

MA201 - Séance 6

Filtres récurrents bayésiens

H. Piet-Lahanier - L. Meyer



Content

1 Le filtrage récursif bayésien

2 Les équations du filtre

3 Linéarisation

4 Approximation Monte Carlo

Modèle : équations et hypohèses

Les variables et les données

- Variables d'état x_k : caractérisent l'évolution du système
- Mesures z_k

Modèle de représentation

Équation d'état : lien entre x_{k+1} et x_k

Équation de mesure : lien entre x_{k+1} et z_{k+1}

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

- f_k : fonction de transition x_k état courant w_k bruit d'état
- h_k : fonction de mesure. v_k bruit de mesure

Modèle : équations et hypoyhèses

Hypothèses Markoviennes

- L'état courant ne dépend que de l'état précédent et du bruit
- Les mesures z_k ne dépendent que de l'état courant et du bruit
- les mesures z_k sont indépendantes conditionnellement aux états

Hypothèses

- Les bruits w_k , bruit de modèle (ou d'état) et v_k , bruit de mesures sont supposés blancs
- les bruits w_k et v_k , $k \geq 0$ sont mutuellement indépendants

Les principes du filtrage récursif

Objectifs

- Estimer les valeurs de l'état $x_{1,\dots,k}$ connaissant les mesures courantes et passées
- Utilisation de la probabilité jointe : $p(x_{0,\dots,k}, z_{1,\dots,k})$
- \Rightarrow Déterminer $p(x_{0,\dots,k}, z_{1,\dots,k})$

Filtrage récursif

- Exploitation des hypothèses Markoviennes
- Densité $p(x_{0,\dots,k}, z_{1,\dots,k})$: sous forme récursive :

$$p(x_{0,\dots,k}, z_{1,\dots,k}) \propto p(z_k, x_k) p(x_k, x_{k-1}) p(x_{1,\dots,k-1}, z_{1,\dots,k-1})$$

Les principes du filtrage récursif

Éléments de démonstration

- Par utilisation du théorème de Bayes
- L'état courant ne dépend que de l'état précédent :

$$p(x_k | x_{0,\dots,k-1}, z_{1,\dots,k}) = p(x_k | x_{k-1})$$
- Les mesures sont indépendantes conditionnellement aux états :

$$p(z_k | z_{1,\dots,k-1}, x_{0,\dots,k}) = p(z_k | x_k)$$

Récursion

$$p(x_{0,\dots,k}, z_{1,\dots,k}) \propto p(z_k, x_k) p(x_k, x_{k-1}) p(x_{0,\dots,k-1}, z_{1,\dots,k-1})$$

Principes

Objectif

Estimer l'état x_k à l'aide de x_{k-1} et z_k

⇒ Établir les expressions :

- Probabilité de transition
- Vraisemblance

⇒ ou une approximation de ces valeurs

Récursion

$$p(x_{0,\dots,k}, z_{1,\dots,k}) \propto p(z_k, x_k) p(x_k, x_{k-1}) p(x_{0,\dots,k-1}, z_{1,\dots,k-1})$$

- Probabilité de transition $p(x_k, x_{k-1})$:
Dépend de la fonction f et des caractéristiques du bruit w_k
- Vraisemblance de la mesure $p(z_k, x_k)$:
Dépend de la fonction h et du bruit v_k

Content

1 Le filtrage récursif bayésien

2 Les équations du filtre

3 Linéarisation

4 Approximation Monte Carlo

Rappel : Filtre de Kalman

Expression exacte dans le cas linéaire gaussien

$$\begin{cases} p(x_k, x_{k-1}) : x_{k+1} &= F_k x_k + w_k \\ p(z_k, x_k) : z_k &= H_k x_k + v_k \end{cases},$$

- *Probabilité de transition* : Propagation d'un vecteur gaussien par une transformation linéaire

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

- *Vraisemblance des mesures*

$v_{k+1} = z_{k+1} - H_k \hat{x}_{k+1|k} \Rightarrow z_{k+1} - H_k \hat{x}_{k+1|k}$ vecteur aléatoire gaussien de covariance V

