

# Semaine 27 - Vecteurs aléatoires

Valentin De Bortoli  
email : valentin.debortoli@gmail.com

Dans tous les exercices  $X$  est une variable d'un espace probabilisé  $(\Omega, \mathcal{A}, P)$  à valeurs dans un espace d'états fini  $\mathcal{X}$ .

## 1 Transformée de Laplace

On définit la transformée de Laplace d'une variable aléatoire  $X$  par  $\forall t \in \mathbb{R}, \phi(t) = E(e^{tX})$ .

- 1 Montrer que  $\phi \in \mathcal{C}^\infty(\mathbb{R})$ . Calculer  $\forall n \in \mathbb{N}, \phi^{(n)}(0)$ .
- 2 Soit  $X$  une variable de Bernoulli de paramètre  $p$ . On définit  $Y = 2X - 1$ . Quelle est la loi de  $Y$  ? On dit que  $Y$  suit une loi de Rademacher.
- 3 Calculer la transformée de Laplace de  $Y$ .

## 2 Matrice de covariance

Soit  $(X_1, \dots, X_n)$  un vecteur aléatoire. On considère  $M \in \mathcal{M}_n(\mathbb{R})$  avec  $\forall (i, j) \in \llbracket 1, n \rrbracket, M(i, j) = \text{Cov}(X_i, X_j)$ .

- 1 Montrer que  $M$  est symétrique définie positive.
- 2 Soit  $(X_1^k)_{k \in \llbracket 1, N \rrbracket}$  un  $N$ -échantillon du vecteur  $X_1$ . On note  $(X_i^k)_{k \in \llbracket 1, N \rrbracket}$ . On pose  $C_{i,j}^N = \frac{1}{N-1} \sum_{k=1}^N (X_i^k - \overline{X_i})(X_j^k - \overline{X_j})$  avec  $\overline{X_i} = \frac{1}{N} \sum_{k=1}^N X_i^k$ . Calculer  $E(C_{i,j}^N)$ .
- 3 Donner une interprétation de  $C_{i,j}^N$ . Pourquoi a-t-on divisé par  $N - 1$  plutôt que  $N$  ?

**Remarque :** Si on appelle  $M_X$  la matrice des  $C_{i,j}^N$  on obtient la matrice de covariance empirique. La diagonalisation de cette matrice et la projection (orthogonale) sur les premiers vecteurs propres permet de représenter des données de grandes dimensions dans des plans (ou espaces) où la covariance est maximisée. Cette technique porte un nom : ACP (Analyse en Composantes Principales) et est très utilisée en sociologie, neuroscience...

## 3 Entropie et information

On appelle entropie de la loi  $p$  (définie sur  $\mathcal{X}$ ), la quantité suivante :

$$H(p) = - \sum_{x \in \mathcal{X}} \log(p(x))p(x)$$

On adoptera la convention  $0 \log(0) = 0$ .

- 1 Calculer l'entropie d'une loi de Bernoulli de paramètre  $p \in [0, 1]$ . Que peut-on en déduire ?
- 2 Montrer que l'entropie est toujours positive et qu'elle est toujours plus petite que  $\log(|\mathcal{X}|)$ . Dans quel cas l'entropie est-elle maximale ? Minimale ?

**3** On suppose que  $H((p, q)) = H(p) + H(q)$  (où  $(p, q)$  désigne la loi jointe de  $p$  et  $q$ ). Quelle relation lie la loi jointe et  $p$  et  $q$  ?

**Remarque :** la notion d'entropie a été introduite par Shannon dans les années 50. Elle est extrêmement utile pour coder de manière optimale une information.

## 4 Kurtosis et asymétrie

Soit  $X$  une variable aléatoire à valeurs dans  $\mathcal{X}$ . On note  $\mu$  sa moyenne et  $\sigma$  son écart-type. On définit l'asymétrie de  $X$  par  $\gamma(X) = E\left(\left(\frac{X-\mu}{\sigma}\right)^3\right)$ . On définit le kurtosis de  $X$  par  $\beta_2(X) = E\left(\left(\frac{X-\mu}{\sigma}\right)^4\right)$

- 1 Calculer l'asymétrie d'une variable de Bernoulli. Donner une interprétation du signe.
- 2 En déduire l'asymétrie d'une variable binomiale.
- 3 Reprendre les questions précédentes pour le kurtosis.

## 5 Introduction aux matrices aléatoires

Soit  $M_2$  une matrice de taille  $2 \times 2$  dont les coefficients sont des Bernoulli indépendantes de paramètre  $p$ . Soit  $M_3$  une matrice de taille  $3 \times 3$  dont les coefficients sont des Bernoulli indépendantes.

