FILTRE DE KALMAN (J, N3)
(26 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

Le **filtre de KALMAN** est un exemple classique de « **filtrage** » d'un **processus stochastique**, mis en oeuvre eg en **théorie du signal**.

(i) On considère un **système aléatoire** linéaire et stationnaire, en **temps** discret (T = **N**), défini par la donnée :

(a) d'un processus Z = $\{(\Omega, \mathcal{T}, P), (\mathbf{R}^G, \mathcal{B}(\mathbf{R}^G)), (Z_t)_{t \in T}\}$, parfois appelé **signal utile**, ou **signal pur** ;

(b) d'un processus Y = $\{(\Omega, \mathcal{T}, P), (\mathbf{R}^G, \mathcal{B}(\mathbf{R}^G)), (Y_t)_{t \in T}\}$, parfois appelé **signal perturbé**, ou **signal observé**, ou **signal mesuré** ;

(c) d'un processus X = $\{(\Omega, \mathcal{T}, P), (\mathbf{R}^K, \mathcal{B}(\mathbf{R}^K)), (X_t)_{t \in T}\}$ appelé **processus des états** (cf **état**, **espace des états**) ;

(d) de deux processus de **bruit blanc** u = $(u_t)_{t \in T}$ et v = $(v_t)_{t \in T}$, le premier étant dit **bruit perturbateur**, ou **bruit de mesure**, et le second **bruit du système** ;

(e) d'un **système d'équations linéaires** qui s'écrit :

(1) $Y_t = Z_t + u_t$ (**équation d'état**),

(2) $Z_t = A X_t$ (**équations de mesure**),

(3) $X_{t+1} = B X_t + v_t$ (**équation de transition** d'état à état, ou entre états),

expressions dans lesquelles $A \in M_{GK}(\mathbf{R})$ est une **matrice de mesure** et $B \in M_K(\mathbf{R})$ une **matrice de transition**, ou **matrice de dynamique** (à distinguer de l'autre notion de **matrice de transition**).

On cherche à transformer Y pour obtenir un **estimateur** optimal (au sens quadratique) $Z^{\sim}$ du signal Z (cf **optimalité**).

(ii) Le **filtre de R.E. KALMAN** est un filtre récursif et linéaire qui permet d'obtenir un « **estimateur** » (ou prévision) $X^{\sim}$ de X, donc, d'après (2), l'estimateur cherché $Z^{\sim}$ de Z.

La construction du filtrage s'effectue comme suit :

(a) on suppose que u et v sont des bruits gaussiens, avec $V\,u_t = \Sigma_{uu}$, $V\,v_t = \Sigma_{vv}$ et $C\,(u_t , v_t) = \Sigma_{uv}$, $\forall\, t \in T$;

(b) le « meilleur » estimateur $X_t^{\sim}$ de X_t minimise la **variance** de l'**erreur** :

(4) $E_t = X_t^{\sim} - X_t$.

Ce n'est autre que l'**espérance conditionnelle** :

(5) $X_t^{\sim} = E(X_t / Y_{t-1}, Y_{t-2}, ...)$;

(c) on montre que :

(6) $X_{t+1}^{\sim} = E(X_{t+1} / Y_{t-1}, ...) + E(X_{t+1} / K_t)$,

où $K_t = Y_t - E(Y_t / Y_{t-1}, ...)$ définit l'**innovation** de Y_t . Par suite :

(7) $$\begin{aligned} X_{t+1}^{\sim} &= B X_t^{\sim} + E(X_{t+1} / K_t), \\ K_t &= Y_t - A X_t^{\sim}. \end{aligned}$$

La **formule d'ajustement (de R.E. KALMAN)** s'écrit finalement :

(8) $X_{t+1}^{\sim} = B X_t^{\sim} + K(t)(Y_t - A X_t^{\sim})$,

où $K(t) \in M_{KG}(\mathbf{R})$, appelée **matrice de R.E. KALMAN**, s'exprime selon :

(9) $K(t) = \{B \Sigma(t) A' + \Sigma_{uv}'\} . \{A \Sigma(t) A' + \Sigma_{uu}\}^{-1}$,

espression dans laquelle on note, $\forall t \in T$, les matrices de dispersion et de covariance selon :

(10) $$\begin{aligned} V u_t &= \Sigma_{uu}, \\ V v_t &= \Sigma_{vv}, \\ C(u_t, v_t) &= \Sigma_{uv}, \end{aligned}$$

et où $\Sigma(t) \in M_K(\mathbf{R})$ est, $\forall t \in T$, déterminée par récurrence selon :

(11) $\Sigma(t+1) = B \Sigma(t) B' + \Sigma_{vv} - \{B \Sigma(t) A' + \Sigma_{uv}'\} \{A \Sigma(t) A' + \Sigma_{uu}\}^{-1} \{A' \Sigma(t) B + \Sigma_{uv}\}$,

expression dans laquelle $\Sigma(0)$ désigne la solution de la **condition initiale** $\Sigma(0) = \Sigma_{vv} + B \Sigma(0) B'$.

