SOLUTIONS

1 Existence d'une solution primale-duale strictement complémentaire

1. • *Montrons que* $B \cap N = \varnothing$.

   On ne peut avoir $i \in B \cap N$, car on aurait alors une solution primale x avec $x_i > 0$ et une solution duale (y, s) avec $s_i > 0$. Comme l'ensemble des solutions primales-duales est un produit cartésien (ce sont des points-selles du lagrangien), le triplet (x, y, s) serait une solution primale-duale qui contredirait la complémentarité.

   • *Montrons que* $B \cup N = [1:n]$ *si, et seulement si, l'on a complémentarité stricte.*

   [⇒] La convexité de $\mathcal{S}_{\text{P}}$ implique qu'il existe $x \in \mathcal{S}_{\text{P}}$ tel que $x_B > 0$. De même, la convexité de $\mathcal{S}_{\text{D}}$ implique qu'il existe $(y, s) \in \mathcal{S}_{\text{D}}$ tel que $s_N > 0$. Alors, $B \cup N = [1:n]$ montre que (x, y, s) est une solution primale-duale est strictement complémentaire (on utilise aussi le fait que l'ensemble des solutions primales-duales est un produit cartésien).

   [⇐] Si (x, y, s) est une solution primale-duale strictement complémentaire, alors $\{i \in [1:n] : x_i > 0\} \subset B$ et $\{i \in [1:n] : s_i > 0\} \subset N$. Par (10.1), on voit que $B \cup N = [1:n]$.

2. L'ensemble admissible de (P_1) est formé des solutions de (P_0) et nous avons supposé qu'il y en avait. Par ailleurs, par définition de B, sa valeur optimale est nulle. Ce problème d'optimisation linéaire a donc les mêmes solutions que celles du problème (P_0).

3. On choisit une solution x_0 de (P_0) avec $(x_0)_B > 0$. Les conditions d'optimalité assure l'existence de (y_0, s_0) tel que

$$\left\{\begin{array}{l} A^{\mathsf{T}} y_0 + s_0 = c \\ A x_0 = b \\ 0 \leqslant s_0 \perp x_0 \geqslant 0. \end{array}\right.$$

   Comme x_0 est aussi solution de (P_1), il existe $(y_1, s_1, t_1) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}$ tel que

$$\left\{\begin{array}{l} A^{\mathsf{T}} y_1 + s_1 + e_{B^c} = t_1 c \\ A x_0 = b \\ 0 \leqslant s_1 \perp x_0 \geqslant 0 \\ 0 \leqslant t_1 \perp (c^{\mathsf{T}} x_0 - \operatorname{val}(P_0)) \leqslant 0. \end{array}\right.$$

   En additionnant les relations correspondantes de ces deux systèmes, on voit que (x, y, s) défini par

$$x := x_0, \qquad y := \frac{y_0 + y_1}{1 + t_1} \qquad \text{et} \qquad s := \frac{s_0 + s_1 + e_{B^c}}{1 + t_1}$$

   est solution primale-duale de (P_0), que $x_B = (x_0)_B > 0$ et que $s_{B^c} > 0$. Il s'agit donc d'une solution strictement complémentaire de (P_0).

2 Polyèdres convexes imbriqués

- [$(i) \Rightarrow (ii)$] À première vue, le problème de savoir quand est-ce que le polyèdre P_A est contenu dans le polyèdre P_B semble compliqué. Le problème se clarifie si l'on remarque que cela revient à montrer que P_A est contenu dans chaque demi-espace $\{x \in \mathbb{R}^n : B_{i:}x \leqslant b_1\}$, pour $i \in [1:m_B]$, où $B_{i:}$ est la i-ième ligne de B. Ce dernier problème peut s'exprimer par l'optimisation linéaire.

