

# Vecteurs gaussiens

Jérémie Bigot

Cours de probabilités MA105  
ISAE/SUPAERO 1A

**Année 2013 - 2014**

# Introduction

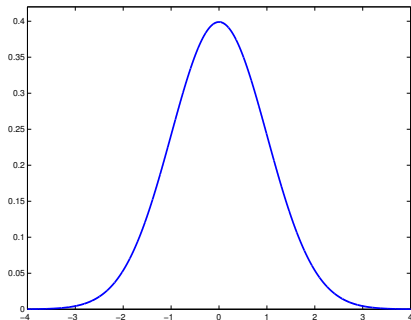
- Les lois de probabilité gaussiennes sont les plus employées en physique et en règle générale en statistique de grande dimension
- Elles sont stables par transformation linéaire
- Elles sont stables par conditionnement
- Ce cours : cas de modèles de dimension finie

- 1 Loi normale sur  $\mathbb{R}$
- 2 Vecteur gaussien sur  $\mathbb{R}^n$
- 3 Conditionnement
- 4 Filtre de Kalman

## Définitions - Loi $\mathcal{N}(0, 1)$

Une v.a.  $X \in \mathbb{R}$  est dite de loi normale  $\mathcal{N}(0, 1)$  si sa densité est

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

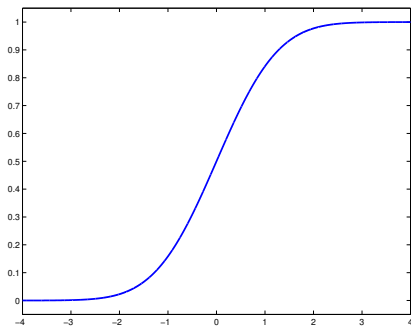


Densité  $f_X(x)$  de la loi normale  $\mathcal{N}(0, 1)$

## Définitions - Loi $\mathcal{N}(0, 1)$

La fonction de répartition de  $X \sim \mathcal{N}(0, 1)$  est

$$F_X(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx, \quad t \in \mathbb{R}.$$



Fonction de répartition  $F_X(t)$  de la loi normale  $\mathcal{N}(0, 1)$

## Définitions - Loi $\mathcal{N}(m, \sigma^2)$

Soit  $m \in \mathbb{R}$  et  $\sigma > 0$ .

Une v.a.  $Y \in \mathbb{R}$  est dite de loi normale  $\mathcal{N}(m, \sigma^2)$  si sa densité est

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-m)^2}{2\sigma^2}}, \quad y \in \mathbb{R}.$$

On a que

$$\mathbb{E}(Y) = m \quad \text{et} \quad \text{Var}(Y) = \sigma^2,$$

ainsi que

$$Y = m + \sigma X,$$

où  $X \sim \mathcal{N}(0, 1)$ .

# Fonction caractéristique

## Définition

Soit  $X \in \mathbb{R}$  une v.a. de densité  $f_X$ . La fonction caractéristique de  $X$  est la fonction à valeurs complexes définie par

$$\forall t \in \mathbb{R}, \phi_X(t) = \mathbb{E}(e^{itX}) = \int_{\mathbb{R}} e^{itx} f_X(x) dx.$$

**Remarques :** la fonction caractéristique  $\phi_X(t)$  est la transformée de Fourier de  $f_X$ . Par la formule d'inversion de Fourier, on a donc que

$$\forall x \in \mathbb{R}, f_X(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \phi_X(t) dt.$$

**La loi d'une v.a.r. réelle (à densité) est donc entièrement déterminée par sa fonction caractéristique.**

# Fonction caractéristique

## Définition

Soit  $X \in \mathbb{R}$  une v.a. de densité  $f_X$ . La fonction caractéristique de  $X$  est la fonction à valeurs complexes définie par

$$\forall t \in \mathbb{R}, \phi_X(t) = \mathbb{E}(e^{itX}) = \int_{\mathbb{R}} e^{itx} f_X(x) dx.$$

## Proposition (Cas gaussien)

Soit  $X$  une variable aléatoire de loi normale  $\mathcal{N}(m, \sigma^2)$ . Alors

$$\forall t \in \mathbb{R}, \phi_X(t) = \exp\left(itm - \frac{1}{2}t^2\sigma^2\right).$$



