# OPT202 – Optimisation Différentiable II

## Contrôle des connaissances

- Durée : 3 h (on ne peut probablement pas tout faire...).
- Dans chaque sujet (ou sous-sujet), on peut admettre des résultats énoncés dans des questions ou sous-questions *précédentes* pour traiter la question ou sous-question courante.
- On peut consulter tous les documents *distribués au cours*, ainsi que ses propres notes.
- Ne pas écrire en rouge, couleur réservée aux observations des correcteurs.
- *Éteindre tout appareil connecté à un réseau.*

## Notations générales

- L'ensemble $[1:n]$ est celui des $n$ premiers entiers non nuls, soit $[1,n] \cap \mathbb{N}$.
- Les inégalités vectorielles doivent se comprendre composante par composante : si $u$ et $v \in \mathbb{R}^n$, l'inégalité $u \leqslant v$ (resp. $u < v$) signifie que, pour tout $i \in [1:n]$, $u_i \leqslant v_i$ (resp. $u_i < v_i$).
- $\mathbb{R}_+ := \{t \in \mathbb{R} : t \geqslant 0\}$ et $\mathbb{R}^n_+ := \{v \in \mathbb{R}^n : v \geqslant 0\}$,
- On note $\overline{\text{Conv}}(\mathbb{E})$ l'ensemble des fonctions convexes fermées propres définies sur un espace vectoriel $\mathbb{E}$ à valeurs dans $\overline{\mathbb{R}}$.

## Partie 1 : Quatre sujets indépendants

**1. Direction de descente en optimisation convexe**. Soient $\mathbb{E}$ un espace euclidien (produit scalaire $\langle \cdot, \cdot \rangle$ et norme associée $\| \cdot \|$), $f : \mathbb{E} \to \mathbb{R}$ une fonction convexe, $f'(x; d)$ sa dérivée directionnelle en $x \in \mathbb{E}$ dans la direction $d \in \mathbb{E}$ et $\partial f(x)$ son sous-différentiel en $x \in \mathbb{E}$.

Montrez que, si $p$ est le projeté orthogonal de 0 sur $\partial f(x)$, alors $f'(x; -p) \leqslant -\|p\|^2$.

**2. Minoration de la valeur optimale par pénalisation extérieure**. Soient $X$ une partie non vide d'un espace vectoriel $\mathbb{E}$ et $f : \mathbb{E} \to \mathbb{R}$ une fonction. On considère le problème d'optimisation

$$(P_X) \qquad \inf_{x \in X} f(x)$$

et le problème de pénalisation associé

$$(P_r) \qquad \inf_{x \in \mathbb{E}} f(x) + r\, p(x),$$

où $r \in \mathbb{R}$ et $p : \mathbb{E} \to \mathbb{R}$ est une fonction vérifiant $p(\cdot) \geqslant 0$ et $p(X) = \{0\}$. On note $\text{val}(P_X)$ et $\text{val}(P_r)$ les valeurs optimales de ces problèmes.

*Sans supposer que les problèmes $(P_X)$ et $(P_r)$ ont une solution*, montrez que

1) $\text{val}(P_r)$ croît lorsque $r$ augmente,
2) pour tout $r \in \mathbb{R}$, $\text{val}(P_r) \leqslant \text{val}(P_X)$.

**3. Conditions d'optimalité par pénalisation quadratique, avec qualification de Mangasarian-Fromovitz**. On considère le problème d'optimisation

$$(P_I) \quad \begin{cases} \min f(x) \\ c(x) \leqslant 0, \end{cases}$$

dans lequel $f : \mathbb{R}^n \to \mathbb{R}$ et $c : \mathbb{R}^n \to \mathbb{R}^m$ sont continûment différentiables. Le problème de pénalisation quadratique associé à $(P_I)$ consiste à minimiser sur $\mathbb{R}^n$ la fonction $\Theta_r : \mathbb{R}^n \to \mathbb{R}$, définie en $x \in \mathbb{R}^n$ par

$$\Theta_r(x) := f(x) + \frac{r}{2}\|c(x)^+\|^2,$$

où $r \geqslant 0$, où, pour $v \in \mathbb{R}^m$, $v^+$ est le vecteur de $\mathbb{R}^m$ dont la $i$-ième composante vaut $\max(0, v_i)$, et où $\|\cdot\|$ désigne la norme $\ell_2$ (euclidienne). On suppose que

$(i)$ pour une suite de $r \to \infty$, la fonction $\Theta_r$ a un point stationnaire noté $\bar{x}_r$, vérifiant donc $\nabla\Theta_r(\bar{x}_r) = 0$,

$(ii)$ $\bar{x}_r \to \bar{x}$ lorsque $r \to +\infty$,

$(iii)$ $\bar{x}$ est un point admissible de $(P_I)$ vérifiant les conditions de qualification (QC-MF) : pour $I^0(\bar{x}) := \{i \in [1:m] : c_i(\bar{x}) = 0\}$, tout vecteur $\alpha \in \mathbb{R}_+^{|I^0(\bar{x})|}$ vérifiant

$$\sum_{i \in I^0(\bar{x})} \alpha_i \nabla c_i(\bar{x}) = 0$$

s'annule.

