SOLUTIONS

1 Optimisation linéaire et dualité

1. (a) Au problème (6.1) nous associons le lagrangien $\ell : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n_+ \to \mathbb{R} \cup \{+\infty\}$ défini par

$$\ell(x, y, s) := c^\mathsf{T} x - y^\mathsf{T}(Ax - b) - s^\mathsf{T} x.$$

Si on prend la convention que la valeur minimale en (6.1) vaut $+\infty$ si le problème n'est pas réalisable, alors (6.1) est équivalent au problème

$$\inf_{x} \sup_{\substack{y \\ s \geqslant 0}} \ell(x, y, s). \tag{6.19}$$

Le problème *dual* de (6.1) est obtenu en inversant l'ordre entre maximisation et minimisation dans (6.19). La fonction $\theta : (y, s) \in \mathbb{R}^m \times \mathbb{R}^n_+ \mapsto \inf_x \ell(x, y, s)$ est la *fonction duale* :

$$\begin{aligned} \theta(y, s) &= \inf_{x \in \mathbb{R}^n} \ell(x, y, s) \\ &= b^\mathsf{T} y + \inf_{x \in \mathbb{R}^n} (c - A^\mathsf{T} y - s)^\mathsf{T} x \\ &= \begin{cases} b^\mathsf{T} y & \text{si } c - A^\mathsf{T} y - s = 0, \\ -\infty & \text{sinon.} \end{cases} \end{aligned}$$

Le problème dual (6.2) s'en déduit si l'on convient que la valeur maximale dans (6.2) vaut $-\infty$ si ce problème est non réalisable.

(b) Dans ce cas, le lagrangien s'écrit $\ell(x, y) := c^\mathsf{T} x - y^\mathsf{T}(Ax - b)$, $y \in \mathbb{R}^m$ et (6.1) devient

$$\inf_{x \geqslant 0} \sup_{y} \ell(x, y).$$

La fonction duale est définie sur $\mathbb{R}^m$ par

$$\begin{aligned} \theta(y) &= \inf_{x \geqslant 0} \ell(x, y) \\ &= b^\mathsf{T} y + \inf_{x \geqslant 0} (c - A^\mathsf{T} y)^\mathsf{T} x \\ &= \begin{cases} b^\mathsf{T} y & \text{si } A^\mathsf{T} y \leqslant c, \\ -\infty & \text{sinon.} \end{cases} \end{aligned}$$

(c) Le problème (6.1) peut s'écrire

$$\inf_{\substack{x \\ Ax = b}} \sup_{s \geqslant 0} c^\mathsf{T} x - s^\mathsf{T} x.$$

Le problème dual associé est donc

$$\sup_{s \geqslant 0} \inf_{\substack{x \\ Ax = b}} (c - s)^\mathsf{T} x.$$

Comme on a supposé $b \in \mathcal{R}(A)$, on a

$$\inf_{\substack{x \\ Ax = b}} (c - s)^\mathsf{T} x = \begin{cases} b^\mathsf{T} y & \text{si } c - s \in \mathcal{R}(A^\mathsf{T}) \text{ et } y \text{ est tel que } c - s = A^\mathsf{T} y \\ -\infty & \text{sinon.} \end{cases}$$

On en déduit que le problème dual est (6.2).

2. Le résultat s'obtient en écrivant le problème dual (6.2) sous forme standard

$$\begin{cases} \inf_{(u,v,s)} -b^\mathsf{T} u + b^\mathsf{T} v \\ \begin{pmatrix} -A^\mathsf{T} & A^\mathsf{T} & -I \end{pmatrix} \begin{pmatrix} u \\ v \\ s \end{pmatrix} = -c \\ (u,v,s) \geqslant 0 \end{cases}$$

où on a décomposé $y = u - v$. Le dual de ce problème n'est autre que (6.1).

3. Si x_0 et (y_0, s_0) sont admissibles pour (6.1) et (6.2) respectivement, on a

$$\begin{aligned} c^\mathsf{T} x_0 - b^\mathsf{T} y_0 &= c^\mathsf{T} x_0 - x_0^\mathsf{T} A^\mathsf{T} y_0 \qquad [Ax_0 = b] \\ &= x_0^\mathsf{T} s_0 \qquad [A^\mathsf{T} y_0 + s_0 = c] \\ &\geqslant 0 \qquad [x_0 \geqslant 0 \text{ et } s_0 \geqslant 0], \end{aligned}$$

ce qui démontre (6.3).

En partant de l'inégalité que l'on vient de démontrer, à savoir

$$\forall\, x_0 \in \mathcal{A}_{\mathrm{P}},\ \forall\, (y_0, s_0) \in \mathcal{A}_{\mathrm{D}} : \qquad b^\mathsf{T} y_0 \leqslant c^\mathsf{T} x_0,$$

où l'on a noté $\mathcal{A}_{\mathrm{P}}$ l'ensemble admissible primal et $\mathcal{A}_{\mathrm{D}}$ l'ensemble admissible dual, on obtient

$$\forall\, (y_0, s_0) \in \mathcal{A}_{\mathrm{D}} : \qquad b^\mathsf{T} y_0 \leqslant \mathrm{val}(P) = \inf_{\substack{x \\ Ax=b \\ x \geqslant 0}} c^\mathsf{T} x \leqslant c^\mathsf{T} x_0$$

puis

$$b^\mathsf{T} y_0 \leqslant \sup_{\substack{(y,s) \\ A^\mathsf{T} y + s = c \\ s \geqslant 0}} b^\mathsf{T} y = \mathrm{val}(D) \leqslant \mathrm{val}(P) = \inf_{\substack{x \\ Ax=b \\ x \geqslant 0}} c^\mathsf{T} x \leqslant c^\mathsf{T} x_0,$$

qui est (6.4). L'inégalité $\mathrm{val}(D) \leqslant \mathrm{val}(P)$ est l'*inégalité de dualité faible*, qui reste vraie dans le cadre beaucoup plus général de la dualité min-max (la dualité lagrangienne en est un cas particulier).