# Fonction caractéristique d'un vecteur aléatoire

## Définition

Soit  $X = (X_1, \dots, X_n)'$  un vecteur aléatoire de dimension  $n$  de densité  $f_X$ . La fonction caractéristique de  $X$  est la fonction à valeurs complexes définie par

$$\begin{aligned}\forall t = (t_1, \dots, t_n)' \in \mathbb{R}^n, \phi_X(t) &= \mathbb{E} \left( e^{i\langle t, X \rangle} \right) = \mathbb{E} \left( e^{i \sum_{k=1}^n t_k X_k} \right) \\ &= \int_{\mathbb{R}^n} e^{i \sum_{k=1}^n t_k x_k} f_X(x_1, \dots, x_n) dx_1 \dots dx_n.\end{aligned}$$

**Remarques :** la fonction caractéristique  $\phi_X(t)$  est la transformée de Fourier de  $f_X$ . Par la formule d'inversion de Fourier, on a donc que

$$\forall x = (x_1, \dots, x_n)' \in \mathbb{R}^n, f_X(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i \sum_{k=1}^n t_k x_k} \phi_X(t) dt.$$

**La loi d'un vecteur aléatoire (à densité) est donc entièrement déterminée par sa fonction caractéristique.**

# Fonction caractéristique et indépendance

## Proposition

Soit  $X = (X_1, \dots, X_n)'$  un vecteur aléatoire de dimension  $n$ . Les coordonnées de  $X$  sont indépendantes si et seulement si la fonction caractéristique de  $X$  est le produit des fonctions caractéristiques de ses coordonnées i.e.

$$\begin{aligned}\forall t = (t_1, \dots, t_n)' \in \mathbb{R}^n, \phi_X(t) &= \mathbb{E} \left( e^{i \sum_{k=1}^n t_k X_k} \right) = \mathbb{E} \left( \prod_{k=1}^n e^{it_k X_k} \right) \\ &= \prod_{k=1}^n \mathbb{E} \left( e^{it_k X_k} \right) \\ &= \prod_{k=1}^n \phi_{X_k}(t_k)\end{aligned}$$

- 1 Loi normale sur  $\mathbb{R}$
- 2 Vecteur gaussien sur  $\mathbb{R}^n$**
- 3 Conditionnement
- 4 Filtre de Kalman

# Définition d'un vecteur gaussien

## Définition

*On dit que  $X = (X_1, \dots, X_n)'$  est un vecteur aléatoire gaussien, si toute combinaison linéaire à coefficients réels des coordonnées  $X_1, \dots, X_n$  de  $X$  suit une loi normale de moyenne  $m$  et de variance  $\sigma^2$  (avec éventuellement  $\sigma^2 = 0$ ) i.e.*

$$\forall u = (u_1, \dots, u_n) \in \mathbb{R}^n, \langle u, X \rangle = \sum_{k=1}^n u_k X_k \sim \mathcal{N}(m, \sigma^2)$$

*où  $m$  et  $\sigma^2$  dépendent du vecteur  $u = (u_1, \dots, u_n)$ .*

**Remarque :** en particulier, chacune des coordonnées  $X_i$  de  $X$  suit une loi normale.

# Exemple de vecteur gaussien

## Proposition

Soit  $X_1, \dots, X_n$  des variables aléatoires réelles **indépendantes** de loi normale  $\mathcal{N}(0, 1)$  i.e. centrée réduite.

Alors, le vecteur  $X = (X_1, \dots, X_n)'$  est un vecteur aléatoire gaussien.

## Exemple de vecteur gaussien

**Preuve :** Soit  $u_1, \dots, u_n$  des réels, et  $Y = \sum_{k=1}^n u_k X_k$  une combinaison linéaire des coordonnées de  $X = (X_1, \dots, X_n)'$ .

En utilisant le fait que les  $X_k$  sont des v.a. indépendantes de loi  $\mathcal{N}(0, 1)$ , on a que la fonction caractéristique de  $Y$  est

$$\begin{aligned} \forall t \in \mathbb{R}, \phi_Y(t) &= \mathbb{E} \left( e^{it \sum_{k=1}^n u_k X_k} \right) = \mathbb{E} \left( \prod_{k=1}^n e^{itu_k X_k} \right) = \prod_{k=1}^n \mathbb{E} \left( e^{itu_k X_k} \right) \\ &= \prod_{k=1}^n \phi_{X_k}(tu_k) = \prod_{k=1}^n \exp \left( -\frac{t^2 u_k^2}{2} \right) \\ &= \exp \left( -\left( \sum_{k=1}^n u_k^2 \right) \frac{t^2}{2} \right). \end{aligned}$$

**Donc** la v.a.  $Y$  suit donc une loi normale d'espérance nulle et de variance  $\sum_{k=1}^n u_k^2$ , ce qui montre que  $X$  est un vecteur gaussien.