Montrez que

1) en $x \in \mathbb{R}^n$ :

$$\Theta_r(x) = \inf_{s \in \mathbb{R}_+^m} f(x) + \frac{r}{2}\|c(x) + s\|^2,$$

2) il existe une suite $\{\lambda_r\} \subset \mathbb{R}_+^m$ telle que $\nabla f(\bar{x}_r) + c'(\bar{x}_r)^\mathsf{T}\lambda_r = 0$,

3) cette suite $\{\lambda_r\}$ est bornée,

[*Indication* : on pourra raisonner par l'absurde, en montrant que, dans le cas contraire, il existe une suite de $\lambda_r$ telle que $\|\lambda_r\| \to \infty$, $\lambda_r/\|\lambda_r\| \to \mu$ et en déduisant une contradiction.]

4) il existe un $\bar{\lambda} \in \mathbb{R}^m$ tel que

$$\begin{cases} \nabla f(\bar{x}) + c'(\bar{x})^\mathsf{T}\bar{\lambda} = 0 \\ 0 \leqslant \bar{\lambda} \perp c(\bar{x}) \leqslant 0. \end{cases}$$

**4. Circonscription ellipsoïdique d'un polyèdre convexe**. Pour une matrice $M \succ 0$ (symétrique définie positive) d'ordre $n$, on note

$$\mathscr{E}(M) := \{x \in \mathbb{R}^n : x^\mathsf{T} M x \leqslant 1\}$$

l'ellipsoïde de $\mathbb{R}^n$ façonné par $M$. Soient $H \succ 0$ et $P$ le polyèdre convexe de $\mathbb{R}^n$ défini par

$$P := \{x \in \mathbb{R}^n : Ax \leqslant b\},$$

où $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$. On suppose que $P \neq \varnothing$. On cherche à montrer l'équivalence suivante

$$P \subset \mathscr{E}(H) \qquad \Longleftrightarrow \qquad \mathscr{E}(H^{-1}) \subset A^\mathsf{T}\{y \in \mathbb{R}_+^m : b^\mathsf{T} y \leqslant 1\}. \tag{1}$$

Supposons, dans un premier temps, que $P \subset \mathscr{E}(H)$.

1) Montrez que, pour tout $v \in \mathscr{E}(H^{-1})$, on a $\sup\{v^\mathsf{T} x : x \in P\} \leqslant 1$.

2) Montrez que $\mathscr{E}(H^{-1}) \subset A^\mathsf{T}\{y \in \mathbb{R}_+^m : b^\mathsf{T} y \leqslant 1\}$.

[*Indication* : on pourra utiliser le problème d'optimisation du point 1, après justification.]

Supposons maintenant que $\mathscr{E}(H^{-1}) \subset A^\mathsf{T}\{y \in \mathbb{R}_+^m : b^\mathsf{T} y \leqslant 1\}$.

3) Montrez que $P \subset \mathscr{E}(H)$.

## Partie 2 : Conjuguée et sous-différentiel d'une fonction spectrale

Pour $x \in \mathbb{R}^n$, on note $[x] \in \mathbb{R}^n$ le vecteur formé des composantes de $x$ en ordre décroissant : il existe une permutation $p : [1:n] \to [1:n]$ (c'est-à-dire une application bijective) telle que $[x]_i = x_{p(i)}$ pour tout $i \in [1:n]$ et

$$[x]_1 \geqslant [x]_2 \geqslant \cdots \geqslant [x]_n.$$

On dit qu'une fonction $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ est *symétrique* si

$$\forall\, x \in \mathbb{R}^n : \qquad f(x) = f([x]).$$

Cela revient à dire que l'on ne modifie pas la valeur de $f(x)$ en permutant les composantes de $x$.

On note $\mathcal{S}^n$ l'ensemble des matrices d'ordre $n$ symétriques, muni du produit scalaire $\langle A, B\rangle = \operatorname{tr} AB$ (la trace de $AB$), dont la norme associée est la norme de Frobenius $A \mapsto \|A\|_F = \langle A, A\rangle^{1/2} = (\operatorname{tr} A^2)^{1/2}$.

Soient $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ une fonction *symétrique* et

$$\lambda : \mathcal{S}^n \to \mathbb{R}^n : A \mapsto \lambda(A) := (\lambda_1(A), \dots, \lambda_n(A)),$$

la fonction donnant les valeurs propres de $A$ en ordre *décroissant* (on rappelle qu'une matrice $A \in \mathcal{S}^n$ a $n$ valeurs propres réelles) :

$$\lambda_1(A) \geqslant \lambda_2(A) \geqslant \cdots \geqslant \lambda_n(A).$$

On appelle *fonction spectrale* une fonction de la forme $f \circ \lambda$, avec $f$ et $\lambda$ comme ci-dessus. Ce sont donc des fonctions définies sur $\mathcal{S}^n$, mais dont les valeurs ne dépendent que du spectre des matrices.

