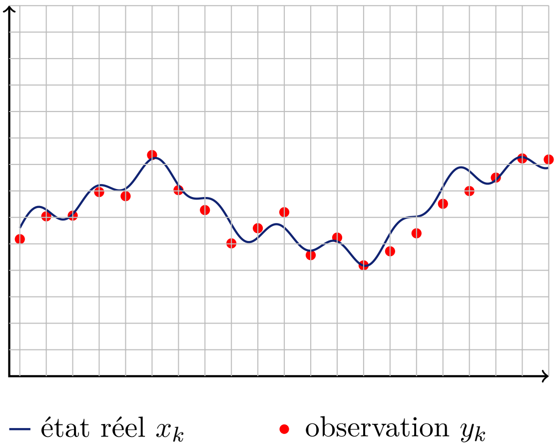
Filtres de Kalman : lissage de données et estimation de paramètres

Philippe Ricka

27 septembre 2018



Que peut-on tirer des observations y_k ?

Modèle

On suppose :

- $\exists (f_k)_k, \forall k \in \mathbb{N},$
$$\boxed{x_{k+1} = f_k(x_k) + w_k}$$
où $w_k \sim \mathcal{N}(0, R_k)$

- $\exists (h_k)_k, \forall k \in \mathbb{N},$
$$\boxed{y_k = h_k(x_k) + v_k}$$
où $v_k \sim \mathcal{N}(0, Q_k)$

Modèle

On suppose :

- $\exists (f_k)_k, \forall k \in \mathbb{N}$,

$$\boxed{x_{k+1} = f_k(x_k) + w_k}$$

où $w_k \sim \mathcal{N}(0, R_k)$

- $\exists (h_k)_k, \forall k \in \mathbb{N}$,

$$\boxed{y_k = h_k(x_k) + v_k}$$

où $v_k \sim \mathcal{N}(0, Q_k)$

Modèle

On suppose :

- $\exists (f_k)_k, \forall k \in \mathbb{N}$,
$$\boxed{x_{k+1} = f_k(x_k) + w_k}$$
où $w_k \sim \mathcal{N}(0, R_k)$

- $\exists (h_k)_k, \forall k \in \mathbb{N}$,
$$\boxed{y_k = h_k(x_k) + v_k}$$
où $v_k \sim \mathcal{N}(0, Q_k)$

Dynamique

$f_k \longleftarrow$ connaissance de la dynamique du phénomène

$w_k \longleftarrow$ ce qu'on *ne connait pas* sur la dynamique : bruit Gaussien

- Souvent, $f_k(x) = f(x, \theta_k)$ où θ_k est un paramètre du modèle
- Si on ne sait rien de la physique, on prend $f_k = Id, \forall k \in \mathbb{N}$

Dynamique

$f_k \longleftarrow$ connaissance de la dynamique du phénomène

$w_k \longleftarrow$ ce qu'on *ne connait pas* sur la dynamique : bruit Gaussien

- Souvent, $f_k(x) = f(x, \theta_k)$ où θ_k est un paramètre du modèle
- Si on ne sait rien de la physique, on prend $f_k = Id, \forall k \in \mathbb{N}$

Dynamique

$$f_k \longleftarrow \text{connaissance de la dynamique du phénomène}$$

$$w_k \longleftarrow \text{ce qu'on } \textit{ne connait pas} \text{ sur la dynamique : bruit Gaussien}$$

- Souvent, $f_k(x) = f(x, \theta_k)$ où θ_k est un paramètre du modèle
- Si on ne sait rien de la physique, on prend $f_k = Id, \forall k \in \mathbb{N}$

Observation

$h_k \longleftarrow$ fonction d'observation : décrit les infos disponibles sur l'état

$v_k \longleftarrow$ erreur de mesure : bruit Gaussien

- Si on observe directement l'état, on a :

$$h_k = Id, \forall k \in \mathbb{N}$$

En bref

On dispose a priori :

- de mesures $(y_k)_k$
- d'une info sur la confiance qu'on accorde aux mesures Q_k
- d'un modèle physique $(f_k)_k$
- d'un estimation initiale de l'état x_0^f

En bref

On dispose a priori :

- de mesures $(y_k)_k$
- d'une info sur la confiance qu'on accorde aux mesures Q_k
- d'un modèle physique $(f_k)_k$
- d'un estimation initiale de l'état x_0^f

En bref

On dispose a priori :

- de mesures $(y_k)_k$
- d'une info sur la confiance qu'on accorde aux mesures Q_k
- d'un modèle physique $(f_k)_k$
- d'un estimation initiale de l'état x_0^f

En bref

On dispose a priori :