4.
   - $[(i) \Rightarrow (ii)]$ Si (i) a lieu, (P_L) est réalisable et borné (pour ce dernier point, on utilise l'inégalité de dualité faible et le fait que (D_L) est réalisable). Alors, (P_L) a une solution par (6.5).
   - $[(ii) \Rightarrow (iii)]$ Si (P_L) a une solution $\bar{x}$, il existe des multiplicateurs $(\bar{y}, \bar{s})$ tels que
     $$\begin{cases} A^\mathsf{T} \bar{y} + \bar{s} = c, \quad \bar{s} \geqslant 0 \\ A\bar{x} = b, \quad \bar{x} \geqslant 0 \\ \bar{x}^\mathsf{T} \bar{s} = 0. \end{cases} \tag{6.20}$$
     Ces conditions sont aussi les conditions d'optimalité de KKT de (D_L), qui a donc une solution (c'est un problème convexe).
   - $[(iii) \Rightarrow (i)]$ Alors (D_L) est clairement réalisable. Par ailleurs Les conditions d'optimalité de (D_L) ci-dessus montrent que (P_L) est aussi réalisable.
   - Supposons que $\bar{x}$ soit une solution de (P_L). Soit $(\bar{y}, \bar{s})$ des multiplicateurs optimaux associés, si bien que l'on a (6.20). Alors $(\bar{y}, \bar{s})$ est admissible pour (D) et (6.3) implique que $c^\mathsf{T} \bar{x} - b^\mathsf{T} \bar{y} = \bar{x}^\mathsf{T} \bar{s} = 0$ (troisième équation de (6.20)). On déduit alors de (6.4) (avec $x_0 = \bar{x}$ et $(y_0, s_0) = (\bar{y}, \bar{s})$) que $\mathrm{val}(D) = \mathrm{val}(P)$ (pas de saut de dualité) et que $(\bar{y}, \bar{s})$ est solution de (D_L).

2 Dualisation d'un problème quadratique

On utilise le lagrangien

$$\ell(x, y, s) = \frac{1}{2}x^\mathsf{T} Qx + c^\mathsf{T} x + y^\mathsf{T}(Ax - b) - s^\mathsf{T} x.$$

Comme le problème primal (P) s'écrit

$$\inf_x \sup_{\substack{y \\ s \geqslant 0}} \ell(x, y, s),$$

le dual est

$$\sup_{\substack{y \\ s \geqslant 0}} \inf_x \ell(x, y, s).$$

Le minimum en x du lagrangien ℓ s'obtient pour

$$x = -Q^{-1}(c + A^\mathsf{T} y - s).$$

On obtient alors le problème dual

$$(D) \quad \sup_{\substack{y \\ s \geqslant 0}} -\frac{1}{2}(c + A^\mathsf{T} y - s)Q^{-1}(c + A^\mathsf{T} y - s) - b^\mathsf{T} y.$$

Ce problème est un problème quadratique sous contraintes de borne, plus simple à résoudre que le problème original. En effet, en plus des contraintes de positivité sur les variables, (P) a également des contraintes d'égalité, ce qui complique assez bien le problème. Le problème dual n'a que des contraintes de positivité sur certaines variables.

3 Dualisation d'un problème linéaire avec contrainte conique

1. Le problème (6.6) s'écrit

$$\inf_{x \in K} \sup_{y \in \mathbb{F}} \langle c, x\rangle + \langle y, b - Ax\rangle.$$

Le problème dual est donc

$$\sup_{y \in \mathbb{F}} \inf_{x \in K} \langle c, x\rangle + \langle y, b - Ax\rangle = \sup_{y \in \mathbb{F}} \left(\langle b, y\rangle + \inf_{x \in K} \langle c - A^* y, x\rangle \right).$$

Il est clair que

$$\inf_{x \in K} \langle c - A^* y, x\rangle = \begin{cases} 0 & \text{si } c - A^* y \in K^+ \\ -\infty & \text{sinon.} \end{cases}$$

En effet, si $c - A^* y \in K^+$, l'infimum est $\geqslant 0$ et en prenant $x = tx_0$ avec $x_0 \in K \neq \emptyset$ et $t \downarrow 0$, on voit que l'infimum est nul; si $c - A^* y \notin K^+$, il existe un $x_0 \in K$ tel que $\langle c - A^*, x_0\rangle < 0$ et en prenant $x = tx_0$ avec $t \to +\infty$, on voit que l'infimum est $-\infty$. On en déduit que le dual s'écrit

$$\begin{cases} \sup_{y \in \mathbb{F}} \langle b, y\rangle \\ c - A^* y \in K^+. \end{cases}$$

Ce dernier problème est équivalent à (6.7).

2. (a) Il suffit d'appliquer le lemme de Farkas $\overline{A(K)} = \{y : A^*y \in K^+\}^+$, avec l'application linéaire $A = I_{\mathbb{E}}$ (l'identité de $\mathbb{E}$).

   (b) Comme $K^{++} = K$, on a

$$\sup_{s\in K^+} -\langle s, x\rangle = \begin{cases} 0 & \text{si } x \in K \\ +\infty & \text{sinon.} \end{cases}$$

   Dès lors, le problème (6.6) peut s'écrire

$$\inf_{x\in\mathbb{E}} \sup_{\substack{y\in\mathbb{F}\\ s\in K^+}} \left(\langle c, x\rangle - \langle y, Ax - b\rangle - \langle s, x\rangle\right).$$

   Celui-ci a pour dual

$$\begin{aligned} &\sup_{\substack{y\in\mathbb{F}\\ s\in K^+}} \inf_{x\in\mathbb{E}} \left(\langle c, x\rangle - \langle y, Ax - b\rangle - \langle s, x\rangle\right) \\ &= \sup_{\substack{y\in\mathbb{F}\\ s\in K^+}} \inf_{x\in\mathbb{E}} \left(\langle b, y\rangle + \langle c - A^*y - s, x\rangle\right) \\ &= \sup_{\substack{y\in\mathbb{F}\\ A^*y+s=c\\ s\in K^+}} \langle b, y\rangle. \end{aligned}$$

   On a donc retrouvé (6.7).