On admettra l'*inégalité de Fan* suivante :

$$\forall\, (A, B) \in \mathcal{S}^n \times \mathcal{S}^n : \qquad \langle A, B\rangle \leqslant \lambda(A)^\mathsf{T} \lambda(B), \tag{2}$$

*avec égalité si, et seulement si, il existe une matrice orthogonale $V$ telle que $V^\mathsf{T} A V = \operatorname{Diag}(\lambda(A))$ et $V^\mathsf{T} B V = \operatorname{Diag}(\lambda(B))$* (rappel : une matrice $V$ est orthogonale si $V^{-1} = V^\mathsf{T}$).

Ci-dessous, $\mathbb{R}^n$ est muni du produit scalaire euclidien. On notera $\varphi^*$ la fonction conjuguée d'une fonction $\varphi$ et $\partial\varphi(x)$ le sous-différentiel d'une fonction convexe $\varphi$.

**5.** Montrez que $f^*$ est symétrique.

**6.** Montrez que $(f \circ \lambda)^* = f^* \circ \lambda$.

[*Indication* : pour l'inégalité $\leqslant$ on pourra utiliser l'inégalité de Fan (2).]

**7.** Montrez que $(f \circ \lambda)^{**} = f^{**} \circ \lambda$.

**8.** Montrez que $(f \circ \lambda) \in \overline{\operatorname{Conv}}(\mathcal{S}^n)$ si, et seulement si, $f \in \overline{\operatorname{Conv}}(\mathbb{R}^n)$.

**9.** Montrez que

$$\forall\, x, y \in \mathbb{R}^n : \qquad x^\mathsf{T} y \leqslant [x]^\mathsf{T} [y].$$

On suppose dorénavant qu'en plus d'être symétrique, $f \in \overline{\mathrm{Conv}}(\mathbb{R}^n)$.

**10.** Montrez que
$$s \in \partial f(x) \qquad \Longrightarrow \qquad [s] \in \partial f([x]).$$

**11.** Montrez que les trois propriétés suivantes, reliant $A$ et $S \in \mathcal{S}^n$, sont équivalentes :

$(i)$ $S \in \partial(f \circ \lambda)(A)$,
$(ii)$ $\langle S, A\rangle = \lambda(S)^\mathsf{T}\lambda(A)$ et $\lambda(S) \in \partial f(\lambda(A))$,
$(iii)$ il existe une matrice orthogonale $V$ et des vecteurs $a$, $s \in \mathbb{R}^n$ tels que
$$V^\mathsf{T} SV = \mathrm{Diag}(s), \qquad V^\mathsf{T} AV = \mathrm{Diag}(a) \qquad \text{et} \qquad s \in \partial f(a).$$

[*Indication* : le plus simple pour nous a été de démontrer successivement les équivalences $(i)$ $\Leftrightarrow$ $(ii)$ et $(ii)$ $\Leftrightarrow$ $(iii)$, en utilisant ce qui a été démontré précédemment.]

**12.** Montrez que $(f \circ \lambda)$ est Gâteaux-différentiable en $A \in \mathcal{S}^n$ si, et seulement si, $f$ est Gâteaux-différentiable en $\lambda(A) \in \mathbb{R}^n$.

[*Indication* : c'est long et gratifiant...]

# OPT202 – Optimisation Différentiable II

## Contrôle des connaissances (solutions succinctes)

**Barème** : 53 points, qui seront ramenés à 20 (on peut évaluer les réponses avec des points non entiers)

## Partie 1 : Quatre sujets indépendants

**1. Barème** : 2 points

Par définition de la projection, on a pour tout $g \in \partial f(x)$ : $\langle g - p, p - 0\rangle \geqslant 0$ ou $\langle g, -p\rangle \leqslant -\|p\|^2$. Par la formule du max, on en déduit que $f'(x; -p) \leqslant -\|p\|^2$.

Remarque. On a donc montré que, si $0 \notin \partial f(x)$, $-p$ est une direction de descente de $f$ en $x$ (alors que $-g$ n'est pas nécessairement une direction de descente pour un sous-gradient $g$ arbitraire). En réalité, on a l'égalité

$$f'(x; -p) = -\|p\|^2,$$

car on a $f'(x; -p) \geqslant \langle g, -p\rangle$ pour tout sous-gradient $g \in \partial f(x)$, donc aussi pour $g = p \in \partial f(x)$, ce qui donne l'inégalité inverse $f'(x; -p) \geqslant -\|p\|^2$.

**2. Barème** : 2 points

1) Soient $r_1 < r_2$ dans $\mathbb{R}$. Pour tout $x \in \mathbb{E}$, on a

$$\operatorname{val}(P_{r_1}) \leqslant f(x) + r_1 p(x) \leqslant f(x) + r_2 p(x), \tag{3}$$

parce que $r_1 < r_2$ et $p(\cdot) \geqslant 0$. En prenant l'infimum en $x \in \mathbb{E}$ à droite, on obtient $\operatorname{val}(P_{r_1}) \leqslant \operatorname{val}(P_{r_2})$, si bien que $\{\operatorname{val}(P_r)\}$ croît avec $r$.

