Formation et Analyse d'Images

James L. Crowley

ENSIMAG 3 Premier Bimestre 2006/2007

Séance 9 24 novmbre 2006

Suivi et estimation de movements
Le filtre de Kalman

Plan de la Séance :

Le Filtre de Kalman

Le filtre de Kalman est un estimateur récursif optimal.

Son origine se trouve dans le problème de l'estimation de la trajectoire des satellites et des aéronefs. Dans un tel problème, on cherche à estimer la position (x, y, z) et la vitesse (v_x, v_y, v_z) à partir de l'observation de deux angles : Site et azimut. (θ, ϕ)

Dans la terminologie de Kalman : la position et vitesse est un vecteur d'état : $\vec{X}$ et l'observation est $\vec{Y}$

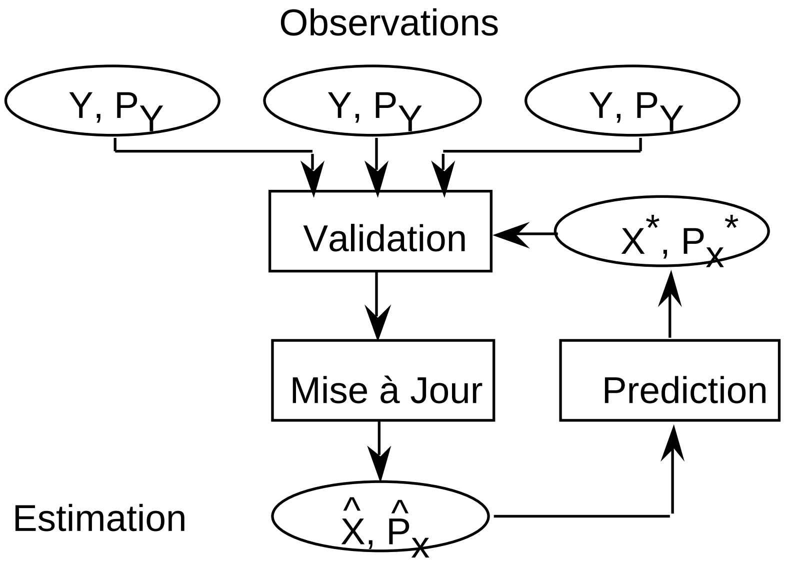
$$\vec{X} = \begin{pmatrix} x \\ y \\ z \\ v_x \\ v_y \\ v_z \end{pmatrix} \qquad \vec{Y} = \begin{pmatrix} \theta \\ \phi \end{pmatrix}$$

Pour chaque instant, t, le filtre donne une estimation $\hat{X}_t$ ainsi que sa précision $\mathbf{P}_t$, à partir de l'observation et l'estimation précédente, $\hat{X}_{t-1}$, $\mathbf{P}_{t-1}$, et l'observation $\vec{Y}$, et sa précision, $\mathbf{P}_y$. Les précisions sont des <u>covariances</u>.

$$\hat{X}_t,\ \mathbf{P}_t \ := F\ \{\ \hat{X}_{t-1},\ \mathbf{P}_{t-1},\ \vec{Y},\ \mathbf{P}_y\ \}$$

Ce permet d'estimer un vecteur de N variable aléatoire à partir de M< N observations. Les précisions sont des 2ieme moments de l'erreur.

Le filtre récursif est un processus cyclique avec 5 phases :
1) Prédiction de l'évolution temporelle de l'état.
2) Observation
3) Prédiction de l'observation
4) Validation de l'observation.
5) Mise à jour de l'estimation.