  Soient $i \in [1:m_B]$. Considérons le problème d'optimisation linéaire en $x \in \mathbb{R}^n$ suivant

  $$\begin{cases} \inf_x \ -B_{i:}x \\ Ax \leqslant a. \end{cases}$$

  Par hypothèse, ce problème est réalisable (car $P_A \neq \varnothing$)) si bien que sa valeur optimale n'est pas $+\infty$. Par ailleurs, si (i) a lieu, sa valeur optimale est $\geqslant -b_i$, où b_i est la i-ième composante de b, car chaque fois que $Ax \leqslant a$, on a $Bx \leqslant b$ et donc $B_{i:}x \leqslant b_i$. Le problème a donc une solution, disons $\bar{x}_i \in \mathbb{R}^n$. Le lagrangien du problème s'écrit

  $$(x, y_i) \in \mathbb{R}^n \times \mathbb{R}^{m_A} \mapsto -B_{i:}x + y_i^{\mathsf{T}}(Ax - a).$$

  Par les conditions d'optimalité, on a l'existence d'un $y_i \in \mathbb{R}^{m_A}$, tel que

  $$B_{i:} = y_i^{\mathsf{T}}A, \qquad 0 \leqslant y_i \perp (A\bar{x}_i - a) \leqslant 0.$$

  Alors, on a

  $$\begin{array}{rcll} y_i^{\mathsf{T}}a & = & y_i^{\mathsf{T}}A\bar{x}_i & \text{[complémentarité]} \\ & = & B_{i:}\bar{x}_i & [B_{i:} = y_i^{\mathsf{T}}A] \\ & \leqslant & b_i & \text{[valeur optimale } \geqslant -b_i]. \end{array}$$

  On obtient la condition (ii) en prenant la i-ième ligne de Y égale à y_i^{T}.

- [$(ii) \Rightarrow (i)$] Supposons que (ii) ait lieu et que $x \in P_A$, si bien que $Ax \leqslant a$. Alors, en utilisant les propriétés de Y données dans (ii), on a

  $$Bx = \underbrace{Y}_{\geqslant 0}\underbrace{Ax}_{\leqslant a} \leqslant Ya \leqslant b.$$

  Donc $x \in P_B$ et (i) a lieu.

3 Alternative de Motzkin non homogène

Voici deux démonstration possibles.

1) *On utilise le lemme de Farkas.*

   - [$(i) \Rightarrow (ii)$] Comme (i) a lieu, il existe $x \in \mathbb{R}$, $s_B \in \mathbb{R}^{m_B}_+$, $s_C \in \mathbb{R}^{m_C}_+$ et $\varepsilon > 0$ tels que

     $$Ax = a, \qquad Bx + s_B = b, \qquad Cx + s_C = c - \varepsilon e,$$

     où e est le vecteur de uns de $\mathbb{R}^{m_C}$. Ceci peut s'écrire

     $$\begin{pmatrix} a \\ b \\ c - \varepsilon e \end{pmatrix} \in \begin{pmatrix} A & 0 & 0 \\ B & I & 0 \\ 0 & 0 & I \end{pmatrix} (\mathbb{R}^n \times \mathbb{R}^{m_B}_+ \times \mathbb{R}^{m_C}_+). \tag{10.3}$$

Par le lemme de Farkas, cette relation est équivalente à la suivante

$$\begin{pmatrix} a \\ b \\ c-\varepsilon e \end{pmatrix} \in \Bigg\{ (\alpha,\beta,\gamma)\in\mathbb{R}^{m_A}\times\mathbb{R}^{m_B}\times\mathbb{R}^{m_C} : \qquad\qquad \begin{pmatrix} A^\mathsf{T} & B^\mathsf{T} & C^\mathsf{T} \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \in \{0_{\mathbb{R}^n}\}\times\mathbb{R}^{m_B}_+\times\mathbb{R}^{m_C}_+ \Bigg\}^+ . \tag{10.4}$$

Dès lors, si $(\alpha,\beta\,\gamma)\in\mathbb{R}^{m_A}\times\mathbb{R}^{m_B}_+\times\mathbb{R}^{m_C}_+$ vérifie $A^\mathsf{T}\alpha+B^\mathsf{T}\beta+C^\mathsf{T}\gamma=0$, on a

$$a^\mathsf{T}\alpha+b^\mathsf{T}\beta+c^\mathsf{T}\gamma\geqslant\varepsilon e^\mathsf{T}\gamma.$$