# Fonction caractéristique d'un vecteur gaussien

## Proposition

Soit  $X = (X_1, \dots, X_n)'$  un vecteur aléatoire gaussien de dimension  $n$ , d'espérance  $m = (m_1, \dots, m_n) \in \mathbb{R}^n$  et de matrice de covariance  $\Sigma$ . Alors, la fonction caractéristique de  $X$  est

$$\forall t \in \mathbb{R}^n, \phi_X(t) = \exp\left(it'm - \frac{1}{2}t'\Sigma t\right).$$

**Remarque :** la loi d'un vecteur gaussien est donc caractérisée par la donnée de son espérance et de sa matrice de covariance.

**Transformation linéaire :** soit  $X \sim \mathcal{N}(m, \Sigma)$  et  $Y = Ax + b$  où  $A \in \mathbb{R}^{m \times n}$  et  $b \in \mathbb{R}^m$ , alors

$$Y \sim \mathcal{N}(m_Y, \Sigma_Y), \text{ avec } m_Y = Am + b \text{ et } \Sigma_Y = A\Sigma A'.$$

## Fonction caractéristique d'un vecteur gaussien

**Preuve :** Soit  $t = (t_1, \dots, t_n) \in \mathbb{R}^n$  et posons  $Y = \sum_{k=1}^n t_k X_k$ . Comme  $Y$  est une combinaison linéaire des coordonnées du vecteur gaussien  $X$ , on a que  $Y$  suit une loi normale d'espérance

$$\mathbb{E}(Y) = \sum_{k=1}^n t_k m_k = t' m,$$

et de variance

$$\text{var}(Y) = t' \Sigma t.$$

On a donc que la fonction caractéristique de  $Y$  est

$$\forall s \in \mathbb{R}, \phi_Y(s) = \exp\left(is(t'm) - \frac{1}{2}s^2 t' \Sigma t\right) = \mathbb{E}(e^{isY}) = \mathbb{E}(e^{is(\sum_{k=1}^n t_k X_k)}),$$

et le résultat s'obtient finalement en remarquant que

$$\phi_X(t) = \mathbb{E}\left(e^{i \sum_{k=1}^n t_k X_k}\right) = \phi_Y(1).$$



# Cas gaussien : indépendance $\equiv$ décorrélation

## Proposition

Soit  $X = (X_1, \dots, X_n)'$  un vecteur aléatoire gaussien d'espérance  $m \in \mathbb{R}^n$  de matrice de covariance  $\Sigma_X = [\text{Cov}(X_i, X_j)]_{1 \leq i, j \leq n}$ .

Les variables aléatoires  $X_1, \dots, X_n$  sont **indépendantes** si et seulement si la matrice  $\Sigma_X$  **est diagonale i.e.**

$$\text{Cov}(X_i, X_j) = 0$$

pour tout  $i \neq j$ .

**Notations :**  $m = (m_1, \dots, m_n)' \in \mathbb{R}^n$  et  $\text{Diag}(\Sigma_X) = (\sigma_1^2, \dots, \sigma_n^2)' \in \mathbb{R}^n$ .

## Cas gaussien : indépendance $\equiv$ décorrélation

**Preuve :** Supposons que la matrice  $\Sigma_X$  soit diagonale.

On a que  $m_k$  et  $\sigma_k^2$  sont l'espérance et la variance de  $X_k$ . On a alors que la fonction caractéristique de  $X$  s'écrit

$$\begin{aligned}\forall t = (t_1, \dots, t_n) \in \mathbb{R}^n, \phi_X(t) &= \exp\left(it'm - \frac{1}{2}t'\Sigma t\right) \\ &= \exp\left(i \sum_{k=1}^n t_k m_k - \frac{1}{2} \sum_{k=1}^n t_k^2 \sigma_k^2\right) \\ &= \prod_{k=1}^n \exp\left(it_k m_k - \frac{1}{2} t_k^2 \sigma_k^2\right) \\ &= \prod_{k=1}^n \phi_{X_k}(t_k).\end{aligned}$$

**Donc** on en déduit que les v.a.  $X_k$  sont indépendantes. La réciproque est immédiate.

# Densité d'un vecteur gaussien

## Proposition

Soit  $X = (X_1, \dots, X_n)'$  un vecteur aléatoire gaussien d'espérance  $m \in \mathbb{R}^n$  et de matrice de covariance  $\Sigma$ .