4 Dualisation d'une contrainte scalaire dans un problème à deux variables

1. Avec le lagrangien $\ell(x, \lambda) = x_2 + \lambda x_1$, le problème primal s'écrit

$$\inf_{x\in X} \sup_{\lambda\in\mathbb{R}} \ell(x, \lambda).$$

   Donc le dual est

$$(D) \quad \sup_{\lambda\in\mathbb{R}} \delta(\lambda),$$

   où la fonction duale s'écrit

$$\delta(\lambda) := \inf_{x\in X} (\lambda x_1 + x_2).$$

   Interprétons ce problème. On a $\lambda x_1 + x_2 = a_\lambda$ pour tout $x \in \mathbb{R}^2$ sur la droite de pente $-\lambda$ passant par $(0, a_\lambda)$. Dès lors, $(0, \delta(\lambda))$ est le point d'intersection avec l'axe des x_2 de la droite de pente $-\lambda$ *rencontrant* X et qui est *la plus basse possible* ou encore le point d'intersection avec l'axe des x_2 de *la plus haute* droite de pente $-\lambda$ qui est *en-dessous* de X (voir la figure 7). Le problème dual cherche à maximiser $\delta(\lambda)$, donc à déterminer la pente des droites qui sont « en-dessous » de X et qui rencontrent l'axe des x_2 le plus haut possible.

2. Supposons que X est un convexe contenant un point d'abscisse > 0 et un point d'abscisse < 0 et que le problème primal a une solution $\bar{x}$.

$$(D) \qquad \sup_{\lambda \in \mathbb{R}} \delta(\lambda)$$

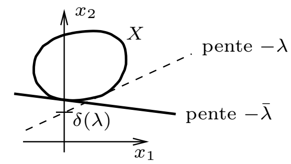


Fig. 7: Le problème dual et sa solution

(a) Montrons que le dual a une solution. Par optimalité, il est clair que $\bar{x} \in \partial X$, si bien qu'il existe une normale ν *non nulle* à X en $\bar{x}$: $\nu^{\mathsf{T}}(x - \bar{x}) \leqslant 0$ pour tout $x \in X$. Comme $\bar{x}_1 = 0$, on obtient

$$\forall\, (x_1, x_2) \in X : \qquad \nu_2 \bar{x}_2 \geqslant \nu_1 x_1 + \nu_2 x_2. \tag{6.21}$$

Montrons que l'on peut prendre $\nu_2 < 0$.

- D'abord $\nu_2 \neq 0$. En effet, dans le cas contraire on aurait pour tout $(x_1, x_2) \in X$: $\nu_1 x_1 \leqslant 0$. Comme X contient deux points x^- et x^+ avec $x_1^- < 0$ et $x_1^+ > 0$, on en déduirait que $\nu_1 = 0$, contredisant le fait que $\nu \neq 0$.
- Si l'on a égalité en (6.21) quel que soit $x \in X$, on peut changer le signe de ν_2 au besoin pour avoir $\nu_2 < 0$, sans changer l'affirmation (6.21).
- Dans le cas contraire, il existe $\tilde{x} \in X$ tel que l'on ait inégalité stricte en (6.21) :

$$\nu_2 \bar{x}_2 > \nu_1 \tilde{x}_1 + \nu_2 \tilde{x}_2.$$

  Alors $\nu_2 < 0$. En effet,
  - si $\tilde{x}_1 = 0$ c'est clair car $\bar{x}_2 \leqslant \tilde{x}_2$ par optimalité de $\bar{x}$;
  - si $\tilde{x}_1 > 0$, par convexité de X et le fait qu'il existe un $x^- \in X$ avec $x_1^- < 0$, on peut trouver un $x \in X$ avec $x_1 = 0$ et vérifiant l'inégalité de (6.21) strictement; on raisonne alors comme dans le premier cas ;
  - si $\tilde{x}_1 < 0$, le raisonnement est similaire au second cas.

Dès lors, en posant $\bar{\lambda} := \nu_1 / \nu_2$, on a

$$\forall\, (x_1, x_2) \in X : \qquad \bar{x}_2 \leqslant \bar{\lambda} x_1 + x_2.$$

Dès lors

$$\delta(\bar{\lambda}) = \inf_{x \in X} (\bar{\lambda} x_1 + x_2) = \bar{x}_2,$$

puisque l'infimum est atteint en $x = \bar{x} = (0, \bar{x}_2) \in X$. Enfin pour λ arbitraire dans $\mathbb{R}$, on a

$$\delta(\lambda) = \inf_{x \in X} \, (\lambda x_1 + x_2) \leqslant (\lambda x_1 + x_2)|_{x = \bar{x}} = \bar{x}_2 = \delta(\bar{\lambda}),$$

si bien que $\bar{\lambda}$ est solution du problème dual.

(b) Enfin, il n'y a pas de saut de dualité, car $\delta(\bar{\lambda}) = \bar{x}_2$, qui est la valeur optimale primale.

Comme $\bar{x}$ est solution primale, comme $\bar{\lambda}$ est solution duale et comme il n'y a pas de saut de dualité, $(\bar{x}, \bar{\lambda})$ est point-selle du lagrangien.

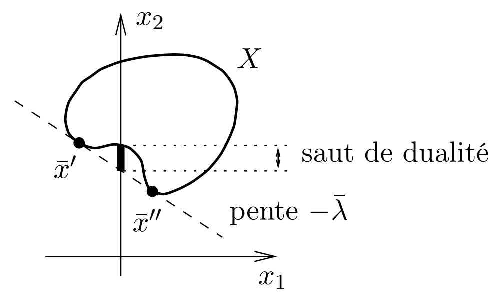


Fig. 8: Un cas avec saut de dualité

3. (a) Si X n'est pas convexe, il peut y avoir un saut de dualité, comme c'est le cas lorsque X est défini comme à la figure 8. Les solutions du problème de Lagrange (6.8) sont les deux points $\bar{x}'$ et $\bar{x}''$ qui ne sont pas solutions de (P). On notera que, dans cet exemple, l'ensemble admissible $\{x \in X : x_1 = 0\}$ est convexe et que le critère est convexe. On peut donc parler d'un « problème convexe » avec saut de dualité.