2) Si, à droite dans (3), on prend l'infimum en $x \in X$, on obtient $\operatorname{val}(P_{r_1}) \leqslant \operatorname{val}(P)$, parce que $p(X) = 0$.

**3. Barème** : 8 points

1) **Barème** : 1 point

Il suffit de montrer que $\min\{\|c(x) + s\|^2 : s \in \mathbb{R}^m_+\} = \|c(x)^+\|^2$, ce qui se déduit du fait que le problème à gauche de l'égalité consiste à projeter $-c(x)$ sur l'orthant positif $\mathbb{R}^m_+$. Comme le projeté orthogonal est $(-c(x))^+ = c(x)^-$ (exercice X.2), la valeur optimale du problème vaut $\|c(x) + c(x)^-\|^2 = \|c(x)^+\|^2$.

Remarque. La fonction $\Theta_r$ s'obtient donc en remplaçant la contrainte $c(x) \leqslant 0$ de $(P_I)$ par les contraintes équivalentes ($c(x) + s = 0$ et $s \geqslant 0$) et en prenant une pénalisation quadratique des contraintes d'égalité tout en gardant explicite les contraintes de positivité $s \geqslant 0$.

2) **Barème** : 1 point

La relation $\nabla\Theta_r(\bar{x}_r) = 0$ s'écrit (exercice X.6)

$$\nabla f(\bar{x}_r) + r\, c'(\bar{x}_r)^{\mathsf{T}} c(\bar{x}_r)^+ = 0.$$

On a donc l'identité

$$\nabla f(\bar{x}_r) + c'(\bar{x}_r)^{\mathsf{T}} \lambda_r = 0, \tag{4}$$

avec $\lambda_r := r c(\bar{x}_r)^+$ qui est bien positif.

3) **Barème** : 3 points (1 point pour la justification des propriétés de la sous-suite $\{\lambda_r\}$

Supposons que $\{\lambda_r\}$ ne soit pas bornée. Alors il existe une sous-suite, encore notée $\{\lambda_r\}$, telle que $\|\lambda_r\| \to \infty$ et $\lambda_r/\|\lambda_r\| \to \mu$ (la suite $\{\lambda_r/\|\lambda_r\|\}$ étant bornée, on peut en extraire une sous-suite convergente), pour des $r \to \infty$. En divisant les deux membres de (4) par $\|\lambda_r\|$ et en passant à la limite, on trouve

$$c'(\bar{x})^\mathsf{T}\mu = 0 \qquad \text{et} \qquad \mu \geqslant 0.$$

Par la définition de $\lambda_r$, $\mu_i = 0$ si $i \in [1:m] \smallsetminus I^0(\bar{x})$. Alors, par (QC-MF), on en déduit que $\mu = 0$, ce qui est en contradiction avec la définition de $\mu$ qui doit être de norme 1.

4) **Barème** : 3 points (2/3 point pour les 3 premiers items, 1 point pour le dernier)

- Comme $\{\lambda_r\}$ est bornée, on peut en extraire une sous suite convergente, encore notée $\{\lambda_r\} \to \bar{\lambda}$.
- En passant à la limite dans (4), on trouve que $\nabla f(\bar{x}_r) + c'(\bar{x}_r)^\mathsf{T}\bar{\lambda} = 0$.
- De la positivité de $\lambda_r$, on déduit celle de $\bar{\lambda}$.
- Enfin, si $c_i(\bar{x}) < 0$, alors $c_i(\bar{x}_r) < 0$ pour $r$ grand; et donc $(\lambda_r)_i = rc_i(\bar{x}_r)^+ = 0$ pour $r$ grand, si bien que $\bar{\lambda}_i = 0$. La complémentarité a donc bien lieu : $\bar{\lambda}^\mathsf{T}c(\bar{x}) = 0$.

**4. Barème** : 8 points

1) **Barème** : 2 points

Soit $v \in \mathscr{E}(H^{-1})$. Pour $x \in \mathscr{E}(H)$, on a

$$v^\mathsf{T}x = (H^{-1/2}v)^\mathsf{T}(H^{1/2}x) \leqslant \|H^{-1/2}v\|\,\|H^{1/2}x\| = (v^\mathsf{T}H^{-1}v)^{1/2}(x^\mathsf{T}Hx)^{1/2} \leqslant 1,$$

où l'on a utilisé l'inégalité de Cauchy-Schwarz. Comme $P \subset \mathscr{E}(H)$, on a aussi $v^\mathsf{T}x \leqslant 1$ pour $x \in P$ et donc

$$\sup_{x \in P} v^\mathsf{T}x \leqslant 1. \tag{5}$$

2) **Barème** : 4 points (1 pour le premier argument, 2 pour le second et 1 pour le dernier)

- On observe d'abord que le problème (5) est linéaire et a une solution (car $P \neq \varnothing$ et le problème est borné).
- Par la dualité forte, son dual a la même valeur : pour tout $v \in \mathscr{E}(H^{-1})$, on a

$$\max_{\substack{x\\ Ax \leqslant b}} v^\mathsf{T}x = \max_x \inf_{y \geqslant 0} v^\mathsf{T}x - y^\mathsf{T}(Ax - b) = \min_{y \geqslant 0} \sup_x (v - A^\mathsf{T}y)^\mathsf{T}x + b^\mathsf{T}y = \min_{\substack{A^\mathsf{T}y = v\\ y \geqslant 0}} b^\mathsf{T}y \leqslant 1$$

