+1}HP_{k+1,k} - P_{k+1,k}(K_{k+1}H)^* \\ &\quad + K_{k+1}(HP_{k+1,k}H^* + R)K_{k+1}^* \end{aligned}$$

Calcul du gain de Kalman

Le gain K_{k+1} est choisi de telle sorte à minimiser l'erreur quadratique de l'estimation donnée par

$$\begin{aligned} E[(\hat{x}_{k+1,k+1} - x_{k+1})^*(\hat{x}_{k+1,k+1} - x_{k+1})] &= \text{tr } E[(\hat{x}_{k+1,k+1} - x_{k+1})^*(\hat{x}_{k+1,k+1} - x_{k+1})] \\ &= \text{tr } E[(\hat{x}_{k+1,k+1} - x_{k+1})(\hat{x}_{k+1,k+1} - x_{k+1})^*] \\ &= \text{tr } P_{k+1,k+1} \end{aligned}$$

L'erreur d'estimation de $\hat{x}_{k+1,k+1}$ est minimale quand

$$\frac{\partial \text{tr } P_{k+1,k+1}}{\partial K_{k+1}} = 0$$

Cette notation désigne les dérivées partielles du scalaire $\text{tr } P_{k+1,k+1}$ par rapport aux termes de K_{k+1} , rangées dans une matrice de même forme que K_{k+1} .

Gain de Kalman

On utilise $\partial \text{tr}(AB)/\partial A = B$ et $\partial \text{tr}(ACA)/\partial A = 2AC$ si C est symétrique.

$$\text{tr } P_{k+1,k+1} = \text{tr } P_{k+1,k} - 2\text{tr} (K_{k+1}HP_{k+1,k}) + \text{tr} (K_{k+1}(HP_{k+1,k}H^* + R)K_{k+1}^*)$$

$$\begin{aligned}\frac{\partial \text{tr } P_{k+1,k+1}}{\partial K_{k+1}} &= -2(HP_{k+1,k})^* + 2K_{k+1}(HP_{k+1,k}H^* + R) \\ &= 0\end{aligned}$$

On obtient

$$K_{k+1} = P_{k+1,k}H^*(HP_{k+1,k}H^* + R)^{-1}$$

et

$$P_{k+1,k+1} = (I - K_{k+1}H)P_{k+1,k}$$

Filtre de Kalman

Données : modèles F et H , covariances Q et R , conditions initiales $\hat{x}_{0,0}$ et $P_{0,0}$.

Prédiction :

- $\hat{x}_{k+1,k} = F\hat{x}_{k,k}$
- $P_{k+1,k} = FP_{k,k}F^* + Q$

Correction :

- $K_{k+1} = P_{k+1,k}H^*(HP_{k+1,k}H^* + R)^{-1}$
- $\hat{x}_{k+1,k+1} = \hat{x}_{k+1,k} + K_{k+1}(y_{k+1} - H\hat{x}_{k+1,k})$
- $P_{k+1,k+1} = (I - K_{k+1}H)P_{k+1,k}$

Initialisation

L'estimation initiale $\hat{x}_{0,0}$ et sa variance $P_{0,0}$ sont nécessaires mais pas toujours disponibles.

- si l'état initial x_0 est connue avec certitude, $x_{0,0} = x_0$ et $P = 0$.
- si possible, utiliser la première mesure et sa covariance.
- à défaut, $x_{0,0} = 0$ et covariance P arbitrairement grande.

Plan

- 1 Introduction
- 2 Modèle de signaux
- 3 Filtre de Kalman
 - Cas Gaussien
 - Cas général
- 4 Exemples
 - Constante
 - Suivi de pendule
- 5 Cas non linéaire
 - Extended Kalman filter
 - Filtrage particulière
- 6 Conclusion

Un exemple très simple : estimation d'une constante

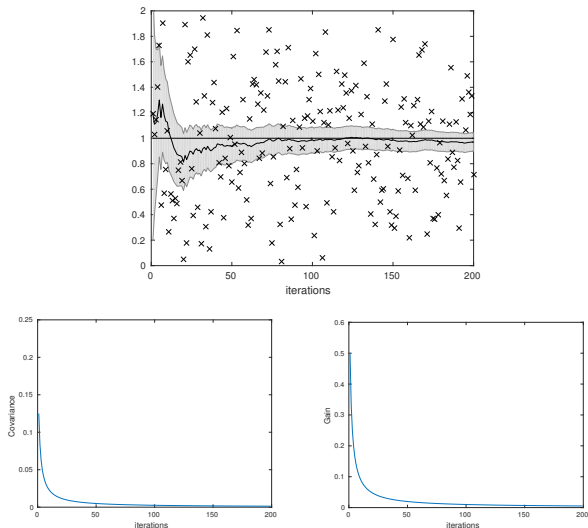
Modèle :

