École Normale Supérieure
FIMFA — Statistique 2015–2016

DEVOIR 2 POUR LE 18 NOVEMBRE

ON PEUT SUPPOSER ACQUIS LE RÉSULTAT D'UN EXERCICE QUAND ON ABORDE UN AUTRE EXERCICE.

DÉFINITIONS ET NOTATIONS

i) Pour un vecteur x de $\mathbb{R}^p$, on note

$$\|x\|_1 := \sum_{i=1}^{p} |x_i| \qquad \|x\|_0 := \sum_{i=1}^{p} \mathbb{I}_{x_i \neq 0} \qquad \|x\| := \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} .$$

ii) Pour $0 \leq s \leq p$,

$$\Theta_s := \{x : x \in \mathbb{R}^p, \|x\|_0 \leq s\} \ ,$$

c'est l'ensemble des vecteurs à au plus s coordonnées non-nulles.

iii) On se propose de reconstruire un signal inconnu $\theta \in \Theta_s$ à partir de $(\langle x_i, \theta\rangle)_{i\leq N}$ où les x_i sont des vecteurs de détection (à valeur dans $\mathbb{R}^p$). On note y le vecteur dont les coordonnées sont les $\langle x_i, \theta\rangle$.
On note A la matrice $N \times p$ dont les lignes sont les (x_i^t) $(A_{i,j} := x_i(j))$. La matrice A est appelée *matrice de détection.*

iv) Une fonction de $f : \mathbb{R}^p \longrightarrow \mathbb{R}^N$ est une ϵ-*isométrie* sur $\mathcal{T} \subseteq \mathbb{R}^p$ si pour tous a, a' de $\mathcal{T}$

$$(1-\epsilon) \leq \frac{\|f(a) - f(a')\|^2}{\|a' - a\|^2} \leq (1+\epsilon)$$

autrement dit si f préserve les distances mutuelles à l'intérieur de $\mathcal{T}$.

v) L'intersection de la sphère unité en dimension p et de Θ_s sera notée A_s^p.

OBJECTIFS

On s'intéresse à la reconstruction de θ à partir de $y \in \mathbb{R}^N$ défini par

$$y := A\theta\,.$$

Nous allons d'abord donner des conditions suffisantes sur la matrice de détection A pour garantir la possibilité de reconstruire $\theta \in \Theta_s$ à partir de θ.

CONDITIONS POUR LA RECONSTRUCTION DE $\theta \in \Theta_s$

Le but du premier exercice est de vérifier quand la méthode de reconstruction

$$\arg\min \|z\|_1 \qquad \text{sous la contrainte } Az = y \qquad (\text{MÉTHODE BP})$$

reconstruit θ.
On utilisera la condition suivante sur la matrice de détection.
Condition NSP : pour tout $S \subset [p] := \{1, \ldots, p\}$ avec $|S| \leq s$, pour tout $z \neq 0$ vérifiant $Az = 0$,

$$\|z_S\|_1 < \|z_{\overline{S}}\|_1$$

avec la convention que z_S (resp $z_{\overline{S}}$) est obtenu à partir de z en annulant tous les coefficients z_i qui ne sont pas dans S (resp. qui sont dans S).

Exercice 1
Montrer que si la condition NSP est vérifée par A, la méthode BP reconstruit θ vérifiant $\|\theta\|_0 \leq s$ à partir de θ vérifiant $y = A\theta$.

Exercice 2
Etablir la réciproque de la propriété précédente.

Exercice 3
Vérifier que si pour $\epsilon > 0$, la matrice A définit une ϵ-isométrie pour Θ_{2s}, alors la reconstruction de tout $\theta \in \Theta_s$ à partir de $y = A\theta$ est possible : il existe une unique solution dans Θ_s à l'équation en θ' $A\theta = A\theta'$.

Exercice 4
Montrer que si A définit une 1/3-isométrie pour Θ_{2s}, on peut reconstruire $\theta \in \Theta_s$ par la méthode BP.
Suggestion : Montrer que dans ce cas, la propriété NS *est vérifiée.*

Construction randomisée de matrices de détection

Exercice 5
Dans cet exercice et dans la suite W envoie $\mathbb{R}^p$ dans $\mathbb{R}^N$, $W : \alpha \mapsto \frac{1}{\sqrt{N}} \left(\langle X_1, \alpha\rangle, \ldots, \langle X_N, \alpha\rangle\right)^t$. Les vecteurs aléatoires $X_1, \ldots, X_N$ sont i.i.d. $\mathcal{N}(0, \mathrm{Id}_p)$.

i) Loi, espérance, variance de $\|W(\alpha)\|^2/\|\alpha\|^2$ (pour tout α non nul de $\mathbb{R}^p$) ?
ii) Majorer la variance de $\|W(\alpha)\|/\|\alpha\|$.
iii) Majorer la variance de $\sup_{\alpha\in\Theta_s} \|W(\alpha)\|/\|\alpha\|$.

