Master 1 MAPI3 - IMA - RO

Travaux dirigés d'optimisation

PIERRE MARÉCHAL

Université Paul Sabatier

pr.marechal@gmail.com

Exercice

On considère le problème d'optimisation

$$(P_1) \quad \begin{cases} \text{Min} & f(x,y) \\ \text{s.c.} & h(x,y) = 0 \end{cases}$$

avec $f(x,y) = y - x^2$ et $h(x,y) = x^2 + y^2 - 1$.

(1) Tracer l'ensemble des points réalisables, ainsi que quelques courbes de niveau de la fonction f.

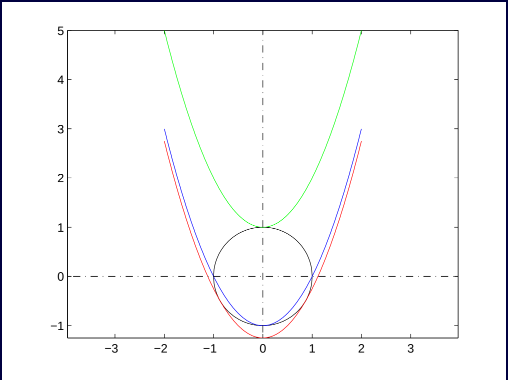
(2) Montrer que le problème (P_1) admet au moins une solution.

(3) Étudier la qualification de la contrainte.

(4) Trouver l'ensemble Ω des points qui vérifient les conditions nécessaires d'optimalité d'ordre 1 de (P_1).

(5) Étudier la nature de chacun de ces points et en déduire l'ensemble des solutions de (P_1).

(1) Sur la figure suivante, on a tracé en trait plein noir l'ensemble des points réalisables, en vert la courbe de niveau 1 de f, en bleu la courbe de niveau -1 de f et en rouge la courbe de niveau $-5/4$ de f.



(2) La fonction f est continue et l'ensemble réalisable est clairement borné et fermé (image réciproque d'un singleton par l'application continue h).

(2) La fonction f est continue et l'ensemble réalisable est clairement borné et fermé (image réciproque d'un singleton par l'application continue h). Le problème (P_1) admet donc au moins une solution.

(2) La fonction f est continue et l'ensemble réalisable est clairement borné et fermé (image réciproque d'un singleton par l'application continue h). Le problème (P_1) admet donc au moins une solution.

(3) Nous allons montrer que la contrainte h est qualifiée en tout point réalisable.

(2) La fonction f est continue et l'ensemble réalisable est clairement borné et fermé (image réciproque d'un singleton par l'application continue h). Le problème (P_1) admet donc au moins une solution.

(3) Nous allons montrer que la contrainte h est qualifiée en tout point réalisable. Pour cela, nous allons montrer que le gradient de h est linéairement indépendant (c'est-à-dire, jamais le vecteur nul) en un point réalisable.

(2) La fonction f est continue et l'ensemble réalisable est clairement borné et fermé (image réciproque d'un singleton par l'application continue h). Le problème (P_1) admet donc au moins une solution.

(3) Nous allons montrer que la contrainte h est qualifiée en tout point réalisable. Pour cela, nous allons montrer que le gradient de h est linéairement indépendant (c'est-à-dire, jamais le vecteur nul) en un point réalisable. Comme $\nabla h(x,y) = 2(x,y)^\top$, il s'annule en l'unique point $(0,0)$, qui n'est pas réalisable.

(4) Notant $\lambda \in \mathbb{R}$ le multiplicateur de Lagrange associé à l'unique contrainte d'égalité et

$$L(x,y;\lambda) = f(x,y) + \lambda h(x,y)$$

le lagrangien associé au problème (P_1), les conditions de Karush, Kuhn et Tucker (KKT) se traduisent ici par

$$\begin{aligned} x^2+y^2-1 &= 0 \qquad (1)\\ -2x+2\lambda x &= 0 \qquad (2)\\ 1+2\lambda y &= 0. \qquad (3) \end{aligned}$$

L'équation (1) représente la réalisabilité du point considéré et les équations (2) et (3) correspondent au fait que le gradient du lagrangien y est nul.