- $x_{k+1} = x_k + u_k$
- $y_k = x_k + v_k$

Filtre :

- $\hat{x}_{k+1,k} = \hat{x}_{k,k}$
- $p_{k+1,k} = p_{k,k} + \sigma_u^2$
- $k_{k+1} = \frac{p_{k+1,k}}{p_{k+1,k} + \sigma_v^2}$
- $\hat{x}_{k+1,k+1} = \hat{x}_{k+1,k} + k_{k+1}(y_{k+1} - \hat{x}_{k+1,k}) = \frac{p_{k+1,k}y_{k+1} + \sigma_v^2\hat{x}_{k+1,k}}{p_{k+1,k} + \sigma_v^2}$
- $p_{k+1,k+1} = \frac{p_{k+1,k}\sigma_v^2}{p_{k+1,k} + \sigma_v^2} < p_{k+1,k}$

Résultats

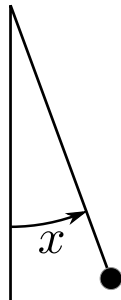


Au fil des itérations, le filtre de Kalman fait de plus en plus confiance à l'état du modèle, et utilise de moins en moins les mesures pour corriger cet état.

Suivi de pendule

Estimation de la position angulaire d'un pendule au cours du temps.

Modèle continu



Modèle continu

Le modèle d'évolution est ici donné sous une forme continue, i.e. un système d'équations différentielles ordinaires :

$$F(X, \dot{X}, \ddot{X}, \dots, X^{(p)}) = 0$$

Dans notre cas $p = 2$ et :

$$\omega_0^2 x + \ddot{x} = 0$$

Ce modèle doit être réécrit

- sous la forme d'un système à l'ordre 1
- à temps discret

Réécriture à l'ordre 1

L'état inclut x et ses dérivées jusqu'à l'ordre $p - 1$:

$$X = (x, \dot{x}).$$

La dérivée \dot{X} s'écrit en fonction de X :

$$\dot{X} = \begin{pmatrix} \dot{x} \\ \ddot{x} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} = AX$$

Système d'ordre 1 de dimension 2

$$\dot{X} = AX$$

Évolution du système

L'évolution du système entre deux temps t_1 et t_2 est donnée par

$$X(t_2) = \exp((t_2 - t_1)A)X(t_1)$$

Exponentielle matricielle :

$$\exp(M) = \sum_{n \in \mathbf{N}} \frac{M^n}{n!}$$

Modèle discret :

$$x_{k+1} = \exp(A\Delta t)x_k$$

Calcul de l'exponentielle matricielle

Si M est nilpotente : $M^p = 0$,

$$\exp(M) = \sum_{n=0}^{p-1} \frac{1}{n!} M^n$$

Si M est diagonalisable : $M = PDP^{-1}$

$$\exp(M) = P \exp(D) P^{-1}$$

et

$$\exp \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n} \end{pmatrix}$$

Suivi de pendule - modélisation

Modèle continu:

$$\ddot{x} + \omega_0^2 x = 0$$

Formulation vectorielle d'ordre 1 avec $X = (x, \dot{x})$:

$$\dot{X} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} X = AX$$

Diagonalisation de $A = PDP^{-1}$:

$$P = \begin{pmatrix} 1 & 1 \\ i\omega_0 & -i\omega_0 \end{pmatrix}, D = \begin{pmatrix} i\omega_0 & 0 \\ 0 & -i\omega_0 \end{pmatrix}$$

Exponentielle de A :

$$F = \exp(A\Delta t) = P \exp(D\Delta t) P^{-1} = \begin{pmatrix} \cos \omega_0 \Delta t & \frac{1}{\omega_0} \sin \omega_0 \Delta t \\ -\omega_0 \sin \omega_0 \Delta t & \cos \omega_0 \Delta t \end{pmatrix}$$

Suivi de pendule - modélisation

Modèle d'évolution discret :

$$F = \exp(A\Delta t) = P \exp(D\Delta t)P^{-1} = \begin{pmatrix} \cos \omega_0 \Delta t & \frac{1}{\omega_0} \sin \omega_0 \Delta t \\ -\omega_0 \sin \omega_0 \Delta t & \cos \omega_0 \Delta t \end{pmatrix}$$

On ne mesure que la position du pendule :

$$H = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

On suppose que l'incertude du modèle ne porte que sur la vitesse angulaire du pendule.

