**HMMA237**

# Filtre linéaire et de Kalmann.

*Cours: Joseph Salmon* *Scribes: El Hadji Leye et Mbaye Cheikh Tidiane*

## Introduction

## 1 Modèle graphique

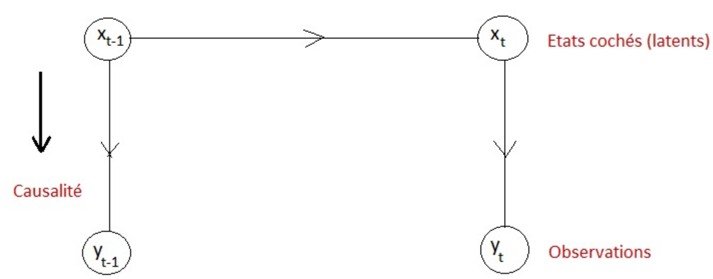



Figure 1: Filtrage.

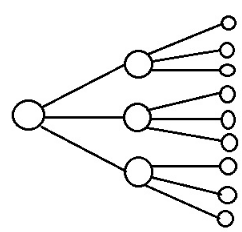


Figure 2: Réseau de neurones.

## 2 Modèles d'états, le cas linéaire avec bruit gaussien.

Soit un processus $(x_t)_{t\in\mathbb{N}}$ où à chaque instant $t$ l'état $x_t \in \mathbb{R}^p$. Le processus $(y_t)_{t\in\mathbb{N}}$ représente les observations et pour tout $t$, $\quad y_t \in \mathbb{R}^q$.

On suppose que l'on observe $y_1, y_2, \ldots, y_s$. Des vecteurs de bruits gaussiens centrés et i.i.d $\varepsilon_t \in \mathbb{R}^p$ et $\xi_t \in \mathbb{R}^q$ ont des variances respectives $\Sigma_\varepsilon$ et $\Sigma_\xi$. De plus on suppose que $\varepsilon_t$ et $\xi_t$ sont indépendants.

**Remark.** *Il est facile de rajouter des variables exogènes si besoin.*

**Exemple.**

- *Marche aléatoire bruitée.*
  $x_t = x_{t-1} + \varepsilon_t$ *où* $\varepsilon_t$ *est de variance* $\sigma_\varepsilon^2$.
  $y_t = x_t + \xi_t$ *où* $\xi_t$ *est de variance* $\sigma_\xi^2$.

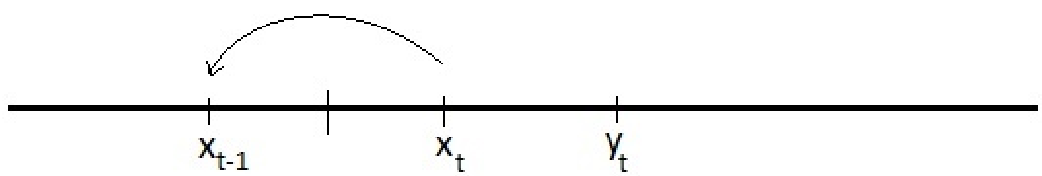



Figure 3: Marche aléatoire.

- *Modèle structurel exponentiel.*

$$\begin{cases} x_t = \varphi x_{t-1} + \varepsilon_t \\ y_t = x_t + s_t + \xi_t \end{cases} \tag{1}$$

*$s_t$ est une saisonnalité trimestrielle et $s_t+s_{t-1}+s_{t-2}+s_{t-3} = \eta_t$ avec $\eta_t$ un gaussien centré de variance $\sigma_\eta^2$. On peut écrire ce modèle sous forme matricielle:*

$$\begin{pmatrix} x_t \\ s_t \\ s_{t-1} \\ s_{t-2} \end{pmatrix} = \begin{pmatrix} \varphi & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ \eta_t \\ 0 \\ 0 \end{pmatrix}$$

*et*

$$y_t = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_t \\ s_t \\ s_{t-1} \\ s_{t-2} \end{pmatrix} + \xi_t.$$

*avec* $\varepsilon_t$, $\xi_t$ *et* $\eta_t$ *sont indépendants.*

$$\mathrm{Var}\begin{pmatrix}\varepsilon_t\\ \eta_t\\ 0\\ 0\end{pmatrix} = \begin{pmatrix}\sigma_\varepsilon^2 & 0 & 0 & 0\\ 0 & \sigma_\eta^2 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$$