Comme (2) est équivalent à $x(\lambda-1)=0$, on a 2 cas à traiter:

- ou bien $x=0$ et dans ce cas (1) donne $y=\pm 1$ et (3) donne $\lambda=-1/(2y)=\mp 1/2$;
- ou bien $\lambda=1$ et dans ce cas (3) donne $y=-1/2$ et (1) donne

$$x=\pm\sqrt{1-\frac{1}{4}}=\pm\frac{\sqrt{3}}{2}.$$

Comme (2) est équivalent à $x(\lambda - 1) = 0$, on a 2 cas à traiter:

- ou bien $x = 0$ et dans ce cas (1) donne $y = \pm 1$ et (3) donne $\lambda = -1/(2y) = \mp 1/2$;
- ou bien $\lambda = 1$ et dans ce cas (3) donne $y = -1/2$ et (1) donne

$$x = \pm\sqrt{1 - \frac{1}{4}} = \pm\frac{\sqrt{3}}{2}.$$

On obtient donc

$$\Omega = \left\{ (0,1), (0,-1), \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right), \left(-\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) \right\}.$$

(5) Pour étudier la nature d'un point (x, y) de Ω, on cherche à comparer la valeur de f en ce point à la valeur de f aux points réalisables voisins.

(5) Pour étudier la nature d'un point (x,y) de Ω, on cherche à comparer la valeur de f en ce point à la valeur de f aux points réalisables voisins. Ces derniers sont nécessairement de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$.

(5) Pour étudier la nature d'un point (x,y) de Ω, on cherche à comparer la valeur de f en ce point à la valeur de f aux points réalisables voisins. Ces derniers sont nécessairement de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$.

Nature du point $(0,1)$.

(5) Pour étudier la nature d'un point (x, y) de Ω, on cherche à comparer la valeur de f en ce point à la valeur de f aux points réalisables voisins. Ces derniers sont nécessairement de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$.

Nature du point $(0,1)$. Ses voisins sont construits comme indiqué ci-dessus, avec

$$y_\varepsilon = 1-\varepsilon \quad \text{et} \quad \varepsilon > 0.$$

En ces points voisins, on a

$$f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) = y_\varepsilon - (1-y_\varepsilon^2)$$

En ces points voisins, on a

$$\begin{aligned} f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1-y_\varepsilon^2) \\ &= 1-\varepsilon-(1-(1-\varepsilon)^2) \end{aligned}$$

En ces points voisins, on a

$$\begin{aligned}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1-y_\varepsilon^2) \\
&= 1-\varepsilon-(1-(1-\varepsilon)^2) \\
&= 1-\varepsilon-(1-(1-2\varepsilon+\varepsilon^2))
\end{aligned}$$

En ces points voisins, on a

$$\begin{aligned}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1-y_\varepsilon^2) \\
&= 1-\varepsilon-(1-(1-\varepsilon)^2) \\
&= 1-\varepsilon-(1-(1-2\varepsilon+\varepsilon^2)) \\
&= 1-3\varepsilon+\varepsilon^2
\end{aligned}$$

En ces points voisins, on a

$$
\begin{aligned}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1-y_\varepsilon^2) \\
&= 1-\varepsilon-(1-(1-\varepsilon)^2) \\
&= 1-\varepsilon-(1-(1-2\varepsilon+\varepsilon^2)) \\
&= 1-3\varepsilon+\varepsilon^2 \\
&= 1+\varepsilon(\varepsilon-3)
\end{aligned}
$$

En ces points voisins, on a

$$
\begin{array}{rcl}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) & = & y_\varepsilon - (1-y_\varepsilon^2) \\
& = & 1-\varepsilon-(1-(1-\varepsilon)^2) \\
& = & 1-\varepsilon-(1-(1-2\varepsilon+\varepsilon^2)) \\
& = & 1-3\varepsilon+\varepsilon^2 \\
& = & 1+\varepsilon(\varepsilon-3) \\
& < & 1 \ = \ f(0,1)
\end{array}
$$

pour tout $\varepsilon \in (0,3)$.