- On en déduit que

$$\forall\, v \in \mathscr{E}(H^{-1}), \quad \exists\, y \in \mathbb{R}^m_+ : \quad A^\mathsf{T}y = v \text{ et } b^\mathsf{T}y \leqslant 1.$$

  Cela donne l'inclusion désirée.
- Au lieu des deux derniers points, on aurait pu aussi utiliser les conditions d'optimalité du problème linéaire : il existe un $y \in \mathbb{R}^m$ tel que

$$v = A^\mathsf{T}y \qquad \text{et} \qquad 0 \leqslant y \perp (Ax - b) \leqslant 0.$$

  Alors $b^\mathsf{T}y = y^\mathsf{T}Ax = v^\mathsf{T}x \leqslant 1$.

3) **Barème** : 2 points

Soit $x \in P$, vérifiant donc $Ax \leqslant b$. On peut supposer que $x \neq 0$ (car $0 \in \mathscr{E}(H)$, clairement). Observons alors que

$$\frac{Hx}{\|H^{1/2}x\|} \in \mathscr{E}(H^{-1}).$$

Par l'hypothèse, on peut trouver un $y \in \mathbb{R}^m$ tel que

$$\frac{Hx}{\|H^{1/2}x\|} = A^{\mathsf{T}} y, \qquad y \geqslant 0 \qquad \text{et} \qquad b^{\mathsf{T}} y \leqslant 1.$$

En multipliant les deux membres de la première identité par $x$, on voit que

$$\frac{x^{\mathsf{T}} Hx}{\|H^{1/2}x\|} = \underbrace{y}_{\geqslant 0}{}^{\mathsf{T}} \underbrace{Ax}_{\leqslant b} \leqslant b^{\mathsf{T}} y \leqslant 1.$$

Donc $x^{\mathsf{T}} Hx \leqslant 1$, si bien que $x \in \mathscr{E}(H)$.

Remarque. On a donc montré qu'un polyèdre convexe est inscrit dans l'ellipsoïde $\mathscr{E}(H)$ si, et seulement si, l'ellipsoïde $\mathscr{E}(H^{-1})$ est inscrit dans un certain polyèdre convexe construit avec les données du premier polyèdre convexe.

## Partie 2 : Conjuguée et sous-différentiel d'une fonction spectrale

**Barème** : 33 points

**5. Barème** : 1 point

Soient $s \in \mathbb{R}^n$ et $p : [1:n] \to [1:n]$ la permutation telle que $s_p = [s]$ ($s_p$ est le vecteur de $\mathbb{R}^n$ tel que $(s_p)_i = s_{p(i)}$). On a

$$\begin{aligned}
f^*(s) &= \sup_{x \in \mathbb{R}^n} s^{\mathsf{T}} x - f(x) \\
&= \sup_{x \in \mathbb{R}^n} s_p^{\mathsf{T}} x_p - f(x_p) \qquad [\text{car } s^{\mathsf{T}} x = s_p^{\mathsf{T}} x_p \text{ et } f(x) = f([x]) = f(x_p)] \\
&= \sup_{x \in \mathbb{R}^n} s_p^{\mathsf{T}} x - f(x) \qquad [\text{car } x \mapsto x_p \text{ est bijective}] \\
&= f^*(s_p) \\
&= f^*([s]).
\end{aligned}$$

On peut aussi raisonner en partant de $f^*([s])$ et en utilisant l'application réciproque de $p$, notée $q$ (donc $p \circ q = I$) :

$$\begin{aligned}
f^*([s]) &= \sup_{x \in \mathbb{R}^n} s_p^{\mathsf{T}} x - f(x) \\
&= \sup_{x \in \mathbb{R}^n} s^{\mathsf{T}} x_q - f(x_q) \qquad [\text{car } s_p^{\mathsf{T}} x = s^{\mathsf{T}} x_q \text{ et } f(x) = f([x]) = f(x_q)] \\
&= \sup_{x \in \mathbb{R}^n} s^{\mathsf{T}} x - f(x) \qquad [\text{car } x \mapsto x_q \text{ est bijective}] \\
&= f^*(s).
\end{aligned}$$

**6. Barème** : 4 points (2 points pour "$\leqslant$" [dont 1 point si on remarque bien que $\mathcal{R}(\lambda) \subset \mathbb{R}^n$ et pas $\mathcal{R}(\lambda) = \mathbb{R}^n$], 2 points pour "$\geqslant$")