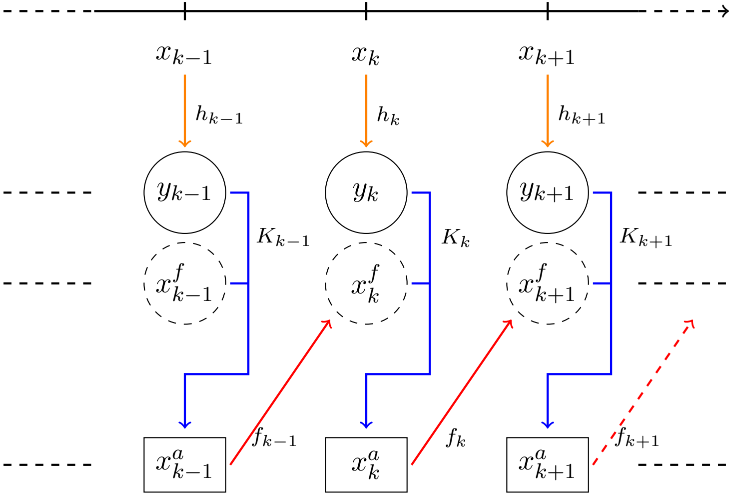
- de mesures $(y_k)_k$
- d'une info sur la confiance qu'on accorde aux mesures Q_k
- d'un modèle physique $(f_k)_k$
- d'un estimation initiale de l'état x_0^f

En bref

On dispose a priori :

- de mesures $(y_k)_k$
- d'une info sur la confiance qu'on accorde aux mesures Q_k
- d'un modèle physique $(f_k)_k$
- d'un estimation initiale de l'état x_0^f

Idée générale



Modèle

Le filtre de Kalman discret :

- f_k est linéaire de matrice A_k
- h_k est linéaire de matrice H_k

$$\begin{cases} x_{k+1} = A_k x_k + w_k \\ y_k = H_k x_k + v_k \end{cases}$$

Modèle

Le filtre de Kalman discret :

- f_k est linéaire de matrice A_k
- h_k est linéaire de matrice H_k

$$\begin{cases} x_{k+1} = A_k x_k + w_k \\ y_k = H_k x_k + v_k \end{cases}$$

Modèle

Le filtre de Kalman discret :

- f_k est linéaire de matrice A_k
- h_k est linéaire de matrice H_k

$$\begin{cases} x_{k+1} = A_k x_k + w_k \\ y_k = H_k x_k + v_k \end{cases}$$

Analyse

Au temps k, on dispose :

- d'une mesure y_k
- d'une prévision de l'état courant x_k^f (issue du temps $k-1$)

Analyse "optimale" des données :

$$\boxed{x_k^a = L_k x_k^f + G_k y_k}$$

où L et G sont des matrices à définir.

Erreurs

On définit :

- l'erreur de prédiction :

$$e_k^f = x_k^f - x_k$$

- l'erreur d'analyse :

$$e_k^a = x_k^a - x_k$$

Proposition :
Si $\mathbb{E}(e_k^f) = 0$ et si $L_k = I - G_k H_k$, alors :

$$\mathbb{E}(e_k^a) = 0$$

Covariances d'erreurs

On définit :

- la covariance de l'erreur de prédiction :

$$P_k^f = \mathbb{E}(e_k^f \cdot^{\mathrm{T}} e_k^f)$$

- la covariance de l'erreur d'analyse :

$$P_k^a = \mathbb{E}(e_k^a \cdot^{\mathrm{T}} e_k^a)$$

Gain de Kalman

Proposition :

$$P_k^a = (I - G_k H_k) P_k^{f\ \mathrm{T}} (I - G_k H_k) + G_k R_k^{\mathrm{T}} G_k$$

$$P_{k+1}^f = A_k P_k^{a\ \mathrm{T}} A_k - Q_k$$

Le gain de Kalman est $K_k = \operatorname{argmin} \left(G \mapsto \mathrm{Tr}(P_k^a) \right)$.