(iii) Le **filtre de R.E. KALMAN** se décrit de façon analogue pour un **système en temps continu**.

(iv) Dans le cadre du **modèle de régression**, le **filtre de R.E. KALMAN** peut s'introduire de la manière suivante. On dispose de deux jeux d'observations $(y_t)_{t=1,...,T}$ à valeurs dans $\mathbf{R}^G$ et $(X_t)_{t=1,...,T}$ à valeurs dans $\mathbf{R}^K$. On suppose ces observations liées par un **modèle linéaire** (**équation d'observation**, ou **équation « observée »**) dans un **espace d'observation** $\mathbf{R}^N$) :

(12) $y_T = X_t b_t + u_t$, avec $u_t \sim \mathcal{N}_G(0, \Sigma_{uu}(t))$ (**loi normale** centrée),

dans laquelle u_t représente l'erreur d'observation et les **paramètres** aléatoires b_t varient selon une équation, dite **équation de système** :

(13) $b_t = M_t b_{t-1} + v_t$, avec $v_t \sim \mathcal{N}_K (0, \Sigma_{vv}(t))$.

On suppose aussi que $M_t \in M_K(\mathbf{R})$, que $\Sigma_{uu}(t)$ et $\Sigma_{vv}(t)$ sont connues, $\forall t \in T$, et que u_t et v_τ sont indépendants (ou seulement non corrélés) si $\tau \neq t$.

Le **théorème de BAYES** conduit à la **loi conditionnelle** a posteriori de b_{t-1} :

(14) $\mathcal{L}(b_{t-1} / y_{t-1}, \ldots) = \mathcal{N}_K (\mu_{t-1}, \Sigma_{t-1})$,

avec $\mu_{t-1} = E(b_{t-1} / y_{t-1}, \ldots)$ et $\Sigma_{t-1} = V(b_{t-1} / y_{t-1}, \ldots)$.

La **procédure récursive de R.E. KALMAN** s'effectue ici en deux étapes. A l'instant $t \in N_T^*$:

(a) avant l'observation de y_t , la loi conditionnelle a priori de b_t se déduit de (13) et (14), ie :

(15) $\mathcal{L}(b_t / y_{t-1}) = \mathcal{N}_K (M_t \mu_{t-1}, M_t \Sigma_{t-1} M_t' + \Sigma_{vv}(t))$;

(b) après observation de y_t , la loi conditionnelle a posteriori de b_t devient :

(16) $\mathcal{L}(b_t / u_t^\wedge, y_{t-1}, \ldots) = \mathcal{N}_K (\mu_t^\wedge, \Sigma_t^\wedge)$,

expression dans laquelle $u_t^\wedge = y_t - y_t^\wedge$ représente l'**erreur de prévision** des moindres carrés portant sur y_t (avec $y_t^\wedge = X_t M_t \mu_{t-1}$), et où :

$$\mu_t^\wedge = M_t \mu_{t-1} + N_t X_t' \{\Sigma_{uu}(t) + X_t N_t X_t'\}^{-1} u_t^\wedge,$$

(17) $$\Sigma_t^\wedge = N_t - N_t X_t' \{\Sigma_{uu}(t) + X_t N_t X_t'\}^{-1} X_t N_t ,$$

$$N_t = M_t \Sigma_{t-1} M_t' + \Sigma_{vv}(t).$$

En effet, dans (16), il est équivalent de conditionner b_t pr à $u_t^\wedge$ ou pr à y_t puisque la donnée de $u_t^\wedge = y_t - y_t^\wedge$ équivaut à celle de y_t ($y_t^\wedge$ est connue car X_t , M_t et μ_{t-1} le sont).

On peut ainsi interpréter le **filtrage de R.E. KALMAN** comme une procédure bayésienne de **mise à jour récursive** :

(a) estimation préliminaire de b_t ;

(b) puis prise en compte de y_t (ou de l'erreur u_t) à l'aide d'une terme correctif dans (17) (dans laquelle le terme précédant $u_t^\wedge$ n'est autre que l'estimateur des moindres carrés obtenu par régression de b_t sur u_t (cf **méthode des moindres carrés**).

La procédure peut alors s'amorcer dès que (μ_0 , Σ_0) est connu.