Techniques :

Prédictions:	Modèle Linéaire
validation :	Distance de Mahalanobis
	(Distance normalisée par Covariance)
Mise à Jour	Estimation linéaire

Notation:

Vecteur d'état a temp t :	$\vec{X}_t$
Estimation du vecteur d'état :	$\hat{X}_t$, $\hat{\mathbf{P}}_t$
Observation :	$\vec{Y}_\tau$, $\mathbf{P}_y$
Prédiction :	X_t^*, $\mathbf{P}_t^*$
Modèle du capteur :	$\mathbf{H}_X^Y$ avec précision $\mathbf{P}_y$
Modèle du processus :	$\phi(\Delta t)$ avec précision $\mathbf{Q}_t$
Affectation :	:=
Définition :	≡

Le modèle de Kalman suppose que les estimations sont corrompues par un bruit $\vec{B}$.

$$\hat{X} = \vec{X} + \vec{B}$$

Le bruit est une variable aléatoire de moyenne nul.
La précision est le deuxième moment de ce bruit.

$$\mathbf{P} \equiv E\{ \vec{B}\ \vec{B}^T\} = E\{(\vec{X} - \hat{X})(\vec{X} - \hat{X})^T\}$$

On ne peut connaître ni $\vec{X}$ ni $\vec{B}$. Kalman nous fourni les estimations $\hat{\mathbf{P}}_t$

Prédiction de l'évolution temporelle de l'état

La phase de prédiction projets $\hat{X}_{t-1}$ et $\hat{\mathbf{P}}_{t-1}$ à temps t afin de produire X_t^* et $\mathbf{P}_t^*$

On inclus les dérivées temporelles des paramètres dans le vecteur d'etat. $\hat{X}_{t-1}$

Soit une propriété, $\hat{x}$, du vecteur $\hat{X}$, avec variance $\hat{\sigma}_x^2$.
La prédiction de premier ordre x*(t) exige la dérivée temporel $\hat{x}'(t-\Delta t)$.

$$\hat{x}'_t \triangleq \frac{\partial \hat{x}_t}{\partial t}$$

Par série de Taylor, on peut écrire :

$$x^*_t := \hat{x}_{t-\Delta t} + \hat{x}'_{t-\Delta t}\, \Delta t + r$$

avec l'hypothese que E{r} = 0.
Soit un vecteur de deux propriétés $\hat{x}_1$ et $\hat{x}_2$ avec une prédiction de premier ordre.

$$\hat{X}_t \equiv \begin{bmatrix} \hat{x}_1 \\ \hat{x}_1' \\ \hat{x}_2 \\ \hat{x}_2' \end{bmatrix}$$

Le prédiction est : $X^*_t := \phi(\Delta t)\hat{X}_{t-\Delta t} + R$

avec le matrice de transition, $\phi(\Delta t) \equiv \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 1 \end{bmatrix}$

La précision de X^*_t est la covariances les propriétés $\hat{x}$

$$\mathbf{P}_t^* \triangleq \begin{bmatrix} \hat{\sigma}_{x_1}^2 & \hat{\sigma}_{x_1'x_1} & \hat{\sigma}_{x_2x_1} & \hat{\sigma}_{x_2'x_1} \\ \hat{\sigma}_{x_1x_1'} & \hat{\sigma}_{x_1'}^2 & \hat{\sigma}_{x_2x_1'} & \hat{\sigma}_{x_2'x_1'} \\ \hat{\sigma}_{x_1x_2} & \hat{\sigma}_{x_1'x_2} & \hat{\sigma}_{x_2}^2 & \hat{\sigma}_{x_2'x_2} \\ \hat{\sigma}_{x_1x_2'} & \hat{\sigma}_{x_1'x_2'} & \hat{\sigma}_{x_2x_2'} & \hat{\sigma}_{x_2'}^2 \end{bmatrix}$$

où

$$\sigma_{xx'} \triangleq E\{ (x - \hat{x})(x' - \hat{x}')\}$$

et

$$\sigma_{x'}^2 \triangleq E\{ (x' - \hat{x}')^2\}$$

La covariance $\mathbf{P}_t^*$ est prédite par $\mathbf{P}_t^* := \phi(\Delta t)\hat{\mathbf{P}}_{t-\Delta t}\,\phi(\Delta t)^T + E\{ R_t R_t^T\}$

Le processus peut avoir les dérivés d'ordre supérieur à le modèle.