Si  $\Sigma$  est **définie positive** alors  $X$  admet pour densité la fonction  $f_X$  définie par

$$\forall x \in \mathbb{R}^n, f_X(x) = \frac{1}{(2\pi)^{n/2}} \frac{1}{\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(x-m)'\Sigma^{-1}(x-m)\right).$$

**Remarque :** analogie avec le cas uni-dimensionnel ( $n = 1$ ). Si  $X \in \mathbb{R}$  suit loi normale  $\mathcal{N}(m, \sigma^2)$ , alors sa densité est

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

## Densité d'un vecteur gaussien

**Preuve :** la matrice  $\Sigma$  (symétrique à valeurs réelles et définie positive) est diagonalisable sous la forme

$$\Sigma = Q\Lambda Q',$$

où  $Q$  est une matrice orthogonale (i.e.  $Q'Q = I_n$ ), et  $\Lambda$  est une matrice diagonale, dont les termes diagonaux  $\lambda_i, i = 1, \dots, n$  sont des réels strictement positifs (valeurs propres de  $\Sigma$ ). Posons

$$Y = Q'(X - m).$$

On a que

$$Y \sim \mathcal{N}(m_Y, \Sigma_Y),$$

avec  $m_Y = \mathbb{E}(Y) = Q'(\mathbb{E}(X) - m) = 0$  et  $\Sigma_Y = Q'\Sigma Q = \Lambda$ .

## Densité d'un vecteur gaussien

**Preuve (suite)** : comme  $Y \sim \mathcal{N}(0, \Lambda)$  avec  $\Lambda$  matrice diagonale, les coordonnées  $Y_i, i = 1, \dots, n$  de  $Y$  sont indépendantes et de loi  $\mathcal{N}(0, \lambda_i)$ .

Par conséquent, la densité du vecteur aléatoire  $Y$  est

$$\begin{aligned} \forall y = (y_1, \dots, y_n) \in \mathbb{R}^n, f_Y(y) &= \prod_{i=1}^n f_{Y_i}(y_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \frac{1}{\lambda_i^{1/2}} \exp\left(-\frac{y_i^2}{2\lambda_i}\right) \\ &= \frac{1}{(2\pi)^{n/2}} \frac{1}{\sqrt{\prod_{i=1}^n \lambda_i}} \exp\left(-\frac{1}{2}y' \Lambda^{-1}y\right). \end{aligned}$$

## Densité d'un vecteur gaussien

**Preuve (suite) :** Soit  $\varphi$  l'application de  $\mathbb{R}^n$  dans  $\mathbb{R}^n$  définie par

$$\varphi(y) = Qy + m.$$

Comme  $Y = Q'(X - m)$ , on a donc que  $X = \varphi(Y)$  et  $Y = \varphi^{-1}(X)$  avec

$$\varphi^{-1}(x) = Q'(x - m)$$

application linéaire de jacobien  $Jac(\varphi^{-1})(x) = \pm 1$  ( $Q$  orthogonale).

Par la formule de changement de variables, la densité de  $X$  est donnée par

$$\begin{aligned} \forall x \in \mathbb{R}^n, f_X(x) &= |Jac(\varphi^{-1})(x)| f_Y(\varphi^{-1}(x)) = f_Y(Q'(x - m)) \\ &= \frac{1}{(2\pi)^{n/2}} \frac{1}{\sqrt{\prod_{i=1}^n \lambda_i}} \exp\left(-\frac{1}{2}(x - m)' Q \Lambda^{-1} Q'(x - m)\right) \\ &= \frac{1}{(2\pi)^{n/2}} \frac{1}{\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(x - m)' \Sigma^{-1}(x - m)\right) \end{aligned}$$

en utilisant le fait que  $\det(\Sigma) = \prod_{i=1}^n \lambda_i$  et que  $\Sigma^{-1} = Q \Lambda^{-1} Q'$ .

- 1 Loi normale sur  $\mathbb{R}$
- 2 Vecteur gaussien sur  $\mathbb{R}^n$
- 3 Conditionnement**
- 4 Filtre de Kalman

## Conditionnement entre vecteurs gaussiens

Soit  $Z$  un vecteur aléatoire gaussien de dimension  $n$  divisé en deux blocs de coordonnées  $X$  et  $Y$  de dimensions  $n_X \geq 1$  et  $n_Y \geq 1$  telles que  $n = n_X + n_Y$ , i.e.

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix}.$$

- Comment caractériser l'espérance conditionnelle  $\mathbb{E}(Y|X)$  de  $Y$  sachant  $X$  ?
- Quelle est la loi conditionnelle de  $Y$  sachant  $X$  ?



# Structure hilbertienne

Soit  $(\Omega, \mathcal{A}, P)$  un espace probabilisé, et  $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$  une v.a. réelle.