L'objectif de l'exercice suivant est de vérifier que si $\mathcal{T}$ est fini, si N est assez grand, avec forte probabilité, W définit une ϵ-isométrie pour $\mathcal{T}$.

Exercice 6
Montrer l'existence d'une constante κ_1 telle que si $\mathcal{T}$ est fini, pour tout $\epsilon, \delta \in]0, 1[$, si

$$N \geq \frac{\kappa_1}{\epsilon^2} \log \frac{2|\mathcal{T}|}{\delta}$$

alors avec probabilité supérieure à $1 - \delta$, W est une ϵ-isométrie.
Suggestion : n'hésitez pas à utiliser les inégalités de concentration gaussiennes et les bornes de Bonferroni (union bound).

Exercice 7
Dans un espace métrique (E, d), on dit qu'un ensemble $G \subseteq E$ est ϵ-séparé si deux points distincts de G sont à distance supérieure ou égale à ϵ. Si $F \subseteq E$, on note $H(\epsilon, F)$ la cardinalité maximale d'un sous-ensemble ϵ-séparé de F.
Montrer que si B^d désigne la boule unité dans $\mathbb{R}^d$

$$H(\epsilon, B^d) \leq d \log\left(1 + \frac{2}{\epsilon}\right) .$$

Suggestion : utiliser un argument de volume.

Pour la suite, on admettra que pour tout $\epsilon < 1/2$, il existe un ensemble $\mathcal{T}_\epsilon$ de points de A^p_{2s}, tel que

$$\log |\mathcal{T}_\epsilon| \leq 2s \log \left(\kappa_2 \frac{p}{2s\epsilon} \right)$$

tel que tout

$$\forall \alpha \in A^{2s}_p, \exists \beta \in \mathcal{T}_\epsilon, \|\alpha - \beta\| \leq \epsilon \text{ et } (\alpha - \beta) \in \Theta_{2s}\,.$$

On notera $\Pi_\epsilon(\alpha)$ un voisin b de $\alpha \in A^p_{2s}$ dans $\mathcal{T}_\epsilon$ tel que $\|\alpha - \beta\| \leq \epsilon$ et $(\alpha - \beta) \in \Theta_s$.

On note

$$U := \sup \left\{ \|W(\alpha)\|^2 - 1 : \alpha \in A^p_{2s} \right\} \qquad V := \sup \left\{ 1 - \|W(\alpha)\|^2 : \alpha \in A^p_{2s} \right\}$$

On vérifie immédiatement que si U et V sont inférieurs à ϵ alors la matrice de détection définie par les $(X_i)_{i \leq N}$ est une ϵ-isométrie pour Θ_{2s}.
On note $\mathcal{D}_\epsilon \subseteq \Theta_{2s}$, l'ensemble défini par

$$\mathcal{D}_\epsilon := \left\{ x - \Pi_\epsilon(x) : x \in A^p_{2s} \right\}\,.$$

Exercice 8
Dans cet exercice $\epsilon \in]0, 1/2[$. Montrer que pour tout $\eta > 0$

$$U \leq (1 + \eta) \left(\sup_{\alpha \in \mathcal{T}_\epsilon} \|W(\alpha)\|^2 - 1 \right) + \left(1 + \frac{1}{\eta} \right) \epsilon^2 U + \eta\,.$$

Suggestion : ne pas oublier $(x + y)^2 \leq (1 + \beta)x^2 + (1 + 1/\beta)y^2$ pour tout $\beta > 0$.
Montrer que si $\sup_{\alpha \in \mathcal{T}_\epsilon} \|W(\alpha)\|^2 - 1 \leq \epsilon$, en choisissant bien η, on peut garantir

$$U \leq 16\epsilon\,.$$

Montrer de façon analogue que si $\sup_{\alpha \in \mathcal{T}_\epsilon} 1 - \|W(\alpha)\|^2 \leq \epsilon$, alors

$$V \leq 16\epsilon\,.$$

Exercice 9
En combinant les résultats précédents, montrer que si la solution ϵ_* de l'équation

$$N\epsilon^2 = 16^2 \kappa_1 \left(2s \log \left(\kappa_2 \frac{16p}{2s\epsilon} \right) + \log \frac{2}{\delta} \right)$$

est inférieure à $1/2$, alors avec une probabilité supérieure à $1 - \delta$, W définit une ϵ_*-isométrie sur Θ_s.