- 1 Calculer l'espérance et la variance de la trace de  $M_2$ .
- 2 Même question pour le déterminant de  $M_2$ .
- 3 Même question pour la trace de  $M_3$ .
- 4 Même question pour le déterminant de  $M_3$ .

**Remarque :** ces questions ont pour but de montrer la complexité calculatoire inhérente à la théorie des matrices aléatoires. Celle-ci s'est fortement développée dans les années 50 avec des applications en physique théorique mais aussi en théorie des nombres.

## 6 La méthode de rejet

Soit  $X$  une variable aléatoire de  $\Omega$  dans  $\mathcal{X}$ . On note  $p$  sa loi. Le but de cet exercice est de proposer une méthode pour simuler un échantillon de la loi  $X$ . On suppose qu'il existe une loi  $q$  (sur le même espace) qui vérifie :

$$\exists C \in \mathbb{R}_+^*, \forall x \in \mathcal{X}, Cq(x) \geq p(x)$$

On suppose que l'on sait échantillonner selon la loi  $q$ . On propose la démarche suivante :

- on tire un élément  $Y$  selon  $q$ .
  - on tire un élément  $U$  selon la loi uniforme sur  $\llbracket 1, q(Y) \rrbracket$ .
  - si  $U \leq p(Y)$  on pose  $Z = Y$  sinon on reprend la première étape.
- 1 Montrer que  $Z$  suit bien la loi  $p$ .
  - 2 Quelle est la probabilité d'acceptation de l'échantillon  $Z$  ? Quel est le rôle de la constante  $C$  ?

**Remarque :** le problème de simulation de variables aléatoires est un domaine de la recherche encore ouvert. De nombreux autres algorithmes existent (méthode d'inversion, méthodes de Metropolis-Hastings, méthodes de Hit-and-Run). L'algorithme de Metropolis-Hastings (le plus utilisé en pratique) fonctionne de la même manière que le rejet mais prend en compte l'information "l'échantillon est rejeté".

## 7 Estimateurs et décomposition biais-variance

On suppose que  $\mathcal{X} = \llbracket 1, N \rrbracket$ . Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon de la loi  $p$ , c'est-à-dire  $n \in \mathbb{N}$  variables aléatoires indépendantes de loi  $p$ . On suppose que la loi  $p$  est uniforme sur  $\mathcal{X}$ . On cherche à estimer  $N$ .

- 1 Proposer une manière d'estimer  $N$ . On note  $E_1$  cette nouvelle variable aléatoire.
- 2 Calculer l'espérance et la variance de  $E_1$ .
- 3 On appelle risque quadratique la quantité  $R(E_1) = E((P - E_1)^2)$ . Montrer que :

$$R(E_1) = V(E_1) + (E(E_1) - P)^2$$

Cette décomposition est appelée biais-variance. Pourquoi ? Calculer le risque quadratique  $R(E_1)$ .

- 4 On note  $E_2 = \max(X_1, \dots, X_n)$ . Calculer l'espérance de  $E_2$  ainsi que sa variance.
- 5 En déduire son risque quadratique  $R(E_2)$  et comparer avec le risque quadratique  $R(E_1)$ .

**Remarque :** le but de cet exercice est de mettre en évidence le fait suivant : il ne sert à rien de considérer un estimateur sans biais si sa variance est trop grande. Il peut être intéressant de considérer des estimateurs avec biais mais de variance petite. La borne de Cramer-Rao donne une borne inférieure pour ce risque quadratique.

## 8 Loïs et moments

- 1 Montrer que si deux variables  $X$  et  $Y$  ont les mêmes moments alors elles ont la même loi (par même moment on entend :  $\forall k \in \mathbb{N}, E(X^k) = E(Y^k)$ ).

**Remarque :** attention néanmoins. Cette propriété n'est plus vraie lorsque  $\mathcal{X}$  n'est pas fini.

## 9 Fonction caractéristique

Soit  $X$  une variable aléatoire sur  $\mathcal{X}$ . On définit  $\phi_X : \mathbb{R} \rightarrow \mathbb{C}$  par  $\phi_X(\xi) = E(e^{i\xi X})$ .

- 1 Montrer que la fonction caractéristique est  $2\pi$  périodique et  $\mathcal{C}^\infty$ .
- 2 Que vaut  $\phi_X(0)$  ?  $\phi'_X(0)$  ?  $\phi''_X(0)$  ?
- 3 Montrer que  $P(X = k) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ik\xi} \phi_X(\xi) d\xi$ . En déduire que la fonction caractéristique caractérise la loi de  $X$ .
- 4 Montrer que si  $X$  et  $Y$  sont indépendantes alors  $\phi_{X+Y} = \phi_X \phi_Y$ . En déduire facilement la fonction caractéristique de la loi binomiale.

**Remarque :** contrairement à la caractérisation par les moments, cette caractérisation est toujours valable.