Equations

Estimé

$$\text{Prédiction} \begin{cases} \hat{x}_{k+1|k} &= F_k \hat{x}_k \\ P_{k+1|k} &= F_k P_k F_k^T + W_k \end{cases} \quad (2)$$

$$\text{Correction} \begin{cases} K_{k+1} &= P_{k+1|k} H_{k+1}^T (H_{k+1} P_{k+1|k} H_{k+1}^T + V_k)^{-1} \\ \hat{x}_{k+1} &= \hat{x}_{k+1|k} + K_{k+1} (z_{k+1} - H_k \hat{x}_{k+1|k}) \\ P_{k+1} &= (I - K_{k+1} H_{k+1}) P_{k+1|k} \end{cases} \quad (3)$$

Cas non linéaire

Equations générales

$$\begin{cases} x_{k+1} &= f_k(x_k, w_k) \\ z_k &= h_k(x_k, v_k) \end{cases}, \quad (4)$$

Non linéarité

- Propagation d'un vecteur Gaussien \neq Gaussien
- Problème calcul de la vraisemblance

\Rightarrow Deux approches possibles

- Linéarisation simple ou multiple
- Approximation Monte Carlo (valeurs moyennes sur un nuage de points)

Content

1 Le filtrage récursif bayésien

2 Les équations du filtre

3 Linéarisation

4 Approximation Monte Carlo

Filtre de Kalman Etendu (EKF), Filtre sans parfum (UKF)

Equations

$$\begin{cases} x_{k+1} &= f(x_k) + w_k \\ z_k &= h(x_k) + v_k \end{cases} \quad (5)$$

- f et h fonctions différentiables connues,
- Bruits additifs
- $\forall k \geq 0$, w_k et v_k suivent des lois normales $\mathcal{N}(0, W)$ et $\mathcal{N}(0, V)$ respectivement.
- w_k et v_k mutuellement indépendants.

Filtre de Kalman Etendu (EKF)

Linéarisation locale

- Approximation de la probabilité de transition
 - Prédiction de x_k
 - Vecteur gaussien de moyenne $F_k \hat{x}_k$ et de covariance $F_k P_k F_k^T + W_k$
 - $F_k = \frac{\partial f(x)}{\partial x} \big|_{x=\hat{x}_k}$ (matrice jacobienne de f calculée en \hat{x}_k)

Linéarisation locale

- Approximation de la vraisemblance
 - z_k mesure d'un vecteur gaussien : $H_k \hat{x}_{k+1|k} + v_k$
 - $H_k = \frac{\partial h(x)}{\partial x} \big|_{x=\hat{x}_{k+1|k}}$ (matrice jacobienne de h prise en $\hat{x}_{k+1|k}$)
 - Estimateur du maximum de vraisemblance :
 - $\hat{x}_{k+1} = \hat{x}_{k+1|k} + K_{k+1}(z_{k+1} - H_k \hat{x}_{k+1|k})$
 - avec $K_{k+1} = P_{k+1|k} H_{k+1}^T (H_{k+1} P_{k+1|k} H_{k+1}^T + V_k)^{-1}$

Filtre de Kalman Etendu (EKF)

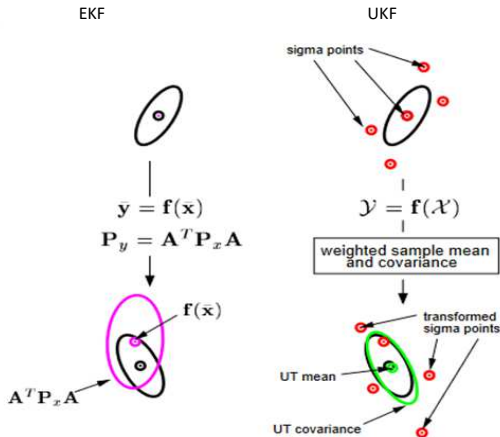
Limites de la linéarisation locale

- Développement de Taylor autour de la valeur moyenne de la gaussienne
- Pour approximation de la probabilité de transition ou la vraisemblance
- Peu efficace pour des non linéarités sévères

Alternative : Filtre sans parfum

- Approximation de la fonction par un ensemble de points
- Points représentatifs de la distribution initiale : *Sigma points*
- Prédiction : Points transformés par la fonction non linéaire
- Reconstruction de l'approximation sur les transformées de ces points

