OPT201 – Optimisation Différentiable I

Contrôle des connaissances

- Durée : 1 h 30 (on ne peut probablement pas tout faire...).
- Dans chaque sujet, on peut admettre des résultats énoncés dans des questions ou sous-questions *précédentes* pour traiter la question ou sous-question courante.
- Ne pas écrire en rouge.
- On peut consulter tous les documents *distribués au cours*, ainsi que ses propres notes.
- *Éteindre tout appareil connecté à un réseau.*

Quatre sujets

1. Infsup. Soient X et Y deux ensembles et $\varphi : X \times Y \to \mathbb{R}$ une fonction. On suppose que

$$\forall\, x \in X: \quad 0 \leqslant \sup_{y\in Y} \varphi(x,y). \tag{1}$$

Dites quelles sont les inégalités ci-dessous que l'on peut déduire de (1) (c'est-à-dire, qui sont certainement vérifiées si (1) l'est) :

$$0 \leqslant \inf_{x\in X} \sup_{y\in Y} \varphi(x,y), \tag{2}$$

$$0 \leqslant \sup_{y\in Y} \inf_{x\in X} \varphi(x,y), \tag{3}$$

$$0 \leqslant \sup_{y\in Y} \sup_{x\in X} \varphi(x,y). \tag{4}$$

Justifiez bien vos réponses ; seules les justifications sont considérées.

2. Calcul de solution. On considère le problème d'optimisation à deux variables $x = (x_1, x_2) \in \mathbb{R}^2$ suivant :

$$\begin{cases} \min_x \ x_1^2 + 3x_1x_2 + x_2^2 \\ x_1x_2 \geqslant 0 \\ (x_1+x_2)^2 \geqslant 2(x_1+x_2) + 3. \end{cases} \tag{5}$$

1. Montrez théoriquement que ce problème a au moins une solution (*théoriquement* veut dire ici *sans calculer une solution*, ce qui est le sujet des points suivants, mais par des arguments de nature topologique).
2. Montrez que les contraintes sont qualifiées en tout point admissible.
3. Calculez tous les points stationnaires du problème.
4. Déterminez la ou les solutions (globales) du problème.

3. Implication d'inégalités affines. Soient $A \in \mathbb{R}^{m\times n}$, $a \in \mathbb{R}^m$, $b \in \mathbb{R}^n$ et $\beta \in \mathbb{R}$. On suppose qu'il existe un x_0 tel que $Ax_0 \geqslant a$. Montrez que les propriétés suivantes sont équivalentes :

(i) tout $x \in \mathbb{R}^n$ vérifiant $Ax \geqslant a$ vérifie aussi $b^\top x \geqslant \beta$,

(ii) il existe un $y \in \mathbb{R}^m_+$ tel que $b = A^\top y$ et $a^\top y \geqslant \beta$.

[*Indication* : Pour l'implication (i) $\Rightarrow$ (ii), on pourra exprimer (i) au moyen d'un problème d'optimisation linéaire.]

4. Alternative de Motzkin non homogène. Soient $A \in \mathbb{R}^{m_A \times n}$ et $B \in \mathbb{R}^{m_B \times n}$ deux matrices ayant un même nombre de colonnes, $a \in \mathbb{R}^{m_A}$ et $b \in \mathbb{R}^{m_B}$ deux vecteurs. On souhaite montrer que les deux affirmations suivantes sont équivalentes (attention à l'inégalité stricte) :

- (i) $\exists\, x \in \mathbb{R}^n$ tel que $Ax = a$ et $Bx < b$,
- (ii) pour tout $(\alpha, \beta) \in \mathbb{R}^{m_A} \times \mathbb{R}^{m_B}$ vérifiant $A^\mathsf{T}\alpha + B^\mathsf{T}\beta = 0$, $\beta \geqslant 0$ et $a^\mathsf{T}\alpha + b^\mathsf{T}\beta \leqslant 0$, on a $a^\mathsf{T}\alpha = 0$ et $\beta = 0$.