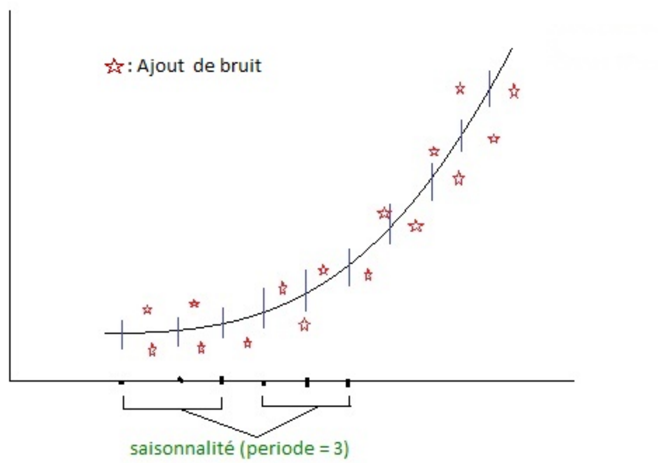
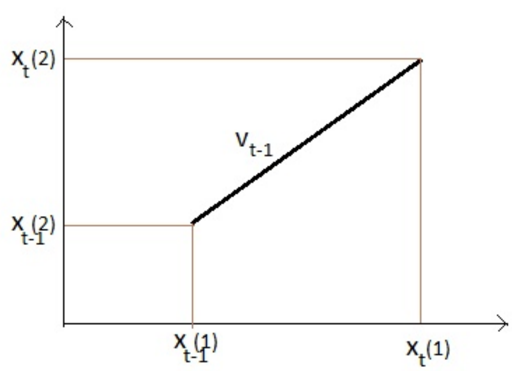



Figure 4: Modèle structurel exponentiel.

**Exemple.** *Poursuite d'objet: exemple farfelu d'un drone qui suit un chat dans un champs...*

$$\begin{pmatrix}x_t(1)\\ s_t(2)\\ v_t(1)\\ v_t(2)\end{pmatrix} = \begin{pmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 0 & 1 & \varphi & 0\\ 0 & 0 & 0 & \varphi\end{pmatrix}\begin{pmatrix}x_{t-1}(1)\\ s_{t-1}(2)\\ v_{t-1}(1)\\ v_{t-1}(2)\end{pmatrix}$$

*les* $x_t(1)$ *et* $x_t(2)$ *constituent les états,* $v_t(1)$ *et* $v_t(2)$ *les vitesses et* $\varphi$ *est l'accélération constante.*



## 3 Modèle d'observation

On modélise les observations sous forme matricielle par

$$\begin{pmatrix} y_t(1) \\ y_t(2) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_t(1) \\ s_t(2) \\ v_t(1) \\ v_t(2) \end{pmatrix} + \mathcal{N}\left(0, \begin{pmatrix} \tau_2 & 0 \\ 0 & \tau_2 \end{pmatrix}\right) \quad . \tag{2}$$

On observe uniquement la position de l'objet au lieu de la vitesse. Intéressons nous au modèle suivant:

$$\begin{cases} x_t = \varphi x_{t-1} + \varepsilon_t \\ y_t = A x_t + \xi_t \end{cases} \tag{3}$$

On initialise l'espace états par $x_0$ qui suit la loi $\mathcal{N}(\mu_0, \Sigma_0)$.

**Notation:** On note l'espérance conditionnelle de $x_t$ sachant $y_1, y_2, \ldots, y_s$ par

$$x_t^s := \mathbb{E}\left[(x_t | y_{1:s})\right] \quad \text{avec} \quad y_{1:s} = (y_1, y_2, \ldots, y_s.)$$

et la covariance conditionnelle en $t_1, t_2$ sachant qu'on a observé jusqu'à $s$ par la matrice de variance conditionnelle

$$\mathsf{P}_{t_1,t_2}^s := \mathbb{E}\left[\left(x_{t_1} - x_{t_1}^s\right)\left(x_{t_2} - x_{t_2}^s\right)^\top\right].$$

On note aussi que $\mathsf{P}_t^s = \mathsf{P}_{t,t}^s$.

**Filtre de Kalman:**

**Definition.** *Le filtre de Kalman est un filtre linéaire aux observations $y_1, \ldots, y_s$. On le définit par:*

$$x_t^t = \sum_{s=1}^{+\infty} B_s y_s \quad \textit{avec} \quad B_s \quad \textit{une matrice.} \tag{4}$$

**Theorem.** *Conditions d'initialisation: $x_0^0 = \mu_0$ et $\mathsf{P}_0^0 = \Sigma_0$.*
*Par récurrence sous l'hypothèse d'évolution décrite en (3), on obtient:*

1. $x_t^{t-1} = \Phi x_{t-1}^{t-1} \in \mathbb{R}^p$ *c'est la prédiction.*

2. $\mathsf{P}_t^{t-1} = \Phi \mathsf{P}_{t-1}^{t-1} \Phi^\top + \Sigma_\varepsilon \in \mathbb{R}^{p \times p}$.

$$\begin{cases} x_t^t = x_t^{t-1} + K_t\left(y_t - A x_t^{t-1}\right) \\ \\ \mathsf{P}_t^t = \left(\mathrm{Id} - K_t A\right) \mathsf{P}_t^{t-1} \end{cases} \quad K_t \textit{ est défini par } \mathsf{P}_t^{t-1} A^\top \left(A \mathsf{P}_t^{t-1} A^\top + \Sigma_\xi\right)^{-1}$$