   (b) Considérons l'ensemble X avec méplat de la figure 9. Le lagrangien a un point-

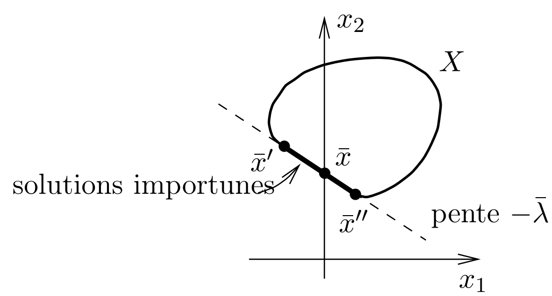


Fig. 9: Un cas sans saut de dualité, mais dont le problème de Lagrange a des solutions importunes

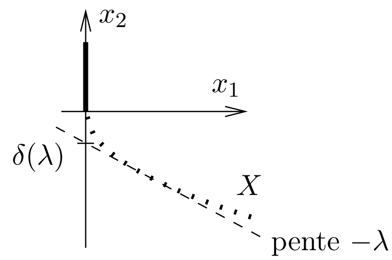
selle (car les problèmes primal et dual ont une solution, $\bar{x}$ et $\bar{\lambda}$ respectivement, et pas de saut de dualité), mais le problème de Lagrange (6.8) admet tout l'intervalle $[\bar{x}', \bar{x}'']$ comme solution. Celui-ci contient l'unique solution primale $\bar{x}$ (comme le prédit la théorie), mais aussi d'autres points qui sont donc des *solutions importunes*.

   (c) Soit $X = \{0\}$. Alors, pour tout $\lambda \in \mathbb{R}$, $\delta(\lambda) = 0$. Donc tout $\lambda \in \mathbb{R}$ est solution duale.

   Plus généralement, il en est de même si $X \subset \{x \in \mathbb{R}^2 : x_1 = 0,\ x_2 \geqslant b\}$ pour un réel b donné.

   (d) Considérons le cas où X est défini par la figure 10.

      i. Clairement, le problème primal n'a pas de solution (en prenant $x_2 \geqslant -\sqrt{x_1}$ dans la définition de X, il en aurait une). La valeur optimale primale est nulle.



$$X = \{x \in \mathbb{R}^2 : x_1 \geqslant 0,\ x_2 > -\sqrt{x_1}\}$$

Fig. 10: Un cas sans solution primale-duale et sans saut de dualité

ii. Par continuité du lagrangien, la fonction duale s'écrit

$$\delta(\lambda) = \inf_{\substack{x_1 \geqslant 0 \\ x_2 \geqslant -\sqrt{x_1}}} (\lambda x_1 + x_2) = \inf_{x_1 \geqslant 0} (\lambda x_1 - \sqrt{x_1}).$$

Si $\lambda \leqslant 0$, on a $\delta(\lambda) = -\infty$. Si $\lambda > 0$, le problème de Lagrange est résolu avec

$$\bar{x}_\lambda = \left(\frac{1}{4\lambda^2}, -\frac{1}{2\lambda}\right).$$

On a donc montré que

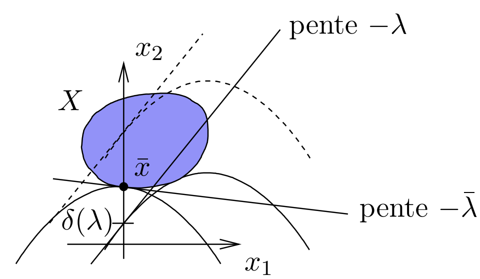
$$\delta(\lambda) = \begin{cases} -\infty & \text{si } \lambda \leqslant 0 \\ -\frac{1}{4\lambda} & \text{si } \lambda > 0. \end{cases}$$

Dès lors le problème dual n'a pas de solution (sa valeur optimale nulle n'est pas atteinte).

iii. Comme les valeurs optimales primale et duale sont toutes les deux nulles, il n'y a pas de saut de dualité.

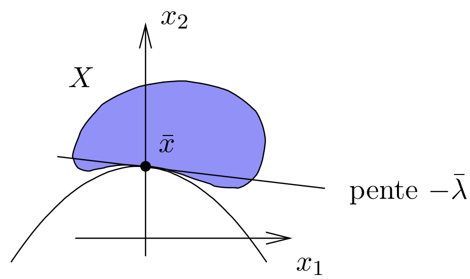
4. (a) Le problème primal s'écrit aussi $\inf_{x \in X} \sup_{\lambda \in \mathbb{R}} \ell_r(x, \lambda)$, où $\ell_r(x, \lambda) = \lambda x_1 + x_2 + \frac{r}{2} x_1^2$. Le dual associé est donc

$$(D) \quad \sup_{\lambda \in \mathbb{R}} \inf_{x \in X} \lambda x_1 + x_2 + \frac{r}{2} x_1^2$$



Le lagrangien augmenté vaut la constante a le long de la parabole concave passant par $(0, a)$, de pente $-\lambda$ en ce point et de courbure $-r$. Dès lors, la valeur de la fonction duale $\delta_r(\lambda) = \inf_{x \in X} \lambda x_1 + x_2 + \frac{r}{2} x_1^2$ est la borne inférieure des a tels que la parabole $x_1 \mapsto -\frac{r}{2} x_1^2 - \lambda x_1 + a$ intersecte X. Le problème dual consiste à maximiser $\delta_r(\lambda)$ en λ, donc à trouver une parabole de courbure $-r$, qui soit sous X, en contact avec X et qui intersecte l'axe des ordonnées le plus haut possible.

(b) Cette dualisation ne donne pas de saut de dualité pour des ensembles qui ne sont pas nécessairement convexes, comme dans l'exemple ci-dessous, alors qu'il n'en serait pas de même pour la dualisation par le lagrangien classique.



5 Identité du minimax de von Neumann

1. Le premier problème consiste à maximiser une fonction continue (linéaire même) sur un compact non vide (le simplexe). Il a donc une solution.