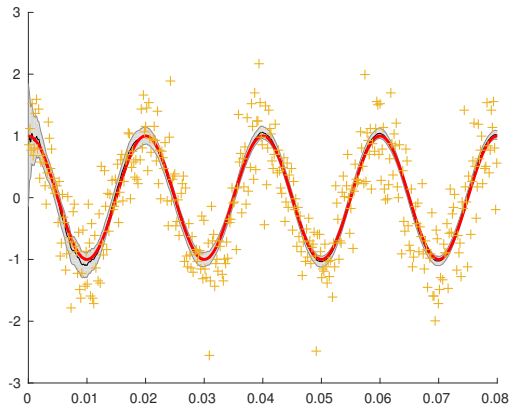
$$Q = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_Q^2 \end{pmatrix}$$

Les mesures sont bruitées :

$$R = \sigma_R^2$$

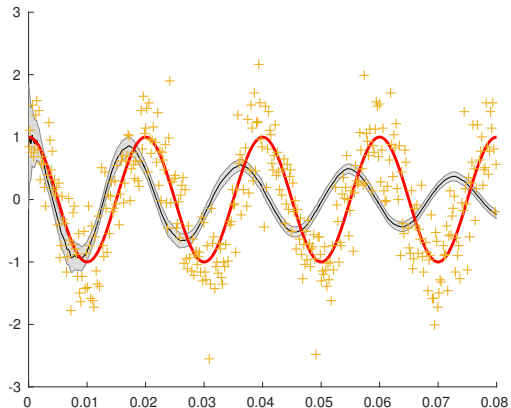
Résultats

Cas où la pulsation du signal et du filtre sont en accord, $\omega = \omega_0$



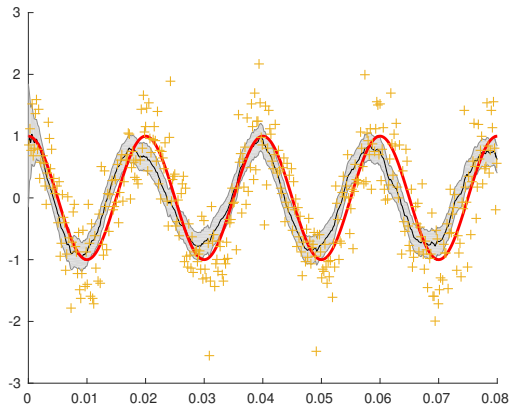
Résultats

Cas où la pulsation du signal et du filtre ne sont pas accordées, $\omega = 1.1\omega_0$



Résultats

Cas où la pulsation du signal et du filtre ne sont pas accordées, $\omega = 1.1\omega_0$, mais avec un fort bruit d'évolution.



On voudrait inclure la pulsation dans l'état : non-linéarités.

Extended Kalman filter

Extension au cas où le modèle n'est pas linéaire :

$$\begin{cases} x_{k+1} = F(x_k) + u_k \\ y_k = H(x_k) + v_k \end{cases}$$

Adaptation simple :

- on utilise directement F et H dans le calcul de l'état a priori et l'innovation
- on utilise les gradients de F et G pour le calcul des covariances et du gain de Kalman.

Application au pendule

État :

$$\tilde{X} = (x, \dot{x}, \omega) = (X, \omega)$$

Évolution :

$$\tilde{X}_{n+1} = \tilde{F}(\tilde{X}_n) = (F(\omega_n)X_n, \omega_n)$$

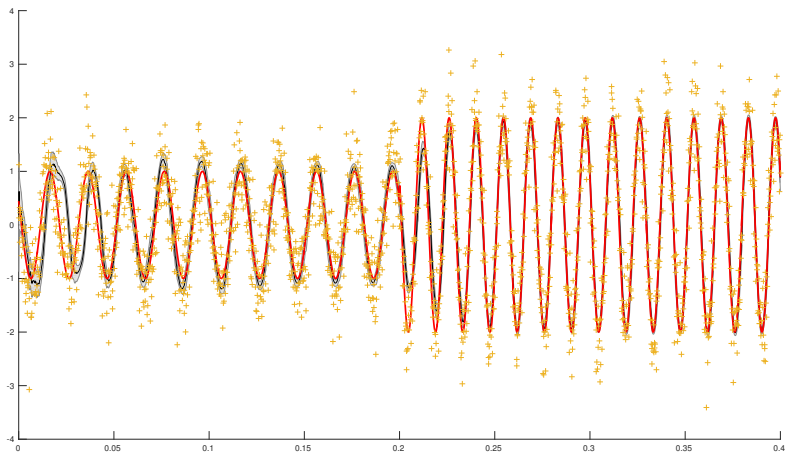
$$F(\omega) = \begin{pmatrix} \cos \omega \Delta t & \frac{1}{\omega} \sin \omega_0 \Delta t \\ -\omega \sin \omega \Delta t & \cos \omega \Delta t \end{pmatrix}$$