Théorème :

$$\begin{aligned} K_k &= P_k^{f\ \mathrm{T}} H_k \left(H_k P_k^{f\ \mathrm{T}} H_k + R_k \right)^{-1} \\ &= P_{xy_k} \qquad\qquad P_{yy_k}^{-1} \end{aligned}$$

En résumé

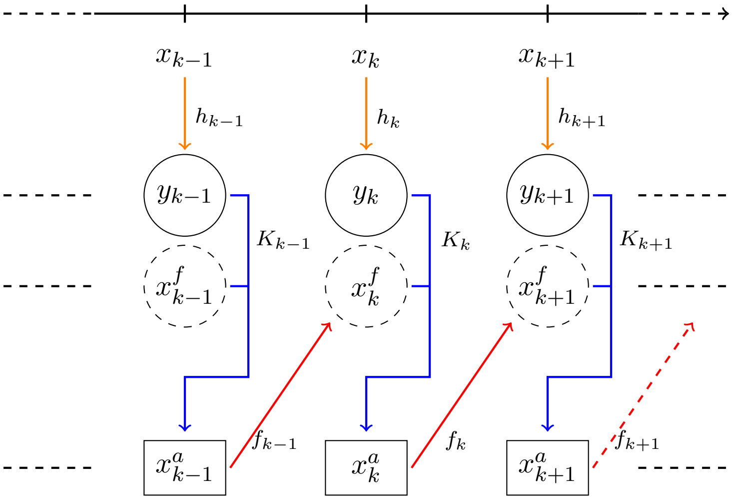
Analyse

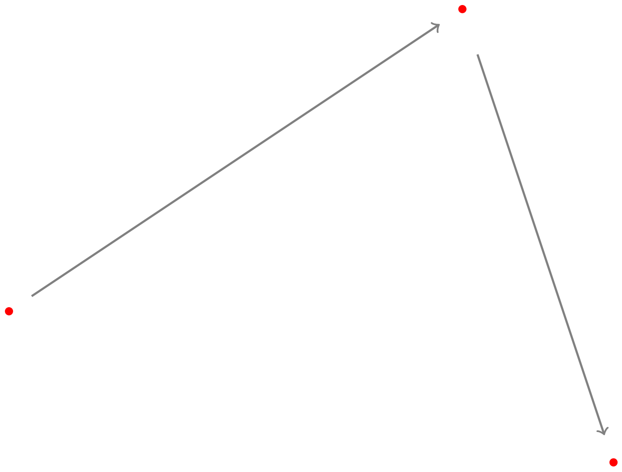
$$\left| \begin{array}{l} K_k = P_{xy_k} P_{yy_k}^{-1} \\ x_k^a = x_k^f + K_k(y_k - H_k x_k^f) \\ P_k^a = (I - K_k H_k) P_k^f \end{array} \right.$$

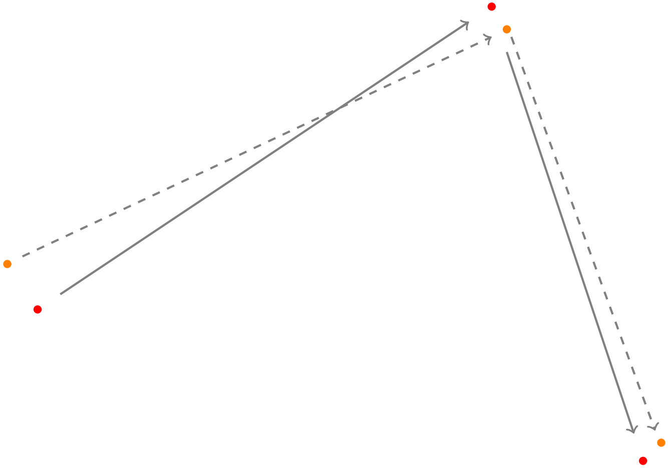
Prévision

$$\left| \begin{array}{l} x_{k+1}^f = A_k x_k^a \\ P_{k+1}^f = A_k P_k^a {}^{\mathrm{T}} A_k + Q_k \end{array} \right.$$

Rappel


x_{k-1}
x_k
x_{k+1}
h_{k-1}
h_k
h_{k+1}
y_{k-1}
y_k
y_{k+1}
K_{k-1}
K_k
K_{k+1}
x^f_{k-1}
x^f_k
x^f_{k+1}
f_{k-1}
f_k
f_{k+1}
x^a_{k-1}
x^a_k
x^a_{k+1}





$\boxed{\text{Incertitude modélisée par une Gaussienne}}$

Deux stratégies :

- Monte-Carlo $\longrightarrow$ Ensemble Kalman Filter (EnKF)
- "précis et efficace" $\longrightarrow$ Uhlmann-Kalman Filter (UKF)

$\boxed{\text{Incertitude modélisée par une Gaussienne}}$

Deux stratégies :

- Monte-Carlo $\longrightarrow$ Ensemble Kalman Filter (EnKF)
- "précis et efficace" $\longrightarrow$ Uhlmann-Kalman Filter (UKF)

$\boxed{\text{Incertitude modélisée par une Gaussienne}}$

Deux stratégies :

- Monte-Carlo $\longrightarrow$ Ensemble Kalman Filter (EnKF)
- "précis et efficace" $\longrightarrow$ Uhlmann-Kalman Filter (UKF)