   Le second problème s'écrit

   $$\inf_{x\in\Delta_n} \varphi(x), \qquad \text{où} \qquad \varphi(x) = \sup_{y\in\Delta_m} y^\mathsf{T} Ax.$$

   Comme Δ_n est un compact non vide, il suffit de montrer que φ est semi-continue inférieurement ; ce qui est bien le cas, car φ est l'enveloppe supérieure de fonctions linéaires (φ est donc aussi convexe et comme elle ne prend que des valeurs finies, elle est continue).

2. Soit $\bar{v} := \min\{v_i : 1 \leqslant i \leqslant n\}$. Si $x \in \Delta_n$, on a

   $$\begin{array}{rcl} v^\mathsf{T} x & \geqslant & (\bar{v}e)^\mathsf{T} x \qquad [x \geqslant 0] \\ & = & \bar{v} \qquad [e^\mathsf{T} x = 1]. \end{array}$$

   Par ailleurs si l'indice i est tel que $v_i = \bar{v}$, on a $e^i \in \Delta_n$ et $v^\mathsf{T} e^i = \bar{v}$.

3. (a) On a les équivalences suivantes :

   $$\begin{array}{ll} t \leqslant \alpha & \\ \iff & \exists\, y \in \Delta_m : \min\{y^\mathsf{T} Ax : x \in \Delta_n\} \geqslant t \qquad [\text{définition de } \alpha \text{ et point 1}] \\ \iff & \exists\, y \in \mathbb{R}^m : y \geqslant 0,\ e^\mathsf{T} y = 1,\ A^\mathsf{T} y \geqslant te \qquad [\text{point 2}] \\ \iff & \exists\, y \in \mathbb{R}^m : y \geqslant 0,\ e^\mathsf{T} y = 1/t,\ A^\mathsf{T} y \geqslant e \qquad [y \curvearrowright ty \text{ et } t > 0] \\ \iff & \inf\{e^\mathsf{T} y : y \geqslant 0,\ A^\mathsf{T} y \geqslant e\} \leqslant 1/t. \end{array} \tag{6.22}$$

   La dernière implication $\Leftarrow$ mérifte quelques explications. D'abord, on peut trouver un $y_1 \geqslant 0$ tel que $A^\mathsf{T} y_1 \geqslant e$ et $e^\mathsf{T} y_1 \leqslant 1/t$ (parce que le problème linéaire dans (6.22) est réalisable et borné et a donc une solution). Alors, comme $y_1 \neq 0$, on peut trouver un $s \geqslant 1$ tel que $y_s := sy_1$ vérifie $y_s \geqslant 0$, $A^\mathsf{T} y_s \geqslant e$ et $e^\mathsf{T} y_s = 1/t$.

   (b) En prenant $t := \alpha$ (égalité à gauche), on obtient $\inf\{e^\mathsf{T} y : y \geqslant 0,\ A^\mathsf{T} y \geqslant e\} \leqslant 1/\alpha$ par l'implication $\Rightarrow$ du point 3a.

   En prenant $t := 1/\inf\{e^\mathsf{T} y : y \geqslant 0,\ A^\mathsf{T} y \geqslant e\}$ (égalité à droite), on obtient $\inf\{e^\mathsf{T} y : y \geqslant 0,\ A^\mathsf{T} y \geqslant e\} \geqslant 1/\alpha$ par l'implication $\Leftarrow$ du point 3a.

(c) C'est un problème d'optimisation linéaire avec une valeur optimale finie.

4. On conclut par dualité forte (il n'y a pas de saut de dualité, car les problèmes linéaires ont une solution) :

$$
\begin{aligned}
\frac{1}{\alpha} &= \min_{\substack{y \geqslant 0 \\ A^{\mathsf{T}} y \geqslant e}} e^{\mathsf{T}} y \\
&= \inf_{y \geqslant 0} \sup_{x \geqslant 0} e^{\mathsf{T}} y + x^{\mathsf{T}}(e - A^{\mathsf{T}} y) \\
&= \sup_{x \geqslant 0} \inf_{y \geqslant 0} e^{\mathsf{T}} x + y^{\mathsf{T}}(e - Ax) \qquad \text{[pas de saut de dualité]} \\
&= \max_{\substack{x \geqslant 0 \\ Ax \leqslant e}} e^{\mathsf{T}} x \\
&= \frac{1}{\beta}.
\end{aligned}
$$

Donc $\alpha = \beta$, ce que l'on voulait démontrer.

6 Méthodes de décomposition par les prix et les ressources

1. (a) Le lagrangien du problème s'écrit pour $x \in X$ et $\lambda \in \mathbb{R}^m$:

$$
\ell(x, \lambda) = \sum_{j=1}^{N} f^j(x^j) + \lambda^{\mathsf{T}} \left(\sum_{j=1}^{N} c^j(x^j) \right) = \sum_{j=1}^{N} \Big(f^j(x^j) + \lambda^{\mathsf{T}} c^j(x^j) \Big) .
$$

Dès lors, le problème de Lagrange

$$
\delta(\lambda) := \inf_{x \in X} \ell(x.\lambda) = \sum_{j=1}^{N} \left(\inf_{x^j \in X^j} \Big[f^j(x^j) + \lambda^{\mathsf{T}} c^j(x^j) \Big] \right) .
$$

se décompose en N minimisations indépendantes :

$$
\inf_{x^j \in X^j} \Big[f^j(x^j) + \lambda^{\mathsf{T}} c^j(x^j) \Big], \quad \text{pour } i = 1, \ldots, N. \tag{6.23}
$$

L'algorithme d'Uzawa fait ensuite un pas $\alpha > 0$ en λ de manière à maximiser la fonction duale $\delta(\lambda)$:

$$
\lambda_+ := (\lambda + \alpha c(x))^+ . \tag{6.24}
$$

(b) Dans le cadre de l'exemple cité, l'approche peut s'interpréter comme suit.

Étape de décomposition. La direction de l'entreprise donne à chaque usine son estimation λ du prix des ressources (λ_i est le prix estimé de la ressource i). Chaque usine j détermine alors par (6.23) la quantité x^j à produire de manière à minimiser le coût total de sa production $f^j(x^j) + \lambda^{\mathsf{T}} c^j(x^j)$ (coût propre plus coût des ressources).