Ces dérivés donnent lieu au "résidu" R.
R est supposée d'être moyenne nul.
Son deuxième moment indique la perte de précision dû au modèle linéaire.

$$X_t^* := \hat{X}_{t-\Delta t} + \frac{\partial X_{t-\Delta t}}{\partial t} \Delta T + R_t.$$

avec

$$\mathbf{Q} \equiv E\{R_t R_t^T\}$$

Prédiction :

$$\text{I :} \quad X_t^* := \phi(\Delta t)\hat{X}_{t-\Delta t} + R_t.$$

$$\text{II :} \quad \mathbf{P}_t^* := \phi(\Delta t)\hat{\mathbf{P}}_{t-\Delta t}\, \phi(\Delta t)^T + \mathbf{Q}_x$$

Modèle du capteur

Un capteur projet l'estimation de l'état du monde vers un vecteur d'observation en o dimentions

$$\text{III : } \quad Y_t^* \; := \; \mathbf{H}_X^Y \, X_t^*$$
$$\text{IV : } \quad \mathbf{P}_y^* \; = \; \mathbf{H}_X^Y \, \mathbf{P}_t^* \, (\mathbf{H}_X^Y)^T$$

La matrice de $\mathbf{H}_X^Y$ est un modèle linéaire du capteur.
Les observations sont corrompues par les bruits aléatoires
ainsi que une erreur d'approximation.

Ceci est représenté par l'erreur du capteur : W_t, calibré au Laboratoire.

On suppose que $E\{W_t\} = 0$
tandis que $\mathbf{P}_y \equiv E\{W_t \, W_t^{\,T}\}$

Par exemple, considère l'observation Y (o=2) d'un état X_t (d = 4).

$$Y_t^* \; := \; \mathbf{H}_X^Y \, X_t^* \quad => \quad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_1' \\ x_2 \\ x_2' \end{bmatrix}$$

$$\mathbf{P}_y^* \; = \; \mathbf{H}_X^Y \, \mathbf{P}_t^* \, (\mathbf{H}_X^Y)^T \; = \; \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{\sigma}_{x_1}^2 & \hat{\sigma}_{x_1'x_1} & \hat{\sigma}_{x_2x_1} & \hat{\sigma}_{x_2'x_1} \\ \hat{\sigma}_{x_1x_1'} & \hat{\sigma}_{x_1'}^2 & \hat{\sigma}_{x_2x_1'} & \hat{\sigma}_{x_2'x_1'} \\ \hat{\sigma}_{x_1x_2} & \hat{\sigma}_{x_1'x_2} & \hat{\sigma}_{x_2}^2 & \hat{\sigma}_{x_2'x_2} \\ \hat{\sigma}_{x_1x_2'} & \hat{\sigma}_{x_1'x_2'} & \hat{\sigma}_{x_2x_2'} & \hat{\sigma}_{x_2'}^2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Aproximations des modèle non-linéaires

Souvent le vrai capteur n'est pas linéaire.

$$\vec{Y} = F(\vec{X})$$

Dans ce cas, on peut faire une approximation linéaire avec la dérivée première, (la Jacobienne), calculé au tour de l'estimation actuelle.

$$\mathbf{H}_X^Y \quad \approx \quad \mathbf{J}_X^Y \quad = \quad \frac{\partial F(\hat{X}_t)}{\partial \hat{X}_t} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} & \frac{\partial y_1}{\partial x_4} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} & \frac{\partial y_2}{\partial x_4} \end{pmatrix}$$

et donc :
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} & \frac{\partial y_1}{\partial x_4} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} & \frac{\partial y_2}{\partial x_4} \end{pmatrix} \begin{pmatrix} x_1 \\ x_1' \\ x_2 \\ x_2' \end{pmatrix}$$

Validation

Validation repose sur une mesure de similarité. Une telle mesure est la "Distance de Mahalanobis" ou "Différence normalisée par covariance"

Soit la prédiction Y_t^* avec sa précision $\mathbf{P}_t^*$

et l'observation $\vec{Y}_\tau$ avec sa précision $\mathbf{P}_y$

L'Innovation : V = $(\vec{Y}_\tau - Y_t^*)$ est l'information obtenu par l'observation.

