Année universitaire 2018/2019

Site : ☐ Luminy ☒ St-Charles ☐ St-Jérôme ☐ Cht-Gombert ☐ Aix-Montperrin ☐ Aubagne-SATIS
Sujet de : ☒ 1er semestre ☐ 2ème semestre ☐ Session 2 Durée de l'épreuve : 3h
Examen de : M2 Nom du diplôme : Master de Mathématiques appliquées statistiques
Code du module : S53MA3C1B Libellé du module : Représentations parcimonieuses des signaux et images
Calculatrices autorisées : NON Documents autorisés : NON

Aucun document n'est autorisé.
Il est conseillé de lire en entier l'énoncé avant de commencer.
Une réponse non justifiée ne rapportera aucun point.

Notations :

— On note $\langle .,. \rangle$ le produit scalaire usuel sur $\mathbb{C}^M$ tel que $\langle x, y \rangle = \sum_{n=0}^{M-1} x_n \overline{y_n}$ pour $x = (x_i)_{i=0,\dots M-1} \in \mathbb{C}^M$ et $y = (y_i)_{i=0,\dots M-1} \in \mathbb{C}^M$. On note $|x|_2^2 = \sum_{n=0}^{M-1} |x_n|^2$ pour $x \in \mathbb{C}^M$.
De même $|x|_1 = \sum_{n=0}^{M-1} |x_i|$.
Attention les normes $|.|_2$ et $|.|_1$ seront définies avec la même notation quel que soit l'espace du type $\mathbb{C}^M$ dans lequel on travaille.

— On note δ^ℓ pour $\ell \in \{0, \dots M-1\}$ le ℓ-ième vecteur de la base canonique de $\mathbb{C}^M$ ou $\mathbb{R}^M$. Il s'agit du vecteur tel que pour $n \in \{0, \dots M-1\}$ on a $\delta_n^\ell = 1$ si $n = \ell$ et 0 sinon. La collection de vecteurs $\{\delta^\ell, \ell \in \{0, \dots, M-1\}\}$ constitue la base canonique de $\mathbb{C}^M$ et aussi de $\mathbb{R}^M$.

— la famille de vecteurs $\widetilde{\mathcal{E}} = \{\tilde{e}^\ell \in \mathbb{C}^M, \ell = 0, \dots, M-1\}$ tels que pour $\ell \in \{0, \dots, M-1\}$ et pour $n \in \{0, \dots M-1\}$ la n-ième coordonnée du vecteur $\tilde{e}^\ell$ s'écrive

$$\tilde{e}_n^\ell = \frac{e^{\frac{2i\pi \ell n}{M}}}{\sqrt{M}}$$

est appelée base de Fourier orthonormalisée discrète.

1 Bases et dictionnaires

Exercice 1 (*environ 6 pts*)

1. *Soit $\{b^0, \dots, b^{N-1}\}$ une base orthonormée de $\mathbb{C}^N$.*
   (a) *Montrer que pour tout $y \in \mathbb{C}^N$ on a*

$$y = \sum_{k=0}^{N-1} \langle y, b^k \rangle b^k \tag{1}$$

   (b) *Montrer que pour tout $y \in \mathbb{C}^N$ on a*

$$|y|_2^2 = \sum_{k=0}^{N-1} |\langle y, b^k \rangle|^2 \tag{2}$$

2. *Soit $\mathcal{B}_1 = \{b^0, \dots, b^{N-1}\}$ et $\mathcal{B}_2 = \{\psi^0, \dots, \psi^{N-1}\}$ deux bases orthonormées de $\mathbb{C}^N$.*

(a) *Montrer que pour tout* $\theta \in [0,1]$ *on a*

$$y = \sum_{k=0}^{N-1} \theta \langle y, b^k \rangle b^k + \sum_{k=0}^{N-1} (1-\theta) \langle y, \psi^k \rangle \psi^k$$

On posera dans la suite $\theta = \frac{1}{2}$ *et on considère donc*

$$y = \sum_{k=0}^{N-1} \frac{\langle y, b^k \rangle}{2} b^k + \sum_{k=0}^{N-1} \frac{\langle y, \psi^k \rangle}{2} \psi^k \tag{3}$$

(b) *On prend* $\mathcal{B}_1 = \{\delta^0, \dots, \delta^{N-1}\}$ *la base canonique et* $\mathcal{B}_2 = \{\tilde{e}^0, \dots, \tilde{e}^{N-1}\}$ *la base de Fourier orthonormalisée discrète. On note* $\mathcal{D} = \mathcal{B}_1 \bigcup \mathcal{B}_2$.

i. *Montrer que* $\mathcal{D}$ *est un dictionnaire de* $\mathbb{C}^N$. *Est-ce une base de* $\mathbb{C}^N$ *?*

ii. *Soit* $k_0 \in \{0, \dots, N-1\}$ *et* $v = \delta^{k_0}$. *Le vecteur* v *est-il parcimonieux dans la base canonique ? Dans la base de Fourier orthonormalisée discrète ? La décomposition (3) fournit-elle une décomposition parcimonieuse de* v *dans* $\mathcal{D}$ *?*

iii. *Donner la décomposition la plus parcimonieuse possible de* v *sur les vecteurs de* $\mathcal{D}$.

(c) *Donner un exemple de vecteur* w *non parcimonieux ni dans la base canonique ni dans la base de Fourier mais qui admet une décomposition parcimonieuse dans le dictionnaire* $\mathcal{D}$.