On considère le problème d'optimisation linéaire en $(\alpha, \beta) \in \mathbb{R}^{m_A} \times \mathbb{R}^{m_B}$ suivant

$$(P) \quad \begin{cases} \inf_{(\alpha,\beta)} \ a^\mathsf{T}\alpha + b^\mathsf{T}\beta \\ A^\mathsf{T}\alpha + B^\mathsf{T}\beta = 0 \\ \beta \geqslant 0. \end{cases}$$

1. Écrivez le dual lagrangien (D) de (P), en utilisant le lagrangien

$$\ell : (\mathbb{R}^{m_A} \times \mathbb{R}^{m_B}) \times (\mathbb{R}^n \times \mathbb{R}^{m_B}) \to \mathbb{R},$$

défini en $((\alpha, \beta), (x, s))$ par

$$\ell((\alpha, \beta), (x, s)) = a^\mathsf{T}\alpha + b^\mathsf{T}\beta - x^\mathsf{T}(A^\mathsf{T}\alpha + B^\mathsf{T}\beta) - s^\mathsf{T}\beta.$$

2. Montrez que si (i) a lieu, alors (D) a une solution et en déduire (ii).
3. Montrez que si (ii) a lieu, alors (P) a une solution et en déduire l'existence d'un $x_0 \in \mathbb{R}^n$ tel que

$$Ax_0 = a \qquad \text{et} \qquad Bx_0 \leqslant b.$$

4. On considère le problème d'optimisation linéaire en $(\alpha, \beta) \in \mathbb{R}^{m_A} \times \mathbb{R}^{m_B}$ suivant

$$(P_1) \quad \begin{cases} \inf_{(\alpha,\beta)} \ -e^\mathsf{T}\beta \\ A^\mathsf{T}\alpha + B^\mathsf{T}\beta = 0 \\ \beta \geqslant 0 \\ a^\mathsf{T}\alpha + b^\mathsf{T}\beta \leqslant 0, \end{cases}$$

où e est le vecteur de $\mathbb{R}^{m_B}$ dont toutes les composantes valent 1.
Montrez que si (ii) a lieu, alors ce problème (P_1) a une solution (α_1, β_1) et en déduire (i).

OPT201 – Optimisation Différentiable I
Contrôle des connaissances (solutions succinctes)

1. Les inégalités (1) disent que la fonction $f : x \in X \mapsto \sup_{y\in Y} \varphi(x, y)$ est minorée par zéro. Alors, on a certainement $0 \leqslant \inf_{x\in X} f(x)$, qui n'est autre que (2).

On a aussi $0 \leqslant \sup_{x\in X} f(x)$, qui, après l'inversion licite des deux suprema, conduit à (4).

L'inégalité (3) n'est pas garantie, car il peut y avoir un saut de dualité.

Voici un contre-exemple de (3) dans lequel $X = \mathbb{R}$, $Y = \{0, 1\}$ et la fonction φ est définie en $(x, y) \in X \times Y$ par

$$\varphi(x,y) = \begin{cases} x^2 - 1 & \text{si } y = 0, \\ (x-2)^2 - 1 & \text{si } y = 1. \end{cases}$$

On a

$$\sup_{y\in\{0,1\}} \varphi(x,y) = \begin{cases} (x-2)^2 - 1 & \text{si } x \leqslant 1, \\ x^2 - 1 & \text{si } x > 1. \end{cases} \qquad \text{et} \qquad \inf_{x\in\mathbb{R}} \varphi(x,y) = \begin{cases} -1 & \text{si } y = 0, \\ -1 & \text{si } y = 1. \end{cases}$$

Donc (1) a bien lieu, mais

$$\sup_{y\in\{0,1\}} \inf_{x\in\mathbb{R}} \varphi(x,y) = -1$$

n'est pas positif.

2.
1. L'ensemble admissible est un fermé (contraintes continues, inégalités non strictes) non vide (contient $(3, 0)$), mais n'est pas borné (continent $(t, 0)$ pour tout $t \geqslant 3$). Le critère f est continu (polynôme), mais n'est pas coercif sur $\mathbb{R}^2$. Montrons que le critère f est coercif *sur l'ensemble admissible* X, ce qui impliquera l'existence d'une solution.