## Définition

*On note par  $L^2(\Omega, \mathcal{A}, P)$  l'espace des v.a. réelles  $X$  de carré intégrable i.e telles que  $\mathbb{E}(X^2) < +\infty$ .*

## Théorème

*L'espace  $L^2(\Omega, \mathcal{A}, P)$  est un espace de Hilbert lorsqu'il est muni du produit scalaire  $\langle X, Y \rangle = \mathbb{E}(XY)$ , et de la norme quadratique associée  $\|X\|_2 = \sqrt{\mathbb{E}(X^2)}$ .*

**Remarque :** si  $\begin{pmatrix} X \\ Y \end{pmatrix} \in \mathbb{R}^2$  est un vecteur gaussien d'espérance nulle. Alors **orthogonalité de  $X$  et  $Y$  i.e.**  $\langle X, Y \rangle = \mathbb{E}(XY) = 0$  et **indépendance de  $X$  et  $Y$**  sont des propriétés équivalentes car

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(XY) = \langle X, Y \rangle.$$

# Structure hilbertienne

Soit  $Z = (Z_1, \dots, Z_n)' \in \mathbb{R}^n$  un vecteur aléatoire gaussien d'espérance nulle, i.e.  $\mathbb{E}(Z_i) = 0$  pour tout  $i = 1, \dots, n$ .

## Définition

Soient  $\{i_1, \dots, i_p\} \subset \{1, \dots, n\}$  avec  $p \leq n$ . Le sous-espace vectoriel (fermé) engendré par les v.a.  $Z_{i_1}, \dots, Z_{i_p}$  est

$$\text{Vect}(Z_{i_1}, \dots, Z_{i_p}) = \overline{\left\{ \sum_{k=1}^p \alpha_k Z_{i_k}, \alpha_k \in \mathbb{R} \right\}}$$

**Remarque :** les éléments de  $\text{Vect}(Z_{i_1}, \dots, Z_{i_p})$  sont des v.a. réelles de loi normale et d'espérance nulle.

# Espérance conditionnelle et projection orthogonale

Soit  $Z = (Z_1, \dots, Z_n)' \in \mathbb{R}^n$  un vecteur aléatoire gaussien d'espérance nulle, i.e.  $\mathbb{E}(Z_i) = 0$  pour tout  $i = 1, \dots, n$ .

## Proposition

Soit  $Y \in L^2(\Omega, \mathcal{A}, P)$  une v.a.r. telle que  $(Y, Z_{i_1}, \dots, Z_{i_p}) \in \mathbb{R}^{p+1}$  est un vecteur aléatoire gaussien d'espérance nulle. Alors, l'espérance conditionnelle de  $Y$  sachant le vecteur  $\tilde{Z} = (Z_{i_1}, \dots, Z_{i_p})$ , notée  $\mathbb{E}(Y|Z_{i_1}, \dots, Z_{i_p})$  ou  $\mathbb{E}(Y|\tilde{Z})$ , est la projection orthogonale de  $Y$  sur l'espace

$$\text{Vect}(Z_{i_1}, \dots, Z_{i_p}) = \overline{\left\{ \sum_{k=1}^p \alpha_k Z_{i_k}, \alpha_k \in \mathbb{R} \right\}}.$$

**Remarques :**  $\mathbb{E}(Y|Z_{i_1}, \dots, Z_{i_p})$  est donc une v.a. gaussienne

**Stabilité de la loi gaussienne par conditionnement !**

# Espérance conditionnelle et projection orthogonale

Soit  $Z \in \mathbb{R}^n$  et  $W \in \mathbb{R}^m$  deux vecteurs aléatoires gaussien indépendants et d'espérance nulle

## Proposition

Soit  $Y \in L^2(\Omega, \mathcal{A}, P)$  une v.a.r. telle que  $(Y, Z', W')$   $\in \mathbb{R}^{n+m+1}$  est un vecteur aléatoire gaussien d'espérance nulle. Alors, l'espérance conditionnelle de  $Y$  sachant le vecteur  $\begin{pmatrix} Z \\ W \end{pmatrix}$ , notée  $\mathbb{E}(Y|Z, W)$ , est la projection orthogonale de  $Y$  sur l'espace

$$\text{Vect}(Z, W) = \overline{\left\{ \sum_{k=1}^n \alpha_k Z_k + \sum_{j=1}^m \beta_j W_j, \alpha_k, \beta_j \in \mathbb{R} \right\}} = \text{Vect}(Z) \oplus \text{Vect}(W)$$

et donc

$$\mathbb{E}(Y|Z, W) = \mathbb{E}(Y|Z) + \mathbb{E}(Y|W).$$

## Loi conditionnelle

Soit  $Z$  un vecteur aléatoire gaussien de dimension  $n$  divisé en deux blocs de coordonnées  $X$  et  $Y$  de dimensions  $n_X \geq 1$  et  $n_Y \geq 1$  telles que  $n = n_X + n_Y$ , i.e.