Filtres de Kalman Etendu et sans parfum (UKF)



Filtre de Kalman sans parfum (UKF)

Filtre sans parfum : Choix des sigma points

Sigma points : Points représentatifs de la distribution initiale ζ_i , $i = 0, \dots, 2L$, On peut prendre $L = n$, n étant la dimension de l'espace (état)

Premier point : ζ_0 correspond à la valeur moyenne de la distribution à transformer (\hat{x}_k pour la transition)

Détermination des autres Sigma Points

- n étant la dimension de l'espace (état)
- n points : $\zeta_i = \zeta_0 + \sqrt{(n + \lambda)} \Sigma_i$, $i = 1, \dots, n$
- n points : $\zeta_i = \zeta_0 - \sqrt{(n + \lambda)} \Sigma_i$, $i = n + 1, \dots, 2n$

Σ_i est la i ème colonne de la matrice Σ : *Matrice Racine* de la covariance,
 \Rightarrow Obtenue par décomposition de Cholesky de la matrice de covariance $\hat{P}_k = \Sigma \times \Sigma^t$

Filtre de Kalman sans parfum (UKF)

Sigma points

Le nombre total de sigma points est de $2L + 1$

Choix de λ : il est important, la valeur doit être assez faible

Expression possible de $\lambda = \alpha^2(n + \kappa) - n$ avec α petit (représente la dispersion autour de la moyenne), κ paramètre souvent choisi nul $\kappa = 0$, β utilisé pour incorporer une connaissance a priori sur la distribution de x : pour des lois gaussiennes $\beta = 2$ est optimal.
E. Wan et R. Van der Merwe : The unscented Kalman filter for nonlinear estimation. Proceedings of the IEEE 2000 adaptive systems for signal proc., commu., and contr. symposium , 2000.

- La pondération pour le calcul de la moyenne du sigma point associé à la moyenne est $\omega_{m0} = \frac{\lambda}{n+\lambda}$
- La pondération pour le calcul de la covariance du sigma point associé à la moyenne est $\omega_{c0} = \frac{\lambda}{n+\lambda} + 1 - \alpha^2 + \beta$ si on utilise l'expression de λ de Wan et
- Les pondérations des autres points sont pour la moyenne et la covariance $\omega_{mi} = \omega_{ci} = \frac{1}{2(n+\lambda)}$
- La somme des pondération est égale à 1 si on n'utilise pas α et β pour ω_{c0}

Filtre de Kalman sans parfum (UKF)

Algorithme

Prédiction (probabilité de transition)

- Calcul des sigma points et de leurs pondérations associés à \hat{x}_k
- Propagation par f pour obtenir $\hat{x}_{k+1/k}$ et $P_{k+1/k} = P_p + W$
- $\hat{x}_{k+1/k} = \sum_{i=0}^{2n} \omega_{mi} f(\zeta_i)$
- Covariance $P_{k+1/k} = \sum_{i=0}^{2n} \omega_{ci} (f(\zeta_i) - \hat{x}_{k+1/k})(f(\zeta_i) - \hat{x}_{k+1/k})^t + W$

Filtre de Kalman sans parfum (UKF)

Algorithme

Correction (Vraisemblance)

- Calcul des sigma points $\tilde{\zeta}_i$ pour $\hat{x}_{k+1/k}$
- Calcul de la mesure prédite : $\hat{z} = \sum_{i=0}^{2n} \omega_{mi} h(\tilde{\zeta}_i)$
- Covariance empirique : $\tilde{P}_c = \sum_{i=0}^{2n} \omega_{ci} (h(\tilde{\zeta}_i) - \hat{z})(h(\tilde{\zeta}_i) - \hat{z})^t + V$
- Calcul de l'erreur de prédiction : $T = \sum_{i=0}^{2n} \omega_{mi} (f(\zeta_i) - \hat{x}_{k+1/k})(h(\tilde{\zeta}_i) - \hat{z})^T$
- Calcul du gain de Kalman : $K = T \times \tilde{P}_c^{-1}$