   Le critère est minoré par $(x_1 + x_2)^2$. Or si $\|x\| \to \infty$ dans X, on a nécessairement $|x_1 + x_2| \to \infty$ (car x_1 et x_2 ont le même signe dans X). On en déduit que le critère tend vers $+\infty$ lorsque $\|x\| \to \infty$ avec $x \in X$.
2. On cherche à voir si (QC-IL) a lieu. Les gradients des contraintes s'écrivent

   $$-\begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \qquad \text{et} \qquad -2(x_1 + x_2 - 1)\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

   Ils sont linéairement indépendants si $x_1 + x_2 \neq 1$ et si $x_1 \neq x_2$.
   Dans l'ensemble admissible, on a toujours $x_1 + x_2 \neq 1$ et les seuls points du bord de l'ensemble admissible où $x_1 = x_2$ sont $x = -e/2$ et $x = 3e/2$ (on a noté $e := (1, 1)$). En ces derniers points, seule la seconde contrainte est active et son gradient est non nul. Donc (QC-IL) a lieu.
3. On utilise le lagrangien $\ell : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ défini en $(x, \lambda) \in \mathbb{R}^2 \times \mathbb{R}^2$ par

   $$\ell(x,\lambda) = x_1^2 + 3x_1x_2 + x_2^2 - \lambda_1(x_1x_2) - \lambda_2(x_1^2 + 2x_1x_2 + x_2^2 - 2x_1 - 2x_2 - 3).$$

   Les conditions de KKT s'écrivent

   $$\begin{cases} 2x_1 + 3x_2 - \lambda_1 x_2 - \lambda_2(2x_1 + 2x_2 - 2) = 0 \\ 3x_1 + 2x_2 - \lambda_1 x_1 - \lambda_2(2x_1 + 2x_2 - 2) = 0 \\ 0 \leqslant \lambda_1 \perp x_1x_2 \geqslant 0 \\ 0 \leqslant \lambda_2 \perp (x_1^2 + 2x_1x_2 + x_2^2 - 2x_1 - 2x_2 - 1) \geqslant 0. \end{cases}$$

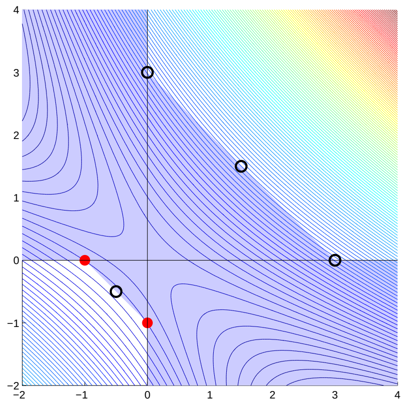

Figure 1: Les solutions (points rouges) du problème (5), quatre autres minima locaux (cercles noirs) et les courbes de niveau du critère (qui est coercif sur l'ensemble admissible, mais pas sur $\mathbb{R}^2$)

Calculons des points stationnaires (ils sont représentés à la figure 1).
On choisit de considérer les quatre cas possibles en fonction de la valeur prise par les multiplicateurs.

- $\boldsymbol{\lambda_1 = \lambda_2 = 0}$. Alors $x = 0$, qui n'est pas admissible.
- $\boldsymbol{\lambda_1 > 0}$ **et** $\boldsymbol{\lambda_2 = 0}$. Alors il faut résoudre
$$\begin{cases} 2x_1 + 3x_2 - \lambda_1 x_2 = 0 \\ 3x_1 + 2x_2 - \lambda_1 x_1 = 0 \\ x_1 x_2 = 0, \end{cases}$$
dont la solution est $x = 0$ (si $x_1 = 0$ alors $x_2 = 0$, si $x_2 = 0$ alors $x_1 = 0$), qui n'est pas admissible.
- $\boldsymbol{\lambda_1 = 0}$ **et** $\boldsymbol{\lambda_2 > 0}$. Alors il faut résoudre
$$\begin{cases} 2x_1 + 3x_2 - \lambda_2(2x_1 + 2x_2 - 2) = 0 \\ 3x_1 + 2x_2 - \lambda_2(2x_1 + 2x_2 - 2) = 0 \\ (x_1 + x_2)^2 - 2(x_1 + x_2) - 3 = 0. \end{cases}$$
La dernière équation montre $x_1 + x_2 = 3$ ou -1. Si $x_1 + x_2 = 3$ les deux premières équations donnent
$$\begin{cases} -x_1 - 4\lambda_2 = -9 \\ x_1 - 4\lambda_2 = -6. \end{cases}$$
On en déduit le point stationnaire
$$x = (3/2, 3/2), \qquad \lambda = (0, 15/8) \qquad \text{et} \qquad f(x) = 45/4. \tag{6}$$
Si $x_1 + x_2 = -1$ les deux premières équations donnent
$$\begin{cases} -x_1 + 4\lambda_2 = 3 \\ x_1 + 4\lambda_2 = 2 \end{cases}$$
On en déduit le point stationnaire
$$x = (-1/2, -1/2), \qquad \lambda = (0, 5/4) \qquad \text{et} \qquad f(x) = 5/4. \tag{7}$$
- $\boldsymbol{\lambda_1 > 0}$ **et** $\boldsymbol{\lambda_2 > 0}$. Alors il faut résoudre
$$\begin{cases} 2x_1 + 3x_2 - \lambda_1 x_2 - \lambda_2(2x_1 + 2x_2 - 2) = 0 \\ 3x_1 + 2x_2 - \lambda_1 x_1 - \lambda_2(2x_1 + 2x_2 - 2) = 0 \\ x_1 x_2 = 0 \\ (x_1 + x_2)^2 - 2(x_1 + x_2) - 3 = 0. \end{cases}$$