Étape de coordination. Au vu des quantités de ressource $c^j(x^j)$ utilisées par chaque usine, la direction adapte par (6.24) le prix λ des ressources de manière réaliser les contraintes portant sur ces ressources.

Cette approche sera d'autant plus utile que les X^j sont « petits »(par exemple des espaces vectoriels de faible dimension ou des ensembles formés de peu de points) et N est grand. Les minimisations dans (6.23) sont alors faciles à réaliser. On notera que le travail de la direction se résume à (6.24) qui ne présente pas de difficulté ....

2. (a) On a clairement

$$\left\{x \in X : \sum_j c^j(x^j) \leqslant 0\right\} = \bigcup_{\substack{p=(p^1,\dots,p^N)\\ \sum_{j=1}^N p^j \geqslant 0}} \left\{x \in X : c^j(x^j) + p^j \leqslant 0,\ \forall j\right\}.$$

Dès lors

$$\begin{aligned} \inf_{c(x)\leqslant 0} f(x) &= \inf_{p:\ \sum_j p^j \geqslant 0} \left(\inf_{x\in X:\ c^j(x^j)+p^j\leqslant 0} \sum_j f^j(x^j)\right) \\ &= \inf_{p:\ \sum_j p^j \geqslant 0} \sum_j \left(\inf_{x^j\in X^j:\ c^j(x^j)+p^j\leqslant 0} f^j(x^j)\right). \end{aligned}$$

D'où le résultat.

On a donc transformé le grand problème (6.10)–(6.12) en N petits problèmes (6.14) avec contraintes, plus faciles à résoudre, et un problème maître (6.13) avec contraintes linéaires, qui est en général non différentiable (les v^j n'ont en effet aucune raison d'être différentiables).

(b) i. Laissons tomber les indices j. Soient $p_1, p_2 \in \mathbb{R}^m$ et $t \in]0,1[$. Il faut montrer que

$$v((1-t)p_1 + tp_2) \leqslant (1-t)v(p_1) + tv(p_2).$$

On peut supposer que $v(p_1)$ et $v(p_2)$ sont finis (sinon le résultat est évident). Soit $\epsilon > 0$. On peut alors trouver x_i $(i = 1,2)$ tels que $c(x_i) + p_i \leqslant 0$ et $f(x_i) \leqslant v(p_i) + \epsilon$. Par convexité de c,

$$c((1-t)x_1 + tx_2) \leqslant (1-t)c(x_1) + tc(x_2) \leqslant -\Big((1-t)p_1 + tp_2\Big),$$

si bien que $(1-t)x_1 + tx_2$ est admissible pour le problème (6.14) avec la perturbation $(1-t)p_1 + tp_2$. Alors en utilisant la convexité de f,

$$\begin{aligned} v((1-t)p_1 + tp_2) &\leqslant f((1-t)x_1 + tx_2) \\ &\leqslant (1-t)f(x_1) + tf(x_2) \\ &\leqslant (1-t)v(p_1) + tv(p_2) + \epsilon. \end{aligned}$$

Comme $\epsilon > 0$ est arbitraire, on a le résultat.

ii. Laissons tomber les indices j. Soit (x,μ) une solution primale-duale de (6.14), correspondant à la perturbation p. Soit q une autre perturbation.

On a

$$
\begin{array}{rcll}
v(q) & = & \inf_{\tilde{x}\in X} \sup_{\tilde{\mu}\geqslant 0} \Big(f(\tilde{x}) + \tilde{\mu}^{\mathsf{T}}(c(\tilde{x})+q) \Big) & \text{[formulation infsup de (6.14)]} \\
& \geqslant & \sup_{\tilde{\mu}\geqslant 0} \inf_{\tilde{x}\in X} \Big(f(\tilde{x}) + \tilde{\mu}^{\mathsf{T}}(c(\tilde{x})+q) \Big) & \text{[dualité faible]} \\
& \geqslant & \inf_{\tilde{x}\in X} \Big(f(\tilde{x}) + \mu^{\mathsf{T}}(c(\tilde{x})+q) \Big) & [\mu \geqslant 0] \\
& = & \inf_{\tilde{x}\in X} \Big(f(\tilde{x}) + \mu^{\mathsf{T}}(c(\tilde{x})+p) \Big) + \mu^{\mathsf{T}}(q-p) & \\
& = & v(p) + \mu^{\mathsf{T}}(q-p) \quad [(x,\mu) \text{ est un point-selle}]. &
\end{array}
$$

iii. En notant $v(p) = v(p^1, \ldots, p^N)$ le critère de (6.13) et en sommant les inégalités obtenues au point ii, on a

$$
\begin{array}{rcl}
v(q) & = & \sum_{j=1}^{N} v^j(q^j) \\
& \geqslant & \sum_{j=1}^{N} v^j(p^j) + \sum_{j=1}^{N} (\mu^j)^{\mathsf{T}}(q^j - p^j) \\
& = & v(p) + \mu^{\mathsf{T}}(q-p),
\end{array}
$$

où $\mu = (\mu^1, \ldots, \mu^N)$. Ce vecteur μ est donc un sous-gradient de v en p.

(c) Dans le cadre de l'exemple cité, l'approche peut s'interpréter comme suit.

Étape de décomposition. La direction envoie aux usines des consignes de ressources à ne pas dépasser : $-p^j$. Les usines déterminent les quantités à produire en minimisant leurs coûts propres tout en respectant les contraintes imposées; c'est le problème (6.14).

Étape de coordination. La direction se charge de résoudre (6.13). Elle utilise pour cela des valeurs $v^j(p^j)$ et des multiplicateurs μ^j envoyés par chaque usine. Elle dispose ainsi d'un sous-gradient $\mu = (\mu^1, \ldots, \mu^N)$ du critère de (6.13).

Remarque.