Mise à jour

- $\hat{x}_{k+1} = \hat{x}_{k+1/k} + K(z - \hat{z})$
- $P_{k+1} = P_{k+1/k} - K\tilde{P}K^t$

Filtre de Kalman sans parfum (UKF)

Avantages

- Linéarisation locale remplacée par approximation sur un ensemble de points
- Choix des points adapté à la répartition Gaussienne
- Calcul exact des propagations
- Nombre de calculs supplémentaires limités
- Equations similaires de celles du filtre de Kalman
- Peu de paramètres à choisir (λ)

Limitations

- Hypothèse gaussienne sur les bruits
- Pas de multimodalités (décomposition sur la matrice de covariance)

Content

1 Le filtrage récursif bayésien

2 Les équations du filtre

3 Linéarisation

4 Approximation Monte Carlo

Cas non linéaire non gaussien multimodalité

Constat

- EKF et UKF : Hypothèse gaussienne et fonctions différentiables
- Pas adapté aux cas non gaussiens avec des transformations non régulières

Approximation

- Approximation de la densité par un échantillon pondéré
- Répartition des points aléatoires (tirages Monte-Carlo)
- Détermination des points à conserver

Le filtre Particulaire (apparu en 1993)

Principe

- Mise en oeuvre des étapes de prédiction correction
- En propageant les particules et en mettant à jour leurs pondérations
- Problème de dégénérescence

Le filtre Particulaire

Objectif

- Approximer la densité a posteriori à partir d'une somme pondérée de particules
- $p(x_{0,\dots,k}, z_{1,\dots,k} = \sum_{i=1}^N \omega_i x^i(x_{0,\dots,k})$
- avec ω_i pondérations et $x^i(x_{0,\dots,k})$ particules
- \Rightarrow Déterminer l'estimé \hat{x}_k
- Mais on ne sait pas représenter la densité a posteriori

Etapas de l'algorithme

- Prédiction : Densité *a priori* $\hat{x}_{k|k_1}$ par la probabilité de transition
- Correction : Densité *a posteriori* de \hat{x}_k connaissant y_k par la vraisemblance et l'*a priori*
- Rééchantillonnage

\Rightarrow Etapes identiques à celles des autres filtres (sauf rééchantillonnage)

Mais à exécuter sur chaque particule

Le filtre Particulaire

Définition des particules

- Choix d'un nombre N grand
- Répartition des particules autour de la valeur estimée mais pas nécessairement
- Définition des $x_k^i = \hat{x}_k + \eta_k^i$ avec η_k^i bruit aléatoire générateur
- Association de pondération $\frac{1}{N}$ pour chaque particule

Prédiction

- Propagation des particules par la probabilité de transition
- $p(x_k|x_{k-1}) = \int p(x_k|x_{k-1}) \sum_{i=1}^N \omega_i x_{k-1}^i dx_{k-1}$
- $p(x_k|x_{k-1}) = \sum_{i=1}^N \omega_i \int p(x_k|x_{k-1}) x_{k-1}^i$
- $p(x_k|x_{k-1}) = \sum_{i=1}^N \omega_i p(x_k|x_{k-1}^i)$
- Prédiction des particules : $x_{k|k-1}^i = f_k(x_{k-1}^i) + \eta_k^i$

Les poids sont inchangés mais leur valeur devient faible pour N grand

Le filtre Particulaire

Correction

- Calcul de la vraisemblance
- $p(z_k | x_{k|k-1}^i) = p(z_k - h(x_{k|k-1}^i))$
- Correction des poids : $\omega_{k|k-1}^i = \omega_{k-1}^i \times p(z_k | x_{k|k-1}^i)$
- Normalisation $\omega_{k|k-1}^i = \frac{\omega_{k-1}^i \times p(z_k | x_{k|k-1}^i)}{\sum_{i=1}^N \omega_{k|k-1}^i}$

Estimateur

- $\hat{x}_k = \sum_{i=1}^N \omega_{k|k-1}^i x_{k|k-1}^i$
- Covariance associée
- $P_k = \sum_{i=1}^N \omega_{k|k-1}^i (x_{k|k-1}^i - \hat{x}_k)(x_{k|k-1}^i - \hat{x}_k)^t$

Le filtre Particulaire

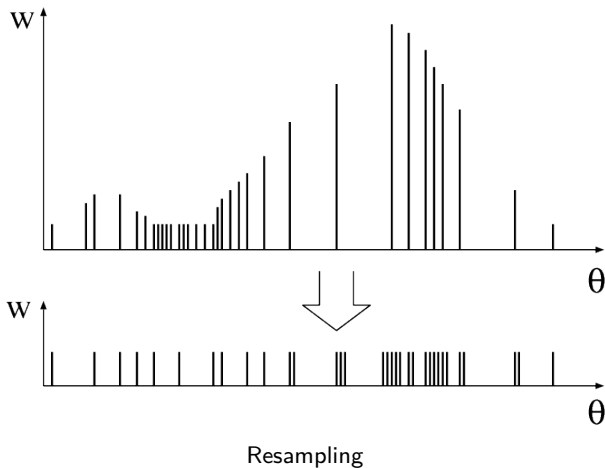
Problème de poids

- Dégénérescence : Les valeurs des poids décroissent et il n'en reste plus qu'un seul non nul
- Si on conserve les mêmes particules : risque de mauvaise représentation de la pdf
- Représentation clairsemée des zones à haute probabilité (un seul point epoids fort)

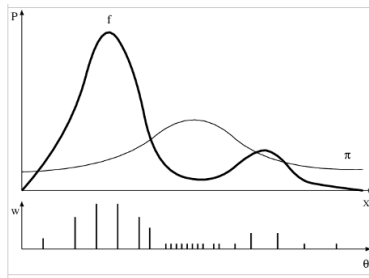
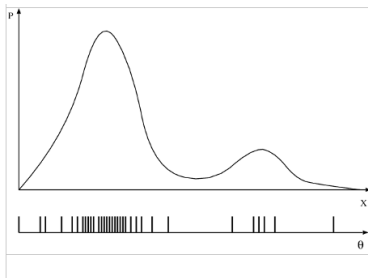
Resampling

- Élimination des particules à poids faibles
- Conservation de particules à poids forts et retraitage de particules supplémentaires
- Nombreuses méthodes possible
- Rééchantillonnage multinomial : On fait un tirage aléatoire multinomial sur les particules en utilisant les poids comme paramètres de la loi
Les particules à poids forts sont tirées plus fréquemment que celles à poids faible
- Appauvrissement : risque de ne conserver qu'un seul type de particules
- Si le nombre de particules différentes est trop faible \Rightarrow retraitage des particules

Le filtre Particulaire : Resampling



Le filtre Particulaire : tirage par importance sampling



À gauche : Représentation d'une distribution de probabilités par des particules tirées aléatoirement selon la loi à représenter.

À droite : Représentation d'une distribution de probabilités par des particules tirées selon une loi d'importance π avec un poids associé $w = \frac{f}{\pi}$

Le filtre Particulaire

Avantages

- Beaucoup de flexibilité
- Peut s'adapter à tous les types de modèles
- Peut traiter les cas multi modaux

Limitations

- Nécessité d'avoir beaucoup de particules : nombre de calculs requis grand
- Problème de dégénérescence des poids : pas de méthodes génériques
- Choix du mode de tirages : facteur impactant sur les résultats

Ce qu'il faut retenir

- Le filtrage récursif bayésien permet d'estimer des vecteurs d'états, pour lesquels on dispose d'une équation de mesures à différents instants.
- Par l'hypothèse de représentation markovienne et d'indépendance des données conditionnellement aux variables d'états, on peut estimer les états de façon récursive
- La récursion est constituée d'une étape de prédiction (densité de transition pour l'évolution de l'état) et de correction (calcul de la vraisemblance de la mesure étant donné la valeur prédite de l'état)
- Suivant la nature du modèle et les hypothèses sur les bruits, on peut disposer d'une expression analytique de l'estimateur optimal au sens du risque quadratique moyen : Filtre de Kalman pour le cas linéaire, gaussien
- Pour les autres types de modèles, on peut effectuer une linéarisation locale (EKF), une linéarisation sur un ensemble de points descriptifs (UKF) ou une caractérisation par génération Monte Carlo de particules (Filtre particulaire)