*3. Le terme de résidu (ou d'erreur entre la prédiction et l'observation) est:*

$$\begin{aligned}\delta_t &= y_t - \mathbb{E}\left(y_t|y_{1:t-1}\right)\\ &= y_t - \mathbb{E}\left(Ax_t|y_{1:t-1}\right)\\ &= y_t - Ax_t^{t-1}\end{aligned}$$

*4. La variance de l'erreur est:*

$$\begin{aligned}\Sigma_{\delta_t} &= \mathrm{Var}(\delta_t)\\ &= \mathrm{Var}(y_t - Ax_t^{t-1})\\ &= \mathrm{Var}(Ax_t + \xi_t - Ax_t^{t-1})\\ &= \mathrm{Var}(\xi_t) + \mathrm{Var}(A(x_t - x_t^{t-1}))\\ &= \Sigma_\xi + A\,\mathrm{Var}(x_t - x_t^{t-1})A^\top\end{aligned}$$

*On retrouve* $\mathsf{P}_t^{t-1} = \mathrm{Var}(x_t - x_t^{t-1})$.

*Proof.* 1.

$$\begin{aligned}x_t^{t-1} &= \mathbb{E}\left(x_t|y_{1:t-1}\right)\\ &= \mathbb{E}\left(\Phi x_{t-1}|y_{1:t-1}\right)\\ &= \Phi\mathbb{E}\left(x_{t-1}|y_{1:t-1}\right)\\ &= \Phi x_{t-1}^{t-1}.\end{aligned}$$

2.

$$\begin{aligned}\mathsf{P}_t^{t-1} &= \mathbb{E}\left[\left(x_t - x_t^{t-1}\right)\left(x_t - x_t^{t-1}\right)^\top\right]\\ &= \mathbb{E}\left[\left(\Phi\left(x_t - x_{t-1}^{t-1}\right) + \varepsilon_t\right)\left(\Phi\left(x_{t-1} - x_{t-1}^{t-1}\right) = \varepsilon_t\right)^\top\right]\\ &= Var\left(\varepsilon_t\right) + \mathrm{Var}(\Phi(x_t - x_{t-1}^{t-1}))\\ &= \Sigma_\varepsilon \;+\; \Phi\mathsf{P}_{t-1}^{t-1}\Phi^\top\end{aligned}$$

3. On a $\mathrm{Cov}(\delta_s, y_t) = 0$ pour tout $t > s$.
   En effet

$$\begin{aligned}\mathrm{Cov}(\delta_s, y_t) &= \mathbb{E}\left[(\delta_t - \mathbb{E}(\delta_t))(y_s - \mathbb{E}(y_s))^\top\right]\\ &= \mathbb{E}\mathbb{E}\left[(\delta_t - \mathbb{E}(\delta_t))(y_s - \mathbb{E}(y_s))^\top|y_{1:s}\right]\\ &= 0 \text{ par mesurabilité.}\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}(x_s, \delta_t|y_{1:t-1}) &= \mathrm{Cov}(x_t, y_t - Ax_x^{t-1}|y_{1:t-1})\\ &= \mathrm{Cov}(x_t - x_t^{t-1}, y_t - Ax_t^{t-1}|y_{1:t-1})\\ &= \mathrm{Cov}(x_t - x_t^{t-1}, A(x_t - x_t^{t}) + \xi_t|y_{1:t-1})\\ &= \mathrm{Var}(x_t - x_t^{t-1})A^\top\\ &= \mathsf{P}_t^{t-1}A^\top.\end{aligned}$$

Ainsi, $\begin{pmatrix} x_t \\ \delta_t \end{pmatrix}$ suit une loi normale $\mathcal{N}\left(\begin{pmatrix} x_t^{t-1} \\ 0 \end{pmatrix}, \begin{pmatrix} \mathsf{P}_t^{t-1} & \mathsf{P}_t^{t-1}A^\top \\ \mathsf{P}_t^{t-1}A^\top & \Sigma_{\delta_t} \end{pmatrix}\right)$ avec $A\mathsf{P}_t^{t-1} = \mathsf{P}_t^{t-1}A^\top$ et $\Sigma_{\delta_t} = A\mathsf{P}_t^{t-1}A^\top + \Sigma_\xi$.
On notera $\mu_{i|j}$ l'espérance de la distribution conditionnelle de $x_i$ sachant $x_j$ et $\Sigma_{i|j}$ sa variance. Les paramètres conditionnels sont alors donnés par les relations suivantes:

$$\begin{cases} \mu_{i|j} = \mu_i + \Sigma_{ij}\Sigma_{jj}^{-1}(x_j - \mu_j) \\ \\ \Sigma_{i|j} = \Sigma_{ii} - \Sigma_{ij}^\top \Sigma_{jj}^{-1}\Sigma_{ij} \end{cases} \tag{5}$$

Ainsi,

$$x_t^t = \mathbb{E}(x_t | y_{1:t}, \delta_t) = x_t^{t-1} + (\mathsf{P}_t^{t-1}A^\top)(\Sigma_j)^{-1}(\delta_t) \ . \tag{6}$$

□