La décomposition par les ressources est en général plus délicate à mettre en œuvre que la relaxation lagrangienne et est, à notre connaissance, moins utilisée en pratique que cette dernière. Il est en effet parfois difficile de s'assurer que les consignes de ressources à ne pas dépasser $-p^j$ n'entraînent pas la vacuité des ensembles admissibles des problèmes (6.14).

Elle a son intérêt lorsque rien ne garantit que les problèmes (6.23) ont une solution. Par exemple, en optimisation linéaire, les X^j sont des espaces vectoriels et les f^j et c^j sont affines, si bien que (6.23) n'a pas de solution en général. Grâce à ses contraintes d'inégalité, le problème (6.14) est alors mieux adapté. Cela dit, la relaxation lagrangienne s'utilise également en optimisation linéaire, mais avec des ensembles X^j qui sont des parties d'espaces vectoriels, définies par des contraintes d'inégalité, de sorte que les problèmes (6.23) soient bien posés.

Pour en apprendre davantage sur la relaxation lagrangienne, on pourra consulter [**?** ; 1996] et [**?** ; 2006, partie II].

7 Approximation de Tchebychev d'un système linéaire surdéterminé

1. Le plus simple est de voir (6.15) comme le problème $\inf\{\|y-b\|_\infty : y \in \mathcal{R}(A)\}$, dans lequel on « projette » le point b sur $\mathcal{R}(A)$. Il ne s'agit pas d'une projection au sens habituel, car la norme $\|\cdot\|_\infty$ ne dérive pas d'un produit scalaire. Cette projection existe (mais elle n'est pas nécessairement unique) puisque le critère est coercif et que l'ensemble admissible est fermé non vide (corollaire 1.4).

2. On utilise l'équivalence des normes dans $\mathbb{R}^m$:

$$\forall y \in \mathbb{R}^m : \quad \frac{1}{\sqrt{m}}\,\|y\|_2 \leqslant \|y\|_\infty \leqslant \|y\|_2. \tag{6.25}$$

Alors

$$\begin{aligned}
\|r_{\rm mc}\|_\infty &\leqslant \|r_{\rm mc}\|_2 && [(6.25)]\\
&\leqslant \|Ax_{\rm tch}-b\|_2 && [\text{optimalité du problème de MCL}]\\
&\leqslant \sqrt{m}\,\|Ax_{\rm tch}-b\|_\infty && [(6.25)]\\
&= \sqrt{m}\,v_{\rm tch}.
\end{aligned}$$

3. Voici deux techniques de dualisation possibles pour ce problème sans contrainte.

- *Première technique* (on écrit (6.15) comme un infsup sans dualisation de contrainte). La norme ℓ_∞ est la norme duale de la norme ℓ_1, si bien que le problème (6.15) s'écrit

$$\inf_{x\in\mathbb{R}^n}\ \sup_{\substack{y\in\mathbb{R}^m\\ \|y\|_1\leqslant 1}}\ y^{\mathsf T}(b-Ax).$$

On peut donc prendre comme dual minmax

$$\sup_{\substack{y\in\mathbb{R}^m\\ \|y\|_1\leqslant 1}}\ \inf_{x\in\mathbb{R}^n}\ y^{\mathsf T}(b-Ax) = \sup_{\substack{y\in\mathbb{R}^m\\ \|y\|_1\leqslant 1}} \left\{\begin{array}{ll} b^{\mathsf T}y & \text{si } A^{\mathsf T}y=0\\ -\infty & \text{sinon,}\end{array}\right.$$

qui est bien (6.17).

- *Deuxième technique* (plus compliquée, mais plus générale dans le sens où elle fait apparaître une contrainte que l'on dualise). On récrit d'abord le problème (6.15) comme le problème avec contrainte

$$\left\{\begin{array}{l} \inf_{(x,z)\in\mathbb{R}^n\times\mathbb{R}^m}\ \|z\|_\infty\\ b-Ax=z.\end{array}\right.$$

La dualisation lagrangienne de la contrainte permet alors d'écrire le problème (6.15) comme un infsup :

$$\inf_{(x,z)\in\mathbb{R}^n\times\mathbb{R}^m}\ \sup_{y\in\mathbb{R}^m}\ \|z\|_\infty + y^{\mathsf T}(b-Ax-z).$$

Le dual de celui-ci s'écrit alors

$$\begin{aligned}
&\sup_{y\in\mathbb{R}^m}\ \inf_{(x,z)\in\mathbb{R}^n\times\mathbb{R}^m}\ \|z\|_\infty + y^{\mathsf T}(b-Ax-z)\\
&= \sup_{y\in\mathbb{R}^m}\Big(b^{\mathsf T}y + \inf_{x\in\mathbb{R}^n}\,(-y^{\mathsf T}Ax) + \inf_{z\in\mathbb{R}^m}\,(\|z\|_\infty - y^{\mathsf T}z)\Big).
\end{aligned} \tag{6.26}$$

D'une part, on a

$$\inf_{x \in \mathbb{R}^n} (-y^\mathsf{T} A x) = \begin{cases} 0 & \text{si } A^\mathsf{T} y = 0 \\ -\infty & \text{sinon.} \end{cases}$$

D'autre part, il est classique que la conjuguée d'une norme est l'indicatrice de la boule unité duale (voir les exercices sur la conjugaison) :

$$\inf_{z \in \mathbb{R}^m} (\|z\|_\infty - y^\mathsf{T} z) = - \sup_{z \in \mathbb{R}^m} (y^\mathsf{T} z - \|z\|_\infty) = \begin{cases} 0 & \text{si } \|y\|_1 \leqslant 1 \\ -\infty & \text{sinon.} \end{cases}$$

En utilisant le résultat de ces deux calculs dans (6.26), on obtient (6.17).

Il est clair que (6.17) a une solution, car le critère est continu et l'ensemble admissible est compact (grâce à la contrainte $\|y\|_1 \leqslant 1$) non vide ($y = 0$ lui appartient).