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix}.$$

L'espérance de  $Z$  s'écrit alors sous la forme

$$\mathbb{E}(Z) = \begin{pmatrix} \mathbb{E}(X) \\ \mathbb{E}(Y) \end{pmatrix},$$

et la matrice de covariance  $\Sigma_Z$  de  $Z$  peut se décomposer de la façon suivante

$$\Sigma_Z = \begin{pmatrix} \Sigma_X & \Sigma_0 \\ \Sigma_0' & \Sigma_Y \end{pmatrix},$$

où

$$\Sigma_0 = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))').$$

# Loi conditionnelle

Soit  $Z$  un vecteur aléatoire gaussien de dimension  $n$  divisé en deux blocs de coordonnées  $X$  et  $Y$  de dimensions  $n_X \geq 1$  et  $n_Y \geq 1$  telles que  $n = n_X + n_Y$ .

La notion de densité conditionnelle vue précédemment pour les couples de variables aléatoires réelles peut être étendue au cas des vecteurs aléatoires.

## Définition

*La densité conditionnelle de  $Y$  sachant  $X = x$  est alors la fonction*

$$f_Y^{X=x}(y) = \frac{f_Z(x, y)}{f_X(x)},$$

# Loi conditionnelle

## Proposition

Supposons que  $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$  soit un vecteur gaussien, et que la matrice de covariance  $\Sigma_X$  de  $X$  est définie positive. Soit  $x \in \mathbb{R}^{n_X}$ .

La loi conditionnelle de  $Y$  sachant  $X = x$  est alors une **loi gaussienne** d'espérance

$$\mathbb{E}(Y|X = x) = \mathbb{E}(Y) + \Sigma'_0 \Sigma_X^{-1} (x - \mathbb{E}(X))$$

et de matrice de covariance

$$\Sigma_Y^{X=x} = \Sigma_Y - \Sigma'_0 \Sigma_X^{-1} \Sigma_0.$$

**Remarque :** on retrouve bien que  $\mathbb{E}(Y|X)$  est une v.a. gaussienne

$$\mathbb{E}(Y|X) = \mathbb{E}(Y) + \Sigma'_0 \Sigma_X^{-1} (X - \mathbb{E}(X)).$$

## Loi conditionnelle

**Cas particulier** : régression linéaire de  $Y \in \mathbb{R}$  par rapport  $X \in \mathbb{R}$

On a que  $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$  est un vecteur aléatoire gaussien de dimension  $n = 2$ , avec  $X$  et  $Y$  v.a. de dimension  $n_X = n_Y = 1$

La matrice de covariance  $\Sigma_Z$  de  $Z$  peut se décomposer de la façon suivante

$$\Sigma_Z = \begin{pmatrix} \Sigma_X & \Sigma_0 \\ \Sigma_0' & \Sigma_Y \end{pmatrix},$$

avec  $\Sigma_X = \text{Var}(X)$ ,  $\Sigma_Y = \text{Var}(Y)$ , et  $\Sigma_0 = \text{Cov}(X, Y)$ .

**Espérance conditionnelle** :  $\mathbb{E}(Y|X) = \mathbb{E}(Y) + \frac{\text{Cov}(X,Y)}{\text{Var}(X)}(X - \mathbb{E}(X))$   
i.e.

$$\mathbb{E}(Y|X) = aX + b,$$

avec  $a = \frac{\text{Cov}(X,Y)}{\text{Var}(X)}$  et  $b = \mathbb{E}(Y) - a\mathbb{E}(X)$ .