Par conséquent, si les hypothèses du point (ii) ont lieu, c'est-à-dire si $(\alpha,\beta\,\gamma)\in\mathbb{R}^{m_A}\times\mathbb{R}^{m_B}_+\times\mathbb{R}^{m_C}_+$ vérifie $A^\mathsf{T}\alpha+B^\mathsf{T}\beta+C^\mathsf{T}\gamma=0$ et $a^\mathsf{T}\alpha+b^\mathsf{T}\beta+c^\mathsf{T}\gamma\leqslant 0$, on a

$$0\geqslant a^\mathsf{T}\alpha+b^\mathsf{T}\beta+c^\mathsf{T}\gamma\geqslant\varepsilon e^\mathsf{T}\gamma\geqslant 0,$$

si bien que les inégalités de cette chaîne sont en réalité des égalités. On en déduit que $\gamma=0$ (parce que $\varepsilon e^\mathsf{T}\gamma=0$ et $\varepsilon e>0$) et aussi que $a^\mathsf{T}\alpha+b^\mathsf{T}\beta=0$ (parce que $a^\mathsf{T}\alpha+b^\mathsf{T}\beta+c^\mathsf{T}\gamma=0$ et $\gamma=0$).

- $[(ii)\Rightarrow(i)]$ Il suffit de trouver un $\varepsilon>0$ tel que l'on ait (10.4), car alors (10.3) a lieu (par le lemme de Farkas) et (i) s'ensuit. Autrement dit, il suffit de montrer que

  $\exists\,\varepsilon>0$, $\forall\,(\alpha,\beta\,\gamma)\in\mathbb{R}^{m_A}\times\mathbb{R}^{m_B}_+\times\mathbb{R}^{m_C}_+$ vérifiant
  $A^\mathsf{T}\alpha+B^\mathsf{T}\beta+C^\mathsf{T}\gamma=0$, on a $a^\mathsf{T}\alpha+b^\mathsf{T}\beta+c^\mathsf{T}\gamma\geqslant\varepsilon e^\mathsf{T}\gamma$.

  On choisit de raisonner par l'absurde en supposant que

  $$\begin{array}{l}\forall\,\varepsilon>0,\ \exists\,(\alpha_\varepsilon,\beta_\varepsilon\,\gamma_\varepsilon)\in\mathbb{R}^{m_A}\times\mathbb{R}^{m_B}_+\times\mathbb{R}^{m_C}_+ \text{ vérifiant} \\ A^\mathsf{T}\alpha_\varepsilon+B^\mathsf{T}\beta_\varepsilon+C^\mathsf{T}\gamma_\varepsilon=0 \text{ et } a^\mathsf{T}\alpha_\varepsilon+b^\mathsf{T}\beta_\varepsilon+c^\mathsf{T}\gamma_\varepsilon<\varepsilon e^\mathsf{T}\gamma_\varepsilon.\end{array} \tag{10.5}$$

  On prend une suite de $\varepsilon\downarrow 0$ et on note

  $$t_\varepsilon:=\max(|a^\mathsf{T}\alpha_\varepsilon+b^\mathsf{T}\beta_\varepsilon|,\|\gamma_\varepsilon\|) \qquad\text{et}\qquad M:=\begin{pmatrix} A^\mathsf{T} & B^\mathsf{T} & C^\mathsf{T} \\ a^\mathsf{T} & b^\mathsf{T} & c^\mathsf{T} \end{pmatrix},$$

  où $\|\cdot\|$ est la norme euclidienne (ce choix raffiné de t_ε est guidé par la forme de la condition (ii) ; le choix plus commun de $t_\varepsilon:=\|(\alpha_\varepsilon,\beta_\varepsilon,\gamma_\varepsilon)\|$ permettrait de passer à la limite comme ci-dessous, mais pas de conclure).

  Observons que $t_\varepsilon>0$, car sinon l'inégalité stricte dans (10.5) ne peut avoir lieu. Observons ensuite qu'avec $z_\varepsilon:=(\alpha_\varepsilon,\beta_\varepsilon\,\gamma_\varepsilon)/t_\varepsilon$, la suite $\{Mz_\varepsilon\}$ est bornée dans $\mathbb{R}^{n+1}$ (mais $\{z_\varepsilon\}$ n'est pas nécessairement bornée), parce que les n premières composantes de Mz_ε sont nulles et sa dernière composante est majorée par

  $$\frac{|a^\mathsf{T}\alpha_\varepsilon+b^\mathsf{T}\beta_\varepsilon|}{t_\varepsilon}+\frac{|c^\mathsf{T}\gamma_\varepsilon|}{t_\varepsilon}\leqslant\frac{|a^\mathsf{T}\alpha_\varepsilon+b^\mathsf{T}\beta_\varepsilon|}{|a^\mathsf{T}\alpha_\varepsilon+b^\mathsf{T}\beta_\varepsilon|}+\frac{\|c\|\,\|\gamma_\varepsilon\|}{\|\gamma_\varepsilon\|}=1+\|c\|.$$