4. Le point $-r_{\text{mc}}/\|r_{\text{mc}}\|_1$ est admissible pour (6.17), car $A^\mathsf{T} r_{\text{mc}} = 0$ (c'est l'équation normale du problème de moindres-carrés linéaire). Alors

$$\begin{aligned} \frac{\|r_{\text{mc}}\|_2^2}{\|r_{\text{mc}}\|_1} &= \frac{-b^\mathsf{T} r_{\text{mc}}}{\|r_{\text{mc}}\|_1} \qquad [A^\mathsf{T} r_{\text{mc}} = 0] \\ &\leqslant \operatorname{val}((6.17)) \qquad [\text{optimalité dans } (6.17)] \\ &\leqslant v_{\text{tch}} \qquad [\text{dualité faible}]. \end{aligned}$$

Pour montrer que

$$\frac{1}{\sqrt{m}} \|r_{\text{mc}}\|_\infty \leqslant \frac{\|r_{\text{mc}}\|_2^2}{\|r_{\text{mc}}\|_1}$$

on utilise le fait que, quel que soit $y \in \mathbb{R}^m$, on a $\|y\|_\infty \leqslant \|y\|_2$ (c'est clair) et $\|y\|_1 \leqslant \sqrt{m}\, \|y\|_2$ (par Cauchy-Schwarz).

8 Projection par relaxation lagrangienne

1. Le problème (6.18) s'écrit

$$\inf_{x \in \mathbb{R}^2} \sup_{\substack{\lambda \in \mathbb{R}^2 \\ \lambda_2 \geqslant 0}} \ell(x, \lambda),$$

où

$$\ell(x, y) = x_1^2 + x_2^2 + \lambda_1(x_1 + x_2 - 1) - \lambda_2 x_2.$$

Le problème dual associé au lagrangien ℓ s'écrit donc

$$\sup_{\substack{\lambda \in \mathbb{R}^2 \\ \lambda_2 \geqslant 0}} \inf_{x \in \mathbb{R}^2} \ell(x, \lambda). \tag{6.27}$$

Comme le problème est convexe et que les contraintes sont qualifiées (en toute solution), il n'y a pas de saut de dualité et on peut résoudre (6.18) en résolvant (6.27). On note

$$\Lambda = \{\lambda \in \mathbb{R}^2 : \lambda_2 \geqslant 0\}.$$

2. Algorithme d'Uzawa. Soit λ^k donné dans Λ. Alors on calcule d'abord x^k solution du problème de Lagrange

$$\inf_{x \in \mathbb{R}^2} \ell(x, \lambda^k),$$

ce qui donne

$$\begin{cases} 2x_1 + \lambda_1 = 0, \\ 2x_2 + \lambda_1 - \lambda_2 = 0, \end{cases}$$

soit

$$x^k = \begin{pmatrix} -\frac{\lambda_1^k}{2} \\ \frac{\lambda_2^k - \lambda_1^k}{2} \end{pmatrix}.$$

Ensuite, on met à jour λ^k par la formule ($\alpha^k > 0$ est un pas)

$$\lambda^{k+1} = P_\Lambda(\lambda^k + \alpha^k c(x^k)),$$

où P_Λ est le projecteur orthogonal sur le convexe Λ, c'est-à-dire $P_\Lambda(\lambda_1, \lambda_2) = (\lambda_1, \lambda_2^+)$. On obtient

$$\begin{aligned} \lambda_1^{k+1} &= \lambda_1^k + \alpha_k(x_1^k + x_2^k - 1) \\ &= (1-\alpha_k)\lambda_1^k + \alpha_k\left(\frac{\lambda_2^k}{2} - 1\right), \\ \lambda_2^{k+1} &= (\lambda_2^k - \alpha_k x_2^k)^+ \\ &= \left(\frac{\alpha_k}{2}\lambda_1^k + \left(1 - \frac{\alpha_k}{2}\right)\lambda_2^k\right)^+. \end{aligned}$$

Si on part de $\lambda^1 = (0,0)$, on trouve

$$\lambda^2 = (-\alpha_1, 0).$$

On en déduit que pour tout $k \geqslant 2$

$$\lambda_1^k = \lambda_1^{k-1} - \alpha_{k-1}(\lambda_1^{k-1} + 1), \quad \lambda_2^k = 0,$$
$$x_1^k = x_2^k = -\frac{\lambda_1^k}{2}.$$

Comme $(\lambda_1^k + 1) = (1 - \alpha_{k-1})(\lambda_1^{k-1} + 1)$, on a $\lambda_k \to (-1, 0)$ si α_k est pris dans un compact de $]0,1[$, tandis que $x^k \to \left(\frac{1}{2}, \frac{1}{2}\right)$. Le point limite est bien solution primale-duale du problème.

3. L'algorithme d'Arrow-Hurwicz s'écrit en un itéré (x^k, λ^k) :

$$\begin{aligned} x^{k+1} &= x^k - \alpha_k^1 \nabla_x \ell(x^k, \lambda^k) \\ \lambda^{k+1} &= P_\Lambda(\lambda^k + \alpha_k^2 \nabla_x \ell(x^{k+1}, \lambda^k)), \end{aligned}$$

où α_k^1 et α_k^2 sont pris dans un compact de $]0,1[$. On a

$$\begin{aligned} x_1^{k+1} &= x_1^k - \alpha_k^1(2x_1^k + \lambda_1^k) \\ &= (1 - 2\alpha_k^1)x_1^k - \alpha_k^1\lambda_1^k \\ x_2^{k+1} &= x_2^k - \alpha_k^1(2x_2^k + \lambda_1^k - \lambda_2^k) \\ &= (1 - 2\alpha_k^1)x_2^k - \alpha_k^1(\lambda_1^k - \lambda_2^k) \\ \lambda_1^{k+1} &= \lambda_1^k + \alpha_k^2(x_1^{k+1} + x_2^{k+1} - 1) \\ \lambda_2^{k+1} &= (\lambda_2^k - \alpha_k^2 x_2^{k+1})^+. \end{aligned}$$

Bibliographie

[1] S. Boyd, L. Vandenberghe (2004). *Convex Optimization*. Cambridge University Press.

[2] L.G. Khachiyan, M.J. Todd (1993). On the complexity of approximating the maximal ellipsoid for a polytope. *Mathematical Programming*, 61, 137–159.