Pour $S \in \mathcal{S}^n$, on a

$$\begin{aligned}
(f \circ \lambda)^*(S) &= \sup_{A \in \mathcal{S}^n} \langle S, A \rangle - f(\lambda(A)) \\
&\leqslant \sup_{A \in \mathcal{S}^n} \lambda(S)^{\mathsf{T}} \lambda(A) - f(\lambda(A)) \qquad [\text{par } (2)] \\
&\leqslant \sup_{x \in \mathbb{R}^n} \lambda(S)^{\mathsf{T}} x - f(x) \qquad [\text{car } \mathcal{R}(\lambda) \subset \mathbb{R}^n] \\
&= f^*(\lambda(S)).
\end{aligned}$$

Pour l'inégalité inverse, on raisonne comme suit. Soit $S = V\,\mathrm{Diag}(\lambda(S))V^\mathsf{T}$ une décomposition spectrale de $S$. Alors

$$\begin{array}{rcll}
(f\circ\lambda)^*(S) & = & \sup\limits_{A\in\mathcal{S}^n} \langle S,A\rangle - f(\lambda(A)) & \\
& \geqslant & \sup\limits_{x\in\mathbb{R}^n} \langle S, V\,\mathrm{Diag}(x)V^\mathsf{T}\rangle - f([x]) & [\text{en prenant } A = V\,\mathrm{Diag}(x)V^\mathsf{T}] \\
& = & \sup\limits_{x\in\mathbb{R}^n} \langle S, V\,\mathrm{Diag}(x)V^\mathsf{T}\rangle - f(x) & [\text{car } f \text{ est symétrique}] \\
& = & \sup\limits_{x\in\mathbb{R}^n} \langle V^\mathsf{T}SV, \mathrm{Diag}(x)\rangle - f(x) & \\
& = & \sup\limits_{x\in\mathbb{R}^n} \langle \mathrm{Diag}(\lambda(S)), \mathrm{Diag}(x)\rangle - f(x) & [\text{définition de } V] \\
& = & \sup\limits_{x\in\mathbb{R}^n} \lambda(S)^\mathsf{T}x - f(x) & \\
& = & f^*(\lambda(S)). &
\end{array}$$

**7. Barème** : 2 points (on peut être tenté par un long cheminement)

Clairement

$$\begin{array}{rcll}
(f\circ\lambda)^{**} & = & (f^*\circ\lambda)^* & [f \text{ symétrique et question } 6] \\
& = & f^{**}\circ\lambda & [f^* \text{ symétrique (par la question } 5) \text{ et question } 6].
\end{array}$$

**8. Barème** : 2 points

On a

$$\begin{array}{rcll}
(f\circ\lambda) \in \overline{\mathrm{Conv}}(\mathcal{S}^n) & \iff & (f\circ\lambda)^{**} = (f\circ\lambda) & [\text{proposition } 3.40] \\
& \iff & f^{**}\circ\lambda = f\circ\lambda & [\text{question } 7] \\
& \iff & f^{**} = f & [\mathcal{R}(\lambda) = \mathbb{R}^n_\geqslant\,;\ f \text{ et } f^{**} \text{ symétriques}] \\
& \iff & f \in \overline{\mathrm{Conv}}(\mathbb{R}^n) & [\text{proposition } 3.40],
\end{array}$$

où on a noté $\mathbb{R}^n_\geqslant := \{x\in\mathbb{R}^n : x_1 \geqslant \cdots \geqslant x_n\}$.

**9. Barème** : 1 point (tant pis si l'étudiant s'embarque dans une démonstration longue)

Il suffit d'appliquer l'inégalité de Fan (2) à $A = \mathrm{Diag}(x)$ et $B = \mathrm{Diag}(y)$.

On peut aussi donner une démonstration directe, mais c'est plus long. Voici une démonstration par récurrence. L'inégalité est clairement vraie pour $n = 1$. Supposons qu'elle le soit pour $n-1 \geqslant 1$ et montrons la pour $n$. En permutant les coordonnées de $x$ et $y$ de la même manière, on peut supposer que $[x] = x$. On a par récurrence

$$x^\mathsf{T}y \leqslant [x]^\mathsf{T}\tilde{y},$$

où $\tilde{y} := (y_1, [(y_2,\dots,y_n)])$. Si $k\in[1\!:\!n]$ est la position de $y_1$ dans $[y]$, on a

$$[y] = (\tilde{y}_2,\dots,\tilde{y}_k, y_1, \tilde{y}_{k+1},\dots,\tilde{y}_n)$$

et

$$\begin{array}{rcl}
[x]^\mathsf{T}\tilde{y} & = & [x]^\mathsf{T}[y] - \sum\limits_{i=1}^{k-1} x_i\tilde{y}_{i+1} - x_k y_1 + x_1 y_1 + \sum\limits_{i=2}^{k} x_i\tilde{y}_i \\
& = & [x]^\mathsf{T}[y] + (x_1 - x_k)y_1 + \sum\limits_{i=2}^{k} (x_i - x_{i-1})\tilde{y}_i \\
& = & [x]^\mathsf{T}[y] - \sum\limits_{i=2}^{k} (x_i - x_{i-1})y_1 + \sum\limits_{i=2}^{k} (x_i - x_{i-1})\tilde{y}_i
\end{array}$$

$$\begin{aligned} &= [x]^\mathsf{T}[y] + \sum_{i=2}^{k} \underbrace{(x_i - x_{i-1})}_{\leqslant 0}\underbrace{(\tilde{y}_i - y_1)}_{\geqslant 0} \\ &\leqslant [x]^\mathsf{T}[y]. \end{aligned}$$

**10. Barème :** 3 points (pas beaucoup d'autres démonstrations sont justes)

Si $s \in \partial f(x)$, on a
$$f(x) + f^*(s) = s^\mathsf{T} x.$$
Par la symétrie de $f$ (hypothèse), celle de $f^*$ (question 5) et la question 9, on a
$$f([x]) + f^*([s]) = s^\mathsf{T} x \leqslant [s]^\mathsf{T}[x].$$
Donc $[s] \in \partial f([x])$.