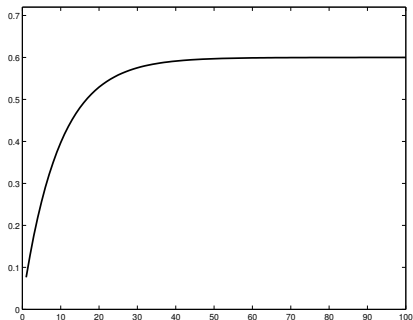


- 1 Loi normale sur  $\mathbb{R}$
- 2 Vecteur gaussien sur  $\mathbb{R}^n$
- 3 Conditionnement
- 4** Filtre de Kalman

## Exemple d'utilisation du filtre de Kalman

Modèle d'évolution de la vitesse d'un véhicule décrit par un réel  $x_t$  au cours du temps  $t \in \mathbb{N}$  qui vérifie une équation du type (modèle idéal) :

$$x_{t+1} = \alpha x_t = \alpha^{t+1} x_0, \text{ avec } \alpha \in ]0, 1[.$$

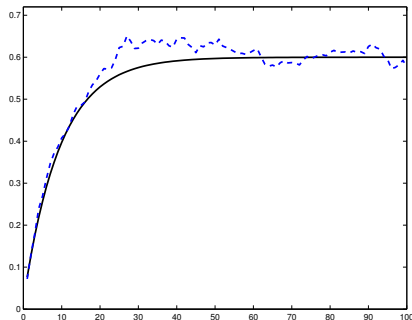


Modèle idéal :  $t \mapsto b + x_t$ , avec  $b = 0.6$ ,  $x_0 = -0.5817$ ,  $\alpha = 0.9$ , et  $t = 1, \dots, 100$

# Exemple d'utilisation du filtre de Kalman

La vitesse réelle  $X_t$  est modélisée en introduisant un terme perturbateur aléatoire

$$X_{t+1} = \alpha X_t + V_{t+1}, \text{ avec } V_{t+1} \text{ v.a. centrée,}$$

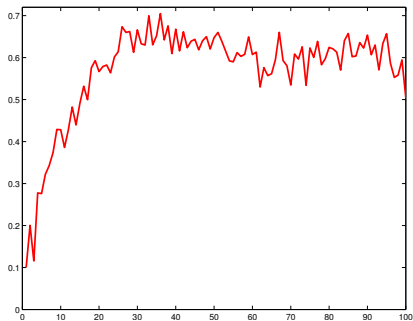


Vitesse réelle :  $t \mapsto b + X_t, t = 1, \dots, 100$

# Exemple d'utilisation du filtre de Kalman

Suivi de la vitesse réelle  $X_t$  à partir de mesures **entachées d'erreurs**  $Y_t$  qui donnent une **approximation** de la vitesse réelle :

$$Y_t = X_t + W_t, \text{ avec } W_t \text{ v.a. centrée,}$$

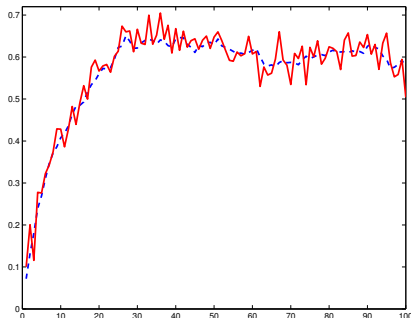


Vitesse mesurée :  $t \mapsto b + Y_t, t = 1, \dots, 100$

# Exemple d'utilisation du filtre de Kalman

Suivi de la vitesse réelle  $X_t$  à partir de mesures  $Y_t$  qui donnent une **approximation** de la vitesse réelle :

$$Y_t = X_t + W_t, \text{ avec } W_t \text{ v.a. centrée,}$$



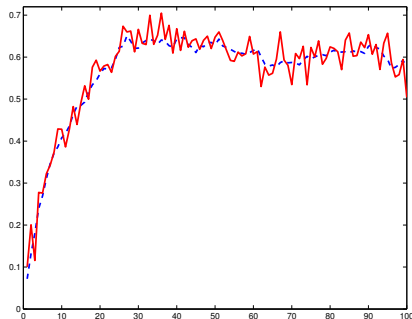
Vitesse réelle :  $t \mapsto b + X_t, t = 1, \dots, 100$

Vitesse mesurée :  $t \mapsto b + Y_t, t = 1, \dots, 100$

# Exemple d'utilisation du filtre de Kalman

Suivi de la vitesse réelle  $X_t$  à partir de mesures  $Y_t$  qui donnent une **approximation** de la vitesse réelle :

$$Y_t = X_t + W_t, \text{ avec } W_t \text{ v.a. centrée,}$$



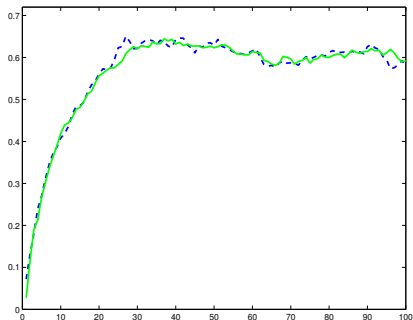
**Question :** connaissant les mesures  $Y_1, \dots, Y_t$ , comment prédire/estimer la vraie position du mobile au temps  $t + 1$  ?