  En extrayant une sous-suite au besoin, on peut donc supposer que $\{Mz_\varepsilon\}$ converge. Mais $\{Mz_\varepsilon\}\subset M(\mathbb{R}^{m_A}\times\mathbb{R}^{m_B}_+\times\mathbb{R}^{m_C}_+)$, qui est un polyèdre convexe (donc un fermé), donc sa limite appartient à ce polyèdre convexe. Alors, on peut trouver $(\alpha,\beta\,\gamma)\in\mathbb{R}^{m_A}\times\mathbb{R}^{m_B}_+\times\mathbb{R}^{m_C}_+$ tel que $Mz_\varepsilon\to M(\alpha,\beta\,\gamma)$.

En divisant les membres de l'égalité et de l'inégalité stricte de (10.5) par t_ε, en observant que $\{e^\mathsf{T}\gamma_\varepsilon/t_\varepsilon\}$ est bornée par $\|e\|$ et en passant à la limite lorsque $\varepsilon \downarrow 0$, on trouve

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

Par (ii), on en déduit que

$$a^\mathsf{T}\alpha + b^\mathsf{T}\beta = 0 \qquad \text{et} \qquad \gamma = 0.$$

Ceci apporte la contradiction attendue car soit $t_\varepsilon = \|\gamma_\varepsilon\|$ pour une infinité de ε et dans ce cas on ne peut pas avoir $\gamma = 0$, soit $t_\varepsilon = |a^\mathsf{T}\alpha_\varepsilon + b^\mathsf{T}\beta_\varepsilon|$ pour une infinité de ε et dans ce cas on ne peut pas avoir $a^\mathsf{T}\alpha + b^\mathsf{T}\beta = 0$.

2) *On utilise la dualité en optimisation linéaire.*

La condition (i) suggère de considérer le problème d'optimisation linéaire suivant

$$\begin{cases} \inf_x \ 0 \\ Ax = a \\ Bx \leqslant b \\ Cx \leqslant c, \end{cases} \tag{10.6}$$

dans lequel la contrainte d'inégalité stricte $Cx < c$ de (i) a été remplacée par une contrainte d'inégalité non stricte pour avoir un ensemble admissible fermé. On cherche alors ce que peut nous apprendre son dual lagrangien. Comme ce problème s'écrit aussi

$$\inf_x \sup_{\substack{\alpha \\ \beta \geqslant 0 \\ \gamma \geqslant 0}} \ \alpha^\mathsf{T}(Ax - a) + \beta^\mathsf{T}(Bx - b) + \gamma^\mathsf{T}(Cx - c),$$

son dual lagrangien s'écrit

$$\begin{aligned} &\sup_{\substack{\alpha \\ \beta \geqslant 0 \\ \gamma \geqslant 0}} \inf_x \ \alpha^\mathsf{T}(Ax - a) + \beta^\mathsf{T}(Bx - b) + \gamma^\mathsf{T}(Cx - c) \\ &\quad = \ \sup_{\substack{\alpha \\ \beta \geqslant 0 \\ \gamma \geqslant 0}} \inf_x \ -a^\mathsf{T}\alpha - b^\mathsf{T}\beta - c^\mathsf{T}\gamma + (A^\mathsf{T}\alpha + B^\mathsf{T}\beta + C^\mathsf{T}\gamma)^\mathsf{T}x \end{aligned}$$

ou encore

$$\begin{cases} \sup_{(\alpha,\beta,\gamma)} \ -a^\mathsf{T}\alpha - b^\mathsf{T}\beta - c^\mathsf{T}\gamma \\ A^\mathsf{T}\alpha + B^\mathsf{T}\beta + C^\mathsf{T}\gamma = 0 \\ \beta \geqslant 0 \\ \gamma \geqslant 0. \end{cases} \tag{10.7}$$

- $[(i) \Rightarrow (ii)]$ Si (i) a lieu, le x donné vérifie les contraintes de (10.6) et ce problème a ce point x comme solution et comme valeur optimale $\operatorname{val}(10.6) = 0$. Alors le problème dual (10.7) a aussi une solution et une valeur optimale nulle.