**11. Barème :** 8 points

- On a
$$\begin{aligned} & S \in \partial(f \circ \lambda)(A) \\ & \iff (f \circ \lambda)(A) + (f \circ \lambda)^*(S) = \langle S, A \rangle \qquad \text{[proposition 3.49]} \\ & \iff f(\lambda(A)) + f^*(\lambda(S)) = \langle S, A \rangle \qquad \text{[par la question 6].} \end{aligned} \tag{6}$$

- $[(i) \Rightarrow (ii)]$ **Barème :** 2 points

  Supposons à présent que $S \in \partial(f \circ \lambda)(A)$. Par la définition de la conjuguée, on a
$$f(\lambda(A)) + f^*(\lambda(S)) \geqslant \lambda(S)^\mathsf{T} \lambda(A).$$
  Alors, l'équivalence (6) ci-dessus et l'inégalité de Fan (2) impliquent que $\langle S, A \rangle = \lambda(S)^\mathsf{T} \lambda(A)$ et que
$$f(\lambda(A)) + f^*(\lambda(S)) = \lambda(S)^\mathsf{T} \lambda(A). \tag{7}$$
  Cette identité montre que $\lambda(S) \in \partial f(\lambda(A))$ (proposition 3.49).

- $[(ii) \Rightarrow (i)]$ **Barème :** 1 point

  Supposons que $\langle S, A \rangle = \lambda(S)^\mathsf{T} \lambda(A)$ et que $\lambda(S) \in \partial f(\lambda(A))$. Alors (7) a lieu, de même que la partie droite de l'équivalence (6). Par celle-ci $S \in \partial(f \circ \lambda)(A)$.

- $[(ii) \Rightarrow (iii)]$ **Barème :** 1 point

  Comme $\langle S, A \rangle = \lambda(S)^\mathsf{T} \lambda(A)$, l'inégalité de Fan (2) est une égalité, si bien qu'il existe une matrice orthogonale $V$ telle que
$$V^\mathsf{T} A V = \operatorname{Diag}(\lambda(A)) \qquad \text{et} \qquad V^\mathsf{T} S V = \operatorname{Diag}(\lambda(S)).$$
  En définissant $a := \lambda(A)$ et $s := \lambda(S)$, et en utilisant $\lambda(S) \in \partial f(\lambda(A))$, on obtient $(iii)$.

- $[(iii) \Rightarrow (ii)]$ **Barème :** 4 points

  <u>Attention</u> : le fait que $A$ et $B$ soit diagonalisables par la même matrice orthogonale $V$ n'implique pas que l'on a égalité dans l'inégalité de Fan (2), car il faudrait aussi que les diagonales obtenues soient dans $\mathbb{R}^n_{\geqslant}$.

On a

$$\begin{aligned}
\lambda(S)^\top\lambda(A) &\leqslant f(\lambda(A)) + f^*(\lambda(S)) \qquad [\text{définition de } f^*] \\
&\leqslant f([a]) + f^*([s]) \qquad [V^\top AV = \mathrm{Diag}(a) \text{ et } V^\top SV = \mathrm{Diag}(s)] \\
&= f(a) + f^*(s) \qquad [f \text{ et } f^* \text{ sont symétriques}] \\
&= s^\top a \qquad [s \in \partial f(a)] \\
&= \langle \mathrm{Diag}(s), \mathrm{Diag}(a)\rangle \\
&= \langle S, A\rangle \qquad [V^\top AV = \mathrm{Diag}(a) \text{ et } V^\top SV = \mathrm{Diag}(s)].
\end{aligned}$$

Dès lors $\langle S, A\rangle = \lambda(S)^\top\lambda(A)$ par l'inégalité de Fan (2). Ensuite $s \in \partial f(a)$ implique que $[s] \in \partial f([a])$ (question 10) et donc $\lambda(S) \in \partial f(\lambda(A))$, car $\lambda(S) = [s]$ et $\lambda(A) = [a]$.

**12. Barème** : 12 points

- [$\Rightarrow$] **Barème** : 6 points (2 pour la sous-différentiabilité de $f$ en $\lambda(A)$, 4 pour la suite)

  Supposons que $(f \circ \lambda)$ soit Gâteaux-différentiable en $A \in \mathcal{S}^n$. Alors son sous-différentiel est un singleton
  $$\partial(f \circ \lambda)(A) = \{S\}.$$
  Par l'implication $(i) \Rightarrow (ii)$ de la question 11, $\lambda(S) \in \partial f(\lambda(A))$, si bien que $f$ est sous-différentiable en $\lambda(A)$.