En ces points voisins, on a

$$\begin{array}{rcl} f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) & = & y_\varepsilon - (1-y_\varepsilon^2) \\ & = & 1-\varepsilon-(1-(1-\varepsilon)^2) \\ & = & 1-\varepsilon-(1-(1-2\varepsilon+\varepsilon^2)) \\ & = & 1-3\varepsilon+\varepsilon^2 \\ & = & 1+\varepsilon(\varepsilon-3) \\ & < & 1 \;=\; f(0,1) \end{array}$$

pour tout $\varepsilon \in (0,3)$. On en déduit donc que $(0,1)$ est un maximum local strict de (P_1).

Nature du point $(0, -1)$.

Nature du point $(0,-1)$. Les points réalisables voisins de $(0,-1)$ sont de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1+\varepsilon$ et $\varepsilon > 0$.

Nature du point $(0,-1)$. Les points réalisables voisins de $(0,-1)$ sont de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1+\varepsilon$ et $\varepsilon > 0$. On a alors:

$$f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) \quad = \quad y_\varepsilon - (1-y_\varepsilon^2)$$

Nature du point $(0,-1)$. Les points réalisables voisins de $(0,-1)$ sont de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1+\varepsilon$ et $\varepsilon > 0$. On a alors:

$$\begin{aligned} f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1-y_\varepsilon^2) \\ &= -1+\varepsilon-(1-(-1+\varepsilon)^2) \end{aligned}$$

Nature du point $(0,-1)$. Les points réalisables voisins de $(0,-1)$ sont de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1+\varepsilon$ et $\varepsilon > 0$. On a alors:

$$\begin{aligned} f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1-y_\varepsilon^2) \\ &= -1+\varepsilon-(1-(-1+\varepsilon)^2) \\ &= -1+\varepsilon-(1-(1-2\varepsilon+\varepsilon^2)) \end{aligned}$$

Nature du point $(0,-1)$. Les points réalisables voisins de $(0,-1)$ sont de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1+\varepsilon$ et $\varepsilon > 0$. On a alors:

$$\begin{aligned}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1-y_\varepsilon^2) \\
&= -1+\varepsilon-(1-(-1+\varepsilon)^2) \\
&= -1+\varepsilon-(1-(1-2\varepsilon+\varepsilon^2)) \\
&= -1-\varepsilon+\varepsilon^2
\end{aligned}$$

Nature du point $(0,-1)$. Les points réalisables voisins de $(0,-1)$ sont de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1+\varepsilon$ et $\varepsilon > 0$. On a alors:

$$\begin{aligned}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1-y_\varepsilon^2) \\
&= -1+\varepsilon-(1-(-1+\varepsilon)^2) \\
&= -1+\varepsilon-(1-(1-2\varepsilon+\varepsilon^2)) \\
&= -1-\varepsilon+\varepsilon^2 \\
&= -1+\varepsilon(\varepsilon-1)
\end{aligned}$$

Nature du point $(0,-1)$. Les points réalisables voisins de $(0,-1)$ sont de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1+\varepsilon$ et $\varepsilon > 0$. On a alors:

$$\begin{aligned}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1-y_\varepsilon^2) \\
&= -1+\varepsilon-(1-(-1+\varepsilon)^2) \\
&= -1+\varepsilon-(1-(1-2\varepsilon+\varepsilon^2)) \\
&= -1-\varepsilon+\varepsilon^2 \\
&= -1+\varepsilon(\varepsilon-1) \\
&< -1 \;=\; f(0,-1)
\end{aligned}$$

pour tout $\varepsilon \in (0,1)$.