Il doit être validé par une distance de Mahalanobis

$$V : \quad d^2 = \frac{1}{2} \{ (\vec{Y}_\tau - Y_t^*)^T (\mathbf{P}_t^* + \mathbf{P}_y)^{-1} (\vec{Y}_\tau - Y_t^*) \}$$

Pour une scalaire :

$$d^2 = \frac{(x^* - y)^2}{2\,(\sigma_x^{*2} + \sigma_y^2)}$$

On considère que X_t^*, $\mathbf{P}_t^*$ définie une probabilité d'une observation Y.

Si il y a K prédictions : X_{kt}^*, $\mathbf{P}_{kt}^*$, l'observation est associée a :

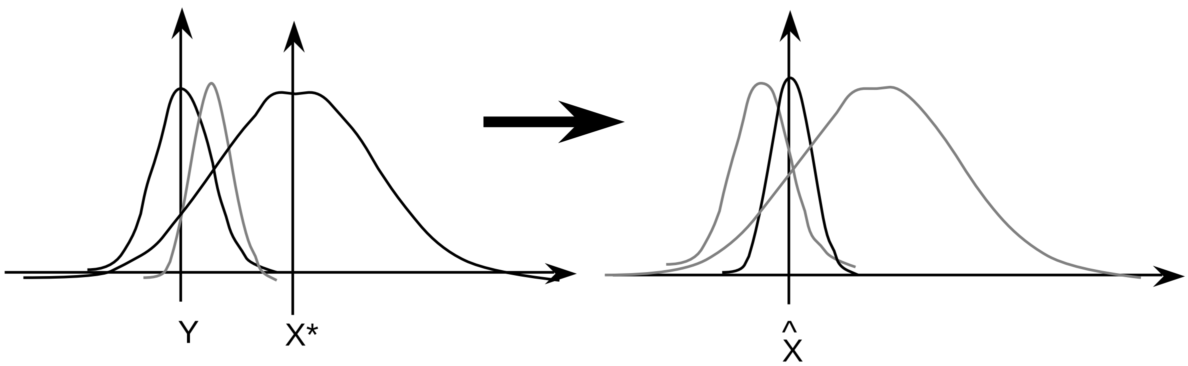
$\mathrm{Arg}_k\mathrm{Min}\,\{d_k^2\}$ tel que $d_k \leq$ seuil.

Le seuil est un multiple de l'écart type.
Son choix depend du problème

Mise à Jour

Contraint d'un modèle par une Observation

Soit un modèle X_t^*, $\mathbf{P}_t^*$ et une observation $\vec{Y}_\tau$, $\mathbf{P}_y$



Y fournis une contrainte sur X^*.
La nouvelle valeur est donnée par un moyen, pondérée par le variance:
Les variances combine comme les résistances parallèles :

$$\hat{\sigma}_x^2 = \frac{\sigma_x^2 \sigma_y^2}{\sigma_y^2 + \sigma_x^2}$$

et

$$\hat{X}_{nt} = \hat{\sigma}_x^2 \left(\frac{X^*}{\sigma_x^{*2}} + \frac{Y}{\sigma_y^2} \right)$$

Formulation Kalman (estimation récursif): Le Gain de Kalman

$$K = \frac{\sigma_x^2}{\sigma_y^2 + \sigma_x^2}$$

K determine la contribution que l'innovation porte à l'esimation $K = \frac{\partial X}{\partial V}$

Puis :

$$\hat{\sigma}_x^2 = \sigma^*{}_x^2 - K \sigma^*{}_x^2$$

et

$$\hat{X} = X^* + K (Y - Y^*)$$

En plusieurs dimensions, le gain de Kalman est une matrice.
Supposons que Y et X sont dans le même espace. (Même nombre de variables.

$$\mathbf{K} := \mathbf{P}_t^* (\mathbf{P}_t^* + \mathbf{P}_y)^{-1}$$

puis :

$$\hat{X}_t = X_t^* + \mathbf{K} (\vec{Y}_\tau - Y_t^*)$$

et

$$\hat{\mathbf{P}}_t = \mathbf{P}_t^* - \mathbf{K} \; \mathbf{P}_t^*.$$

Et si X et Y ne sont pas du même repère,
il faut le transformation $\mathbf{H}_X^Y$ et son inverse $\mathbf{H}_Y^X = \mathbf{H}_X^{Y\,T}$