**Réponse :** utilisation de l'espérance conditionnelle (filtre de Kalman)

# Exemple d'utilisation du filtre de Kalman

**Prédiction/prévision** de la vitesse réelle  $X_{t+1}$  par le filtre de Kalman :

$$\tilde{X}_{t+1} = \mathbb{E}(X_{t+1} | Y_0, \dots, Y_t)$$



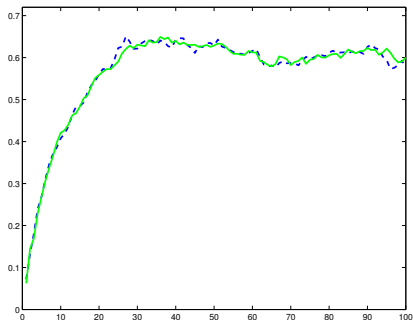
Vitesse réelle :  $t \mapsto b + X_t, t = 1, \dots, 100$

Vitesse prédite :  $t \mapsto b + \tilde{X}_t, t = 1, \dots, 100$

# Exemple d'utilisation du filtre de Kalman

**Estimation** de la vitesse réelle  $X_{t+1}$  par le filtre de Kalman :

$$\hat{X}_{t+1} = \mathbb{E}(X_{t+1} | Y_0, \dots, Y_{t+1})$$



Vitesse réelle :  $t \mapsto b + X_t, t = 1, \dots, 100$

Vitesse estimée :  $t \mapsto b + \hat{X}_t, t = 1, \dots, 100$



## Filtre de Kalman - Cadre général

Déplacement d'un véhicule qui peut être modélisé par une équation différentielle (**modèle théorique**) :

$$\frac{dx(t)}{dt} = f(x, t), \quad x(0) = x_0,$$

où  $x(t) \in \mathbb{R}^n$  est un vecteur donnant la position du mobile à l'instant  $t \in \mathbb{R}^+$ .

Discrétisation de cette équation différentielle (approximation d'Euler) :

$$x_{t+1} = x_t + \Delta_t f(x_t),$$

et approximation de la solution par une équation linéarisée :

$$x_{t+1} = \underbrace{A_t}_{n \times n} x_t.$$

## Filtre de Kalman - Cadre général (linéaire)

Modèle d'évolution de la trajectoire **réelle** à l'aide d'un vecteur aléatoire  $X_t \in \mathbb{R}^n$  :

$$X_0 = 0, \quad X_{t+1} = \underbrace{A_{t+1}}_{n \times n} X_t + \underbrace{m_{t+1}}_{n \times 1} + \underbrace{B_{t+1}}_{n \times n} V_{t+1}, \quad \text{avec } V_{t+1} \sim \mathcal{N}(0, I_n),$$

où  $m_{t+1} + B_{t+1}V_{t+1}$  est un terme perturbateur aléatoire.

A chaque instant  $t$  on mesure  $X_t$  totalement ou partiellement avec un bruit de mesure gaussien. Le résultat de la mesure est donc :

$$Y_t = \underbrace{C_t}_{n \times n} X_t + \underbrace{D_t}_{n \times n} W_t, \quad \text{avec } W_t \sim \mathcal{N}(0, I_n).$$

**Question** : comment construire un “bon estimateur”  $\hat{X}_{t+1}$  de  $X_{t+1}$  en tenant compte de la suite de mesures  $Y_1, \dots, Y_{t+1}$  ?

## Filtre de Kalman - Cadre général (linéaire)

Modèle d'évolution d'un vecteur aléatoire  $X_t \in \mathbb{R}^n$  :

$$X_0 = 0, \quad X_{t+1} = \underbrace{A_{t+1}}_{n \times n} X_t + \underbrace{m_{t+1}}_{n \times 1} + \underbrace{B_{t+1}}_{n \times n} V_{t+1}, \quad \text{avec } V_{t+1} \sim \mathcal{N}(0, I_n).$$

A chaque instant  $t$  on mesure  $X_t$  totalement ou partiellement avec un bruit de mesure gaussien. Le résultat de la mesure est donc :

$$Y_t = \underbrace{C_t}_{n \times n} X_t + \underbrace{D_t}_{n \times n} W_t, \quad \text{avec } W_t \sim \mathcal{N}(0, I_n).$$

**Réponse :** utilisation de l'espérance conditionnelle

$$\hat{X}_{t+1} = \mathbb{E}(X_{t+1} | Y_1, \dots, Y_{t+1})$$