Si $x_2 = 0$, on doit résoudre

$$\begin{cases} 2x_1 - \lambda_2(2x_1 - 2) = 0 \\ 3x_1 - \lambda_1 x_1 - \lambda_2(2x_1 - 2) = 0 \\ x_1^2 - 2x_1 - 3 = 0, \end{cases}$$

qui a pour solution $x_1 = -1$ ou 3. On en déduit les deux points stationnaires

$$x = (-1, 0), \qquad \lambda = (1, 1/2) \qquad \text{et} \qquad f(x) = 1, \tag{8}$$
$$x = (3, 0), \qquad \lambda = (1, 3/2) \qquad \text{et} \qquad f(x) = 9. \tag{9}$$

Si $x_1 = 0$, on doit résoudre

$$\begin{cases} 3x_2 - \lambda_1 x_2 - \lambda_2(2x_2 - 2) = 0 \\ 2x_2 - \lambda_2(2x_2 - 2) = 0 \\ x_2^2 - 2x_2 - 3 = 0, \end{cases}$$

qui a pour solution $x_2 = -1$ ou 3. On en déduit les deux points stationnaires

$$x = (0, -1), \qquad \lambda = (1, 1/2) \qquad \text{et} \qquad f(x) = 1, \tag{10}$$
$$x = (0, 3), \qquad \lambda = (1, 3/2) \qquad \text{et} \qquad f(x) = 9. \tag{11}$$

4. Comme les solutions (globales) du problème sont parmi les points stationnaires trouvés (car les contraintes sont qualifiées en tout point de l'ensemble admissible), il suffit de comparer les valeurs de l'objectif en ces points. On voit sur (6)-(11) que la valeur optimale est 1 et que les minima globaux sont

$$x = (-1, 0), \qquad \lambda = (1, 1/2) \qquad \text{et} \qquad f(x) = 1,$$
$$x = (0, -1), \qquad \lambda = (1, 1/2) \qquad \text{et} \qquad f(x) = 1.$$

3. $[(i) \Rightarrow (ii)]$

Considérons le problème d'optimisation linéaire

$$\begin{cases} \inf b^\mathsf{T} x \\ Ax \geqslant a. \end{cases}$$

Par hypothèse, ce problème est réalisable (par x_0) si bien que sa valeur optimale n'est pas $+\infty$. Par ailleurs, si (i) a lieu, la valeur optimale est $\geqslant \beta$. Le problème a donc une solution, disons $\bar{x} \in \mathbb{R}^n$. Le lagrangien du problème s'écrit $(x, \lambda) \mapsto b^\mathsf{T} x + y^\mathsf{T}(a - Ax)$. Par les conditions d'optimalité, on a l'existence d'un $y \in \mathbb{R}^m$, tel que

$$b = A^\mathsf{T} y, \qquad 0 \leqslant y \perp (A\bar{x} - a) \geqslant 0.$$

La condition (ii) a donc bien lieu, puisque l'on a aussi

$$a^\mathsf{T} y = (A\bar{x})^\mathsf{T} y = \bar{x}^\mathsf{T} b \geqslant \beta.$$

$[(ii) \Rightarrow (i)]$

Supposons que (ii) ait lieu et que x vérifie $Ax \geqslant a$. Alors

$$b^\mathsf{T} x = \underbrace{y}_{\geqslant 0}{}^\mathsf{T} \underbrace{Ax}_{\geqslant a} \geqslant y^\mathsf{T} a \geqslant \beta.$$

Donc (i) a lieu.