2 Problème : Étude de l'algorithme de Frank-Wolfe

On considère $x \in \mathbb{R}^N$ et un dictionnaire $\mathcal{D} = \{\phi^0, \dots, \phi^{K-1}\}$ de $\mathbb{R}^N$. Dans tout ce qui suit on suppose que x admet une décomposition dans $\mathcal{D}$ telle que

$$x = \sum_{k=0}^{K-1} \beta_k \phi^k \text{ avec } |\beta|_1 \leq 1 \tag{4}$$

On note D la matrice qui a pour vecteurs colonnes les vecteurs ϕ^k exprimés dans la base canonique et D^* la transposée de D.

On considère $F : \alpha \mapsto |D\alpha - x|_2^2$.

L'algorithme de Frank-Wolfe (ou algorithme de « gradient conditionnel »), qui est souvent utilisé en apprentissage statistique, propose de calculer une solution au problème

$$\begin{aligned} &\text{Trouver } \tilde{\alpha} \in \mathbb{R}^K \text{ tel que} \\ F(\tilde{\alpha}) &= \min\{F(\alpha) : \alpha \in \mathbb{R}^K \text{ et } |\alpha|_1 \leq 1\} \end{aligned} \tag{5}$$

2.1 Préliminaire mathématique

Exercice 2 (*environ 4 pts*)

1. *Donner un exemple d'un vecteur* x *qui vérifie (4).*
2. *Montrer que si* x *vérifie (4) alors* $F(\beta) = 0$ *et* β *est une solution du problème (5).*
3. *Soit* $u \in \mathbb{R}^K$ *et* $h \in \mathbb{R}^K$. *Montrer que*

$$F(u) - F(u-h) = -2\langle x - Du, Dh \rangle + g(h) \tag{6}$$

où $g(h) \leq 0$ *pour tout* $h \in \mathbb{R}^K$.

En déduire que pour tout $\alpha \in \mathbb{R}^K$

$$F(\alpha) \leq 2\langle x - D\alpha, D(-\alpha + \beta)\rangle \tag{7}$$

2.2 Algorithme de Frank-Wolfe

L'algorithme de Frank-Wolfe est un algorithme itératif qui s'appuie sur l'équation (6). Il fonctionne de la façon suivante.

1. Initialisation : $m := 0$, on prend α^0 qui est le vecteur nul.
2. Par récurrence sur $m \in \mathbb{N}^*$
   (a) Calcul des produits scalaires pour tout k $c_k = \langle x - D\alpha^{m-1}, D\delta^k\rangle$
   (b) choix du meilleur vecteur de la base canonique $k_m = \arg\max_k |\langle x - D\alpha^{m-1}, D\delta^k\rangle|$
   (c) calcul du pas $\gamma_m = \frac{2}{m+1}$
   (d) mise à jour du signe $s_m = \text{sign}(c_{k_m}) \in \{-1, 1\}$
   (e) mise à jour de l'itérée $\alpha^m = (1 - \gamma_m)\alpha^{m-1} + \gamma_m s_m \delta^{k_m} = \alpha^{m-1} + \gamma_m \left(s_m \delta^{k_m} - \alpha^{m-1}\right)$

La dernière partie du sujet permet de montrer que $F(\alpha^m) \to F(\beta) = 0$ pour $m \to +\infty$.

Exercice 3 (*environ 5 pts*)

1. *Dans l'algorithme ci-dessus quelles sont les dimensions des vecteurs $(c_k)_k$, α^m et δ^k ? Combien vaut $D\delta^k$?*
2. *Indiquer comment calculer $c = (c_k)_k$ défini ligne 2a, sans boucle* `for`*, à partir de $x - D\alpha^{m-1}$ et D^*.*
3. *Indiquer comment calculer k_m défini ligne 2b à l'aide de Python et comment calculer α^m défini ligne 2e.*
4. *Quels tests d'arrêt de l'algorithme proposez-vous de mettre en place ?*

2.3 Convergence de l'algorithme

Exercice 4 (*environ 8 pts*)

1. *En utilisant si besoin l'inégalité pour a et b deux réels positifs $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ montrer que pour tout vecteur $\alpha \in \mathbb{R}^M$ on a $|\alpha|_2 \leq |\alpha|_1$.*
2. *Montrer qu'avec α^0 le vecteur nul on a pour tout $m \in \mathbb{N}$ $|\alpha^m|_1 \leq 1$ avec α^m défini à la ligne 2e dans la description de l'algorithme de Frank-Wolfe.*
3. *Soit $h \in \mathbb{R}^K$ tel que $|h|_1 \leq 1$. Soit $m \in \mathbb{N}$. Montrer que*

$$\langle x - D\alpha^{m-1}, Dh\rangle \leq s_m \langle x - D\alpha^{m-1}, D\delta^{k_m}\rangle \tag{8}$$

4. *En déduire à l'aide de (7) que pour tout $m \in \mathbb{N}^*$*

$$F(\alpha^{m-1}) \leq 2\langle x - D\alpha^{m-1}, D(s_m\delta^{k_m} - \alpha^{m-1})\rangle$$

5. *Montrer que $F(\alpha^m) \leq F(\alpha^{m-1}) - \gamma_m 2\langle x - D\alpha^{m-1}, D(s_m\delta^{k_m} - \alpha^{m-1})\rangle + C\gamma_m^2$ où C est une constante qui dépend de $\|D^*D\|$ et que l'on précisera.*
6. *Montrer par récurrence que $F(\alpha^m) \leq \frac{4C}{m+3}$ pour $m \geq 1$. Conclure.*

Voilà c'est fini !

Une dernière citation, attribuée à P. Hein avant de partir en vacances et de vous souhaiter de très bonnes fêtes de fin d'année

Problems worthy of attack prove their worth by fighting back !

ce qui est traduit par

Un problème digne d'attaque montre sa valeur en ripostant !

Et donc très bonnes fêtes de fin d'année ! !