  Soit $(\alpha, \beta, \gamma) \in \mathbb{R}^{m_A} \times \mathbb{R}^{m_B}_+ \times \mathbb{R}^{m_C}_+$ vérifiant $A^\mathsf{T}\alpha + B^\mathsf{T}\beta + C^\mathsf{T}\gamma = 0$ et $a^\mathsf{T}\alpha + b^\mathsf{T}\beta + c^\mathsf{T}\gamma \leqslant 0$ comme dans (ii). Alors (α, β, γ) est solution du problème dual (10.7) et on a

  $$a^\mathsf{T}\alpha + b^\mathsf{T}\beta + c^\mathsf{T}\gamma = 0. \tag{10.8}$$

Enfin, en retranchant $(A^\mathsf{T}\alpha + B^\mathsf{T}\beta + C^\mathsf{T}\gamma)^\mathsf{T} x = 0$ (où x est la solution primale donnée par (i)) et (10.8), on obtient

$$\underbrace{(Ax-a)}_{=0}{}^\mathsf{T}\alpha + \underbrace{(Bx-b)}_{\leqslant 0}{}^\mathsf{T}\underbrace{\beta}_{\geqslant 0} + \underbrace{(Cx-c)}_{<0}{}^\mathsf{T}\underbrace{\gamma}_{\geqslant 0} = 0.$$

On en déduit que $\gamma = 0$. Alors (10.8) délivre $a^\mathsf{T}\alpha + b^\mathsf{T}\beta = 0$.

- $[(ii) \Rightarrow (i)]$ Le problème dual (10.7) est réalisable et sa valeur optimale est positive (en prenant $\alpha = 0$, $\beta = 0$ et $\gamma = 0$). Si (ii) a lieu, alors $\operatorname{val}(10.7) = 0$ et le problème dual (10.7) a une solution. Dès lors, le problème primal (10.6) a aussi une solution (dualité forte). En prenant des solutions primales-duales strictement complémentaires (il en existe toujours en OL), disons x_0 et $(\alpha_0, \beta_0, \gamma_0)$, on doit avoir $Cx_0 < c$, car $\gamma_0 = 0$ par (ii). On a donc trouvé un point x_0 qui vérifie (i).

4 Ellipsoïde circonscrit de volume minimal

1. Soit $H \in \mathcal{S}^n$. En développant $c_y(X+H) = \|(X+H)y\|^2$, on obtient

$$c_y(X+H) = c_y(X) + y^\mathsf{T} HXy + y^\mathsf{T} XHy + \|Hy\|^2.$$

   Les termes linéaires en H forment la dérivée directionnelle $c_y'(X)\cdot H$, qui s'écrit aussi (on utilise le fait que $\operatorname{tr} AB = \operatorname{tr} BA$)

$$c_y'(X)\cdot H = \operatorname{tr}(Xyy^\mathsf{T} H + yy^\mathsf{T} XH) = \langle Xyy^\mathsf{T} + yy^\mathsf{T} X, H\rangle. \tag{10.9}$$

   On en déduit le résultat.

2. La matrice examinée s'écrit

$$M = \sum_{i=1}^m y^i (y^i)^\mathsf{T} = YY^\mathsf{T},$$

   où

$$Y := \begin{pmatrix} y^1 & \cdots & y^m \end{pmatrix}. \tag{10.10}$$

   On en déduit que, pour $v \in \mathbb{R}^n$, $v^\mathsf{T} Mv = \|Y^\mathsf{T} v\|^2 \geqslant 0$.

   Dès lors, M est définie positive si, et seulement si, Y est surjective, c'est-à-dire si, et seulement si, les vecteurs y^i engendrent $\mathbb{R}^n$.

3. (a) La matrice εI est dans $\mathscr{X}$ pour $\epsilon \in \,]0, \min_i \|y^i\|^{-1}]$.