4. 1. Le problème primal peut s'écrire

$$\inf_{(\alpha,\beta)\in\mathbb{R}^{m_A}\times\mathbb{R}^{m_B}} \sup_{\substack{(x,s)\in\mathbb{R}^n\times\mathbb{R}^{m_B}\\ s\geqslant 0}} \ell((\alpha,\beta),(x,s)).$$

Son dual lagrangien s'écrit donc

$$\sup_{\substack{(x,s)\\ s\geqslant 0}} \inf_{(\alpha,\beta)} \ell((\alpha,\beta),(x,s)) \quad = \quad \sup_{\substack{(x,s)\\ s\geqslant 0}} \inf_{(\alpha,\beta)} (a-Ax)^{\mathsf{T}}\alpha + (b-Bx-s)^{\mathsf{T}}\beta,$$

qui est le problème en $x \in \mathbb{R}^n$ suivant

$$(D) \quad \begin{cases} \sup_x \ 0 \\ Ax = a \\ Bx \leqslant b. \end{cases}$$

2. Si (i) a lieu, (D) est réalisable et borné, il a donc le x donné dans (i) comme solution et sa valeur optimale est $\mathrm{val}(D) = 0$. Alors (P) a aussi une solution et sa valeur optimale $\mathrm{val}(P) = 0$ (pas de saut de dualité). Alors, si $(\alpha,\beta) \in \mathbb{R}^{m_A} \times \mathbb{R}^{m_B}$ vérifie

$$A^{\mathsf{T}}\alpha + B^{\mathsf{T}}\beta = 0, \qquad \beta \geqslant 0 \qquad \text{et} \qquad a^{\mathsf{T}}\alpha + b^{\mathsf{T}}\beta \leqslant 0,$$

(α,β) est solution de (P), si bien que $a^{\mathsf{T}}\alpha + b^{\mathsf{T}}\beta = \mathrm{val}(P) = 0$ et on a la complémentarité (les solutions primales-duales forment un produit cartésien)

$$0 \leqslant \beta \perp (b - Bx) > 0.$$

Donc $\beta = 0$ et, par conséquent, $a^{\mathsf{T}}\alpha = 0$. On a montré (ii).

3. Le problème (P) est réalisable et de valeur optimale $\mathrm{val}(P) \leqslant 0$ (en prenant $\alpha = 0$ et $\beta = 0$).

   Si (ii) a lieu, sa valeur optimale est finie ($\mathrm{val}(P) = 0$). Dès lors le problème a une solution et aussi une solution duale (x_0, s_0) qui vérifie

$$Ax_0 = b \qquad \text{et} \qquad Bx_0 \leqslant b.$$

4. Le problème (P_1) est réalisable (en prenant $\alpha = 0$ et $\beta = 0$). Lorsque (ii) a lieu, sa valeur optimale est finie ($\mathrm{val}(P_1) = 0$). Il a donc une solution primale (α_1, β_1) et une solution duale (x_1, s_1, t_1) qui vérifie

$$\begin{cases} Ax_1 = t_1 a \\ Bx_1 + s_1 + e = t_1 b \\ 0 \leqslant s_1 \perp \beta_1 \geqslant 0 \\ 0 \leqslant t_1 \perp (a^{\mathsf{T}}\alpha_1 + b^{\mathsf{T}}\beta_1) \leqslant 0. \end{cases}$$

   - Si $t_1 > 0$, alors on prend $x = x_1/t_1$ qui vérifie $Ax = a$ et $Bx < b$.
   - Si $t_1 = 0$, alors on prend $x = x_0 + x_1$ qui vérifie $Ax = a$ et $Bx < b$.

   On a montré (i).