  Soit à présent $\tilde{s} \in \partial f(\lambda(A))$. Il s'agit de montrer que $\tilde{s} = \lambda(S)$. On sait, par l'implication $(i) \Rightarrow (ii)$ de la question 11 $[(i) \Rightarrow (ii)]$ et les conditions d'égalité dans l'inégalité de Fan (2), qu'il existe une matrice orthogonale $V$ telle que
  $$V^\top SV = \mathrm{Diag}(\lambda(S)) \qquad \text{et} \qquad V^\top AV = \mathrm{Diag}(\lambda(A)).$$
  On introduit $\tilde{S} \coloneqq V\,\mathrm{Diag}(\tilde{s})V^\top$. Alors,
  $$V^\top\tilde{S}V = \mathrm{Diag}(\tilde{s}), \qquad V^\top AV = \mathrm{Diag}(\lambda(A)) \qquad \text{et} \qquad \tilde{s} \in \partial f(\lambda(A)).$$
  Par la question 11 $[(iii) \Rightarrow (i)]$, $\tilde{S} \in \partial(f \circ \lambda)(A)$, si bien que $\tilde{S} = S$ (unicité du sous-gradient de $(f \circ \lambda)$ en $A$). Maintenant, $\mathrm{Diag}(\lambda(S)) = V^\top SV = V^\top \S V = \mathrm{Diag}(\tilde{s})$ montre que $\tilde{s} = \lambda(S)$.

- [$\Leftarrow$] **Barème** : 6 points (2 pour la sous-différentiabilité de $(f \circ \lambda)$ en $A$, 4 pour la suite)

  Supposons que $f$ soit Gâteaux-différentiable en $\lambda(A) \in \mathbb{R}^n$. Alors son sous-différentiel est un singleton
  $$\partial f(\lambda(A)) = \{s\}.$$
  Soit $V$ orthogonale telle que $A = V\,\mathrm{Diag}(\lambda(A))V^\top$. On introduit $S \coloneqq V\,\mathrm{Diag}(s)V^\top$, si bien que
  $$V^\top SV = \mathrm{Diag}(s), \qquad V^\top AV = \mathrm{Diag}(\lambda(A)) \qquad \text{et} \qquad s \in \partial f(\lambda(A)).$$
  Par l'implication $(iii) \Rightarrow (i)$ de la question 11, $S \in \partial(f \circ \lambda)(A)$, si bien que $(f \circ \lambda)$ est sous-différentiable en $A$.

  Montrons à présent que
  $$\partial(f \circ \lambda)(A) = \{\tilde{V}\,\mathrm{Diag}(s)\tilde{V}^\top : \tilde{V} \text{ est orthogonale et } A = \tilde{V}\,\mathrm{Diag}(\lambda(A))\tilde{V}^\top\}. \tag{8}$$

  - [$\subset$] Soit $\tilde{S} \in \partial(f \circ \lambda)(A)$. Par l'implication $(i) \Rightarrow (ii)$ de la question 11, puis par les conditions d'égalité dans l'inégalité de Fan, il existe une matrice orthogonale $\tilde{V}$ telle que
    $$\tilde{V}^\top A\tilde{V} = \mathrm{Diag}(\lambda(A)), \qquad \tilde{V}^\top\tilde{S}\tilde{V} = \mathrm{Diag}(\lambda(\tilde{S})) \qquad \text{et} \qquad \lambda(\tilde{S}) \in \partial f(\lambda(A)).$$
    On en déduit que $\lambda(\tilde{S}) = s$ (unicité du sous-gradient de $f$ en $\lambda(A)$) et donc que $\tilde{S} = \tilde{V}\,\mathrm{Diag}(\lambda(\tilde{S}))\tilde{V}^\top = \tilde{V}\,\mathrm{Diag}(s)\tilde{V}^\top$.

∘ [⊃] Si $\tilde{S}$ est dans l'ensemble à droite dans (8), on a pour une matrice orthogonale $\tilde{V}$ :

$$\tilde{V}^{\top}\tilde{S}\tilde{V} = \mathrm{Diag}(s), \qquad \tilde{V}^{\top}A\tilde{V} = \mathrm{Diag}(\lambda(A)) \qquad \text{et} \qquad s \in \partial f(\lambda(A)).$$

Par l'implication $(iii) \Rightarrow (i)$ de la question 11, on en déduit que $\tilde{S} \in \partial(f \circ \lambda)(A)$.

Montrons à présent que $\partial(f \circ \lambda)(A)$ est un singleton. Cet ensemble est convexe (c'est un sous-différentiel) et, par (8),

$$\forall\, S \in \partial(f \circ \lambda)(A), \quad \|S\|_F^2 = \|s\|_2^2.$$

Alors $(f \circ \lambda)(A)$ est nécessairement un singleton, car $A \in \mathcal{S}^n \mapsto \|A\|_F^2 = \langle A, A\rangle$ est une fonction strictement convexe.