MORALITÉ : LE HASARD FAIT BIEN LES CHOSES.

Exercice 10
Soit $X_1, \ldots, X_n$ un échantillon i.i.d. de $\mathcal{N}(\mu_x, \sigma_x^2)$ et $Y_1, \ldots, Y_m$ un échantillon i.i.d. de $\mathcal{N}(\mu_y, \sigma_y^2)$. Les espérances et variances sont inconnues. On veut tester

1. $H_0 : \mu_x = \mu_y$ contre
2. $H_1 : \mu_x \neq \mu_y$

sous le postulat $\sigma_x^2 = \sigma_y^2$ (mais σ_x^2 inconnu.)

i) Ecrire la log-vraisemblance d'une paire d'échantillons sous H_0 et sous H_1.
ii) Calculer l'estimateur au maximum de vraisemblance de $\mu_y - \mu_x$. Loi de cet estimateur.
iii) Proposer un intervalle de niveau de confiance $1 - \alpha$ pour $\mu_y - \mu_x$.
iv) Proposer un test de rapport de vraisemblance de niveau α pour H_1 contre H_0.

Exercice 11
Dans cet exercice, $\begin{pmatrix} X \\ Y \end{pmatrix}$ est un vecteur Gaussien d'espérance $\mu = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}$ et de covariance connue $C = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. L'objectif est d'estimer θ_1. On dispose de n tirages indépendants distribués selon $\mathcal{N}(\mu, C)$: $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \ldots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$.
Dans le scenario 1, θ_2 est connu du statisticien, dans le scenario 2, θ_2 n'est pas connu (c'est un *paramètre de nuisance* par opposition à θ_1 qui est un *paramètre d'intérêt*).

1. Quelle est la loi conditionnelle de X sachant Y ? Quels sont les paramètres de cette loi qui sont connus dans les scénarios 1 et 2 ?
2. Proposer un estimateur de θ_1 au maximum de vraisemblance dans le scenario 1. Est il biaisé ? Quel est son risque quadratique ?
3. Proposer un estimateur de θ_1 au maximum de vraisemblance dans le scenario 2. Est il biaisé ? Quel est son risque quadratique ?
4. Calculer l'information de Fisher dans les deux scenarios. Les deux estimateurs réalisent-ils la borne de Cramer-Rao ?

Exercice 12
Dans cet exercice, on considère un modèle exponentiel univarié où

$$p_\theta(x) = \exp\left(\theta T(x) - \eta(\theta)\right)$$

est la densité de P_θ par rapport une mesure de probabilité ν sur $\mathbb{R}$ (T est une fonction monotone, donc mesurable de $\mathbb{R}$ dans $\mathbb{R}$). On définit Θ comme l'intervalle d'intérieur non vide sur lequel $\int_{\mathbb{R}} \exp(\theta T(x))\nu(\mathrm{d}x) < \infty$. Dans la suite θ_0 appartient à l'intérieur de Θ.
Dans la suite $X_1, \ldots, X_n$, sont i.i.d. selon P_θ, $\theta \in \Theta$ inconnu.
On veut développer un test pour distinguer $H_0 : \theta \leq \theta_0$ et $H_1 : \theta > \theta_0$.

1. Proposer une statistique de test S pour distinguer P_{θ_0} de $P_{\theta'}$ ($\theta' > \theta_0$). Préciser la forme de la région critique choisie pour obtenir un niveau donné.
2. Si maintenant vous souhaitez utiliser la statistique S pour distinguer H_0 de H_1, comment choisir une région critique pour obtenir un niveau donné ?
3. Si $\theta_1 > \theta'$, pouvez-vous concevoir un test pour distinguer P_{θ_0} de P_{θ_1} de niveau α, mais plus puissant en θ_1 que le test générique développé pour tester H_1 contre H_0 ? Justifier.