Nature du point $(0,-1)$. Les points réalisables voisins de $(0,-1)$ sont de la forme $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1+\varepsilon$ et $\varepsilon > 0$. On a alors:

$$\begin{array}{rcl} f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) & = & y_\varepsilon - (1-y_\varepsilon^2) \\ & = & -1+\varepsilon-(1-(-1+\varepsilon)^2) \\ & = & -1+\varepsilon-(1-(1-2\varepsilon+\varepsilon^2)) \\ & = & -1-\varepsilon+\varepsilon^2 \\ & = & -1+\varepsilon(\varepsilon-1) \\ & < & -1 \;=\; f(0,-1) \end{array}$$

pour tout $\varepsilon \in (0,1)$. On en déduit donc que $(0,-1)$ est un maximum local strict de (P_1).

Nature des points $(\sqrt{3}/2, -1/2)$ et $(-\sqrt{3}/2, -1/2)$.

Nature des points $(\sqrt{3}/2, -1/2)$ et $(-\sqrt{3}/2, -1/2)$. Leurs voisins sont de la forme (déjà donnée) $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1/2 + \varepsilon$ et $\varepsilon \neq 0$.

Nature des points $(\sqrt{3}/2,-1/2)$ et $(-\sqrt{3}/2,-1/2)$. Leurs voisins sont de la forme (déjà donnée) $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1/2+\varepsilon$ et $\varepsilon \neq 0$. On a alors:

$$f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) \quad = \quad y_\varepsilon - (1-y_\varepsilon^2)$$

Nature des points $(\sqrt{3}/2, -1/2)$ et $(-\sqrt{3}/2, -1/2)$. Leurs voisins sont de la forme (déjà donnée) $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1/2 + \varepsilon$ et $\varepsilon \neq 0$. On a alors:

$$\begin{aligned} f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1 - y_\varepsilon^2) \\ &= -\frac{1}{2} + \varepsilon - \left(1 - (-\frac{1}{2} + \varepsilon)^2\right) \end{aligned}$$

Nature des points $(\sqrt{3}/2, -1/2)$ et $(-\sqrt{3}/2, -1/2)$. Leurs voisins sont de la forme (déjà donnée) $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1/2 + \varepsilon$ et $\varepsilon \neq 0$. On a alors:

$$\begin{aligned}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1 - y_\varepsilon^2) \\
&= -\frac{1}{2} + \varepsilon - \left(1 - (-\frac{1}{2} + \varepsilon)^2\right) \\
&= -\frac{1}{2} + \varepsilon - \left(1 - (\frac{1}{2} - \varepsilon)^2\right)
\end{aligned}$$

Nature des points $(\sqrt{3}/2, -1/2)$ et $(-\sqrt{3}/2, -1/2)$. Leurs voisins sont de la forme (déjà donnée) $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1/2 + \varepsilon$ et $\varepsilon \neq 0$. On a alors:

$$\begin{aligned}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1 - y_\varepsilon^2) \\
&= -\frac{1}{2} + \varepsilon - \left(1 - (-\frac{1}{2} + \varepsilon)^2\right) \\
&= -\frac{1}{2} + \varepsilon - \left(1 - (\frac{1}{2} - \varepsilon)^2\right) \\
&= -\frac{1}{2} + \varepsilon - \left(1 - (\frac{1}{4} - \varepsilon + \varepsilon^2)\right)
\end{aligned}$$

Nature des points $(\sqrt{3}/2, -1/2)$ et $(-\sqrt{3}/2, -1/2)$. Leurs voisins sont de la forme (déjà donnée) $(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon)$ avec $y_\varepsilon = -1/2 + \varepsilon$ et $\varepsilon \neq 0$. On a alors:

$$\begin{aligned}
f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= y_\varepsilon - (1 - y_\varepsilon^2) \\
&= -\frac{1}{2} + \varepsilon - \left(1 - (-\frac{1}{2} + \varepsilon)^2\right) \\
&= -\frac{1}{2} + \varepsilon - \left(1 - (\frac{1}{2} - \varepsilon)^2\right) \\
&= -\frac{1}{2} + \varepsilon - \left(1 - (\frac{1}{4} - \varepsilon + \varepsilon^2)\right