   (b) Si $\mathscr{X}$ n'est pas borné, on peut trouver une suite de matrice $X_k \in \mathscr{X}$, telles que $\|X_k\| \to \infty$. On peut aussi supposer que $X_k/\|X_k\| \to X$ (on extrait une sous-suite au besoin; $\mathcal{S}^n$ est de dimension finie !). Bien sûr $\|X\| = 1$. D'autre part, en passant à la limite dans $\|X_k y^i\| \leqslant 1$, on obtient $Xy^i = 0$ pour tout i, ce qui s'écrit aussi $XY = 0$. La surjectivité de Y conduit alors à $X = 0$, en contradiction avec $\|X\| = 1$.

   (c) Soit $\mathscr{X}_0 := \{X \in \mathscr{X} : f(X) \leqslant f(X_0)\}$. Comme $\mathcal{S}^n$ est un espace métrique (c'est un espace euclidien !), il suffit de montrer que pour toute suite de matrices $\{X_k\} \subset \mathscr{X}_0$ telles que $X_k \to X$, on a $X \in \mathscr{X}_0$.

En passant à la limite dans $\|X_k y^i\| \leqslant 1$, on obtient $\|Xy^i\| \leqslant 1$. D'autre part, de l'inégalité $f(X_k) \leqslant f(X_0)$, on déduit que $\det X_k \geqslant C_0 := \exp(-f(X_0)) = \det X_0 > 0$, si bien que $\det X \geqslant C_0 > 0$ (par la continuité du déterminant). Les valeurs propres de X sont donc toutes strictement positives (par la continuité des valeurs propres on savait déjà aussi qu'elles étaient positives), donc $X \in \mathscr{X}$. On a aussi $f(X) \leqslant f(X_0)$, car f est continue sur $\mathscr{X}$. Donc $X \in \mathscr{X}_0$.

(d) D'après les points précédents, $\mathscr{X}_0$ est un compact non vide ($\mathcal{S}^n$ est de dimension finie) et f est continue sur $\mathscr{X}_0$; donc le problème $\min\{f(X) : X \in \mathscr{X}_0\}$ a au moins une solution $\bar{X}$. Celle-ci est évidemment aussi une solution de (10.2).

(e) L'unicité vient de ce que $\mathscr{X}$ est convexe (convexité de $\mathcal{S}^n_{++}$ et convexité de la norme ℓ_2) et de la stricte convexité de f sur $\mathscr{X}$ (pour tout $X \in \mathscr{X}$ et tout $H \in \mathcal{S}^n$ non nul, $f''(X) \cdot H^2 = \operatorname{tr}(X^{-1/2}HX^{-1/2})^2 > 0$ [formule vue]).

4. Il suffit de montrer qu'il existe $D \in \mathcal{S}^n$ telle que $c'_{y^i}(\bar{X}) \cdot D < 0$ pour les i vérifiant $\|\bar{X}y^i\| = 1$. D'après (10.9), on a

$$c'_{y^i}(\bar{X}) \cdot D = (y^i)^\mathsf{T} \bar{X} D y^i + (y^i)^\mathsf{T} D \bar{X} y^i.$$

On voit qu'il suffit de prendre $D = -\bar{X}$, puis qu'alors $c'_{y^i}(\bar{X}) \cdot D = -2\|\bar{X}y^i\|^2 < 0$ ($y^i \neq 0$ pour les indices considérés et $\bar{X}$ est inversible).

5. Soit λ_i le multiplicateur associé à la contrainte $\|Xy^i\|^2 \leqslant 1$, si bien que le lagrangien est défini sur $\mathcal{S}^n \times \mathbb{R}^m$ par

$$\ell(X, \lambda) = -\log\det X + \sum_{i=1}^m \lambda_i \left(\|Xy^i\|^2 - 1\right).$$

En une solution primale-duale $(\bar{X}, \bar{\lambda})$, on a

$$0 = \nabla_X \ell(\bar{X}, \bar{\lambda}) = -\bar{X}^{-1} + \sum_{i=1}^m \bar{\lambda}_i \left(\bar{X}y^i(y^i)^\mathsf{T} + y^i(y^i)^\mathsf{T}\bar{X}\right).$$

Les conditions de KKT s'écrivent alors

$$\begin{cases} \bar{X}^{-1} = \bar{X}Y\bar{\Lambda}Y^\mathsf{T} + Y\bar{\Lambda}Y^\mathsf{T}\bar{X} \\ \|\bar{X}y^i\| \leqslant 1 \quad \text{et} \quad \bar{\lambda}_i \geqslant 0, \quad \text{pour tout } i \\ \bar{\lambda}_i(\|\bar{X}y^i\| - 1) = 0, \end{cases} \tag{10.11}$$

où $\bar{\Lambda} = \operatorname{diag}(\bar{\lambda}_1, \ldots, \bar{\lambda}_m)$ et Y est définie en (10.10).