Sigma-points

État estimé $x_k^f \mapsto \mathcal{N}(x_k^f, \Sigma_{x_k}) \longrightarrow$ on propage une gaussienne, représentée par un ensemble pondéré de $2n+1$ points :

$$\begin{aligned}
X_0 &= x_k^f & W_0 &= \frac{\kappa}{n+\kappa} \\
X_i &= x_k^f + \left[\sqrt{(n+\kappa)\Sigma_{\mathbf{x}}}\right]_i & W_i &= \frac{1}{2(n+\kappa)} \\
X_{n+i} &= x_k^f - \left[\sqrt{(n+\kappa)\Sigma_{\mathbf{x}}}\right]_i & W_{n+i} &= \frac{1}{2(n+\kappa)}
\end{aligned}$$

On note de plus $Y_i = h_k(X_i)$ et $\overline{Y}$ leur moyenne.

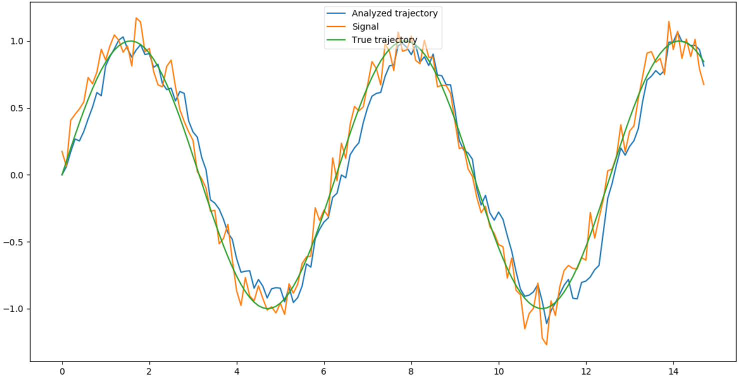
Gain de Kalman

Dans ce contexte, on définit :

$$P_{xy_k} = \sum_{i=0}^{2n+1} W_i (X_i - x_k^f)(Y_i - \overline{Y})$$

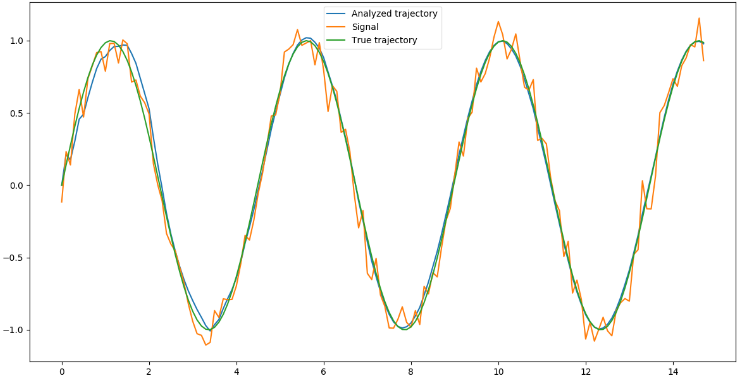
$$P_{yy_k} = \sum_{i=0}^{2n+1} W_i (Y_i - \overline{Y})(Y_i - \overline{Y}) + R_k$$

Estimation de l'état, marche aléatoire)



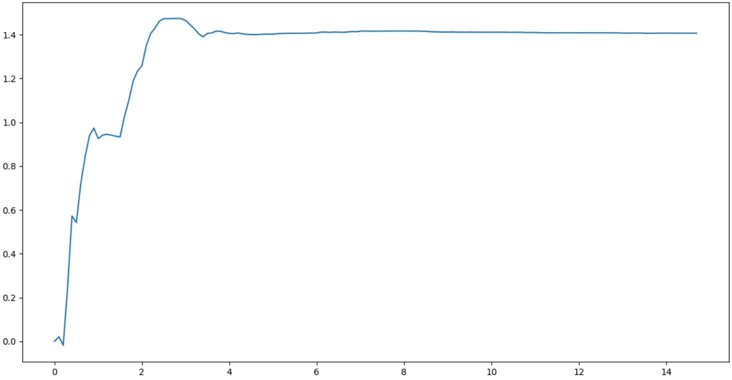
$t \mapsto \sin(t)$	$f(x) = x$	$h(x) = x$	erreur L^2 relative
$dt = 0.1$	$R_k = 0$	$Q_k = 0.1$	0.22

Estimation de l'état, physique connue



$$\boxed{\text{erreur } L^2 \text{ relative : } 0.14}$$

Estimation du paramètre



$$\boxed{\text{erreur quadratique finale : } 4.8\ 10^{-3}}$$

Merci pour votre attention !