$$\text{grad } \tilde{F}(\tilde{X}) = \begin{pmatrix} \cos \omega \Delta t & \frac{1}{\omega} \sin \omega \Delta t & \Delta t \left(\frac{\dot{x}}{\omega} \cos \omega \Delta t - x \sin \omega \Delta t \right) \\ -\omega \sin \omega \Delta t & \cos \omega \Delta t & -\Delta t (x \omega \cos \omega \Delta t + \dot{x} \sin \omega \Delta t) \\ 0 & 0 & 1 \end{pmatrix}$$

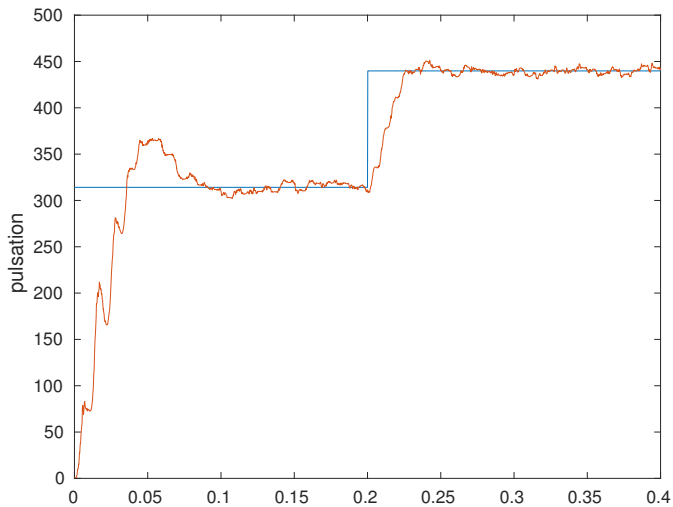
Observation linéaire :

$$H = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$$

Application au pendule



Application au pendule - pulsation



Filtrage particulière

Le filtrage particulière est basé sur une approche Monte-Carlo de l'estimation bayésienne, i.e. sur

$$E(f(x)) \approx \frac{1}{L} \sum_{l=1}^L f(x_l) \quad (1)$$

où les x_l suivent la loi de x , ou plus généralement, avec un échantillonnage d'importance,

$$E(f(x)) \approx \frac{1}{L} \sum_{n=1}^L w(x_l) f(x_l) \quad (2)$$

tels que $x_l \sim q(x)$ et $w(x)q(x) = p(x)$.

Espérance et covariance peuvent se mettre sous cette forme.

Bootstrap filter

On dispose d'échantillons de $x_n|y_{1:n}$.

Étape de prédiction :

$$\hat{x}_l \sim p(x_{n+1}|y_{1:n}) \quad (3)$$

Pas toujours possible mais facile si $x_{n+1} = F(x_n) + u_n$.

Étape de correction :

On calcule les poids

$$w_l = \frac{1}{Z} p(x_{n+1}|y_{n+1}) \quad (4)$$

tels que $\sum_l w_l = 1$.

Rééchantillonnage : on remplace les x_l par des échantillons selon la loi empirique des \hat{x}_l pondérés par les poids w_l .

Conclusion

Le filtre de Kalman est adapté aux situations où

- le système est linéaire
- les observations sont linéaires
- les erreurs sont des bruits blancs gaussiens additifs

Dans ce cas, le filtrage de Kalman est optimal.

Son implémentation est aisée et efficace.

Slides et code sur <http://gilleschardon.fr/fc/kalman/>.

Bibliographie

- Kalman, Rudolph Emil, A New Approach to Linear Filtering and Prediction Problems, Transactions of the ASME–Journal of Basic Engineering, Vol. 82, Series D, pp.= 35-45, 1960
- A. H. Jazwinski, Stochastic process and filtering theory, Academic Press, New York, USA, 1970
- B. D. O. Anderson and J. B. Moore, Optimal Filtering, Prentice-Hall, Englewood Cliffs, NJ, USA, 1979
- Special Issue on applications of Kalman filtering, vol. AC-28 of IEEE Transactions on Automatic and Control, mars 1983