6. Si $m = n$, Y est carrée inversible et $S := Y^\mathsf{T} Y \succ 0$. On vérifie alors que la solution primale-duale s'écrit

$$\bar{X} = YS^{-3/2}Y^\mathsf{T} \qquad \text{et} \qquad \bar{\Lambda} = \frac{1}{2}I.$$

En effet,

- $YS^{-1}Y^\mathsf{T} = I$, car $S^{-1/2}Y^\mathsf{T}YS^{-1/2} = I$ et donc $S^{-1/2}Y^\mathsf{T}$ est l'inverse de $YS^{-1/2}$ et $I = (YS^{-1/2})(S^{-1/2}Y^\mathsf{T}) = YS^{-1}Y^\mathsf{T}$;

- $\bar{X}^{-1} = YS^{-1/2}Y^\mathsf{T}$, car $\bar{X}YS^{-1/2}Y^\mathsf{T} = YS^{-3/2}Y^\mathsf{T}YS^{-1/2}Y^\mathsf{T} = YS^{-1}Y^\mathsf{T} = I$;
- $\bar{X}Y\bar{\Lambda}Y^\mathsf{T} + Y\bar{\Lambda}Y^\mathsf{T}\bar{X} = \frac{1}{2}YS^{-1/2}Y^\mathsf{T} + \frac{1}{2}YS^{-1/2}Y^\mathsf{T} = \bar{X}^{-1}$, si bien que la première condition d'optimalité est vérifiée;
- $\bar{X}y^i = \bar{X}Ye^i = YS^{-1/2}e^i$, donc $\|\bar{X}y^i\|^2 = (e^i)^\mathsf{T}S^{-1/2}Y^\mathsf{T}YS^{-1/2}e^i = 1$ et les contraintes sont toutes actives.

Il n'y a pas d'autre solution primale, car les conditions de KKT sont suffisantes et que (10.2) n'a qu'une solution.

Remarque Ce problème a déjà été discuté par John [3 ; 1948] dans son travail sur les conditions d'optimalité et repris par de nombreux auteurs, dont Dyer [2 ; 1992]. En l'absence d'expression analytique de l'ellipsoïde de volume minimal, celui-ci pourra être calculé numériquement par divers algorithmes : un algorithme de points intérieurs qui trouve une ε-solution en $O(m^{3.5}\log(m/\varepsilon))$ opérations [4 ; 1996], une combinaison des méthodes de points intérieurs et d'activation de contraintes pour grands problèmes [5 ; 2004], l'algorithme de Frank et Wolfe [1 ; 2008]. Pour un état de l'art sur le sujet, voir [6 ; 2016].

Bibliographie

[1] S.D. Ahipasaoglu, P. Sun, M.J. Todd (2008). Linear convergence of a modified Frank-Wolfe algorithm for computing minimum-volume enclosing ellipsoids. *Optimization Methods and Software*, 23, 5–19.

[2] M.E. Dyer (1992). A class of convex programs with applications to computational geometry. In *Proceedings of the 8th ACM Symposium on Computational Geometry*, pages 9–15.

[3] F. John (1948). Extremum problems with inequalities as subsidiary conditions. In K.O. Friedrichs, O.E. Neugebauer, J.J. Stokes, éditeurs, *Studies and Essays, Courant Anniversary Volume*, pages 186–204. Wiley Interscience, New York.

[4] L.G. Khachiyan (1996). Rounding of polytopes in the real number model of computations. *Mathematics of Operations Research*, 21, 307–320.

[5] P. Sun, R.M. Freund (2004). Computation of minimum-volume covering ellipsoids. *Operations Research*, 52, 690–706.

[6] M.J. Todd (2016). *Minimum-Volume Ellipsoids – Theory and Algorithms.* MOS-SIAM Series on Optimization 23. SIAM and MPS, Philadelphia.