Équations VI, VII et VIII du Filtre de Kalman :

VI : $\mathbf{K} := \mathbf{P}_t^* \; \mathbf{H}_X^{Y\,T} [\mathbf{H}_X^Y \; \mathbf{P}_t^* \mathbf{H}_X^{Y\,T} + \mathbf{P}_y]^{-1}$

VII : $\hat{X}_t := X_t^* + \mathbf{K} [\vec{Y}_t - \mathbf{H}_X^Y X_t^*]$

VIII : $\hat{\mathbf{P}}_t := \mathbf{P}_t^* - \mathbf{K} \mathbf{H}_X^Y \mathbf{P}_t^*$

Le Matrice de Gain K

Le gain **K** est une matrice qui donne le poids pour une mise à jour de x_d à partir de chaque composant de l'innovation v_o

$$\hat{X}_t := \mathbf{P}_t^* - \mathbf{K} \mathbf{H}_X^Y \mathbf{P}_t^*$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} k_{x1y1} & k_{x1y2} \\ k_{x2y1} & k_{x2y2} \\ k_{x3y1} & k_{x3y2} \\ k_{x4y1} & k_{x4y2} \end{pmatrix} \begin{pmatrix} y_1 - H_X^1 X_t^* \\ y_2 - H_X^2 X_t^* \end{pmatrix} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \\ \frac{\partial x_3}{\partial y_1} & \frac{\partial x_3}{\partial y_2} \\ \frac{\partial x_4}{\partial y_1} & \frac{\partial x_4}{\partial y_2} \end{pmatrix} \begin{pmatrix} y_1 - H_X^1 X_t^* \\ y_2 - H_X^2 X_t^* \end{pmatrix}$$

$$\mathbf{K} := \mathbf{P}_t^* \; \mathbf{H}_X^{Y\,T} [\mathbf{H}_X^Y \; \mathbf{P}_t^* \mathbf{H}_X^{Y\,T} + \mathbf{P}_y]^{-1}$$

$$\begin{pmatrix} k_{x1y1} & k_{x1y2} \\ k_{x2y1} & k_{x2y2} \\ k_{x3y1} & k_{x3y2} \\ k_{x4y1} & k_{x4y2} \end{pmatrix} = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \sigma_{13}^2 & \sigma_{14}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \sigma_{23}^2 & \sigma_{24}^2 \\ \sigma_{31}^2 & \sigma_{32}^2 & \sigma_{33}^2 & \sigma_{34}^2 \\ \sigma_{41}^2 & \sigma_{42}^2 & \sigma_{43}^2 & \sigma_{44}^2 \end{pmatrix} \begin{pmatrix} h_{11} & h_{21} \\ h_{12} & h_{22} \\ h_{13} & h_{23} \\ h_{14} & h_{24} \end{pmatrix} \begin{pmatrix} \bullet & \bullet \\ \bullet & \bullet \end{pmatrix}^{-1}$$

ou

$$\begin{pmatrix} \bullet & \bullet \\ \bullet & \bullet \end{pmatrix} = \left(\begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \end{pmatrix} \begin{pmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \sigma_{13}^2 & \sigma_{14}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \sigma_{23}^2 & \sigma_{24}^2 \\ \sigma_{31}^2 & \sigma_{32}^2 & \sigma_{33}^2 & \sigma_{34}^2 \\ \sigma_{41}^2 & \sigma_{42}^2 & \sigma_{43}^2 & \sigma_{44}^2 \end{pmatrix} \begin{pmatrix} h_{11} & h_{21} \\ h_{12} & h_{22} \\ h_{13} & h_{23} \\ h_{14} & h_{24} \end{pmatrix} + \begin{pmatrix} \sigma_{11}^2 & \sigma_{12}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 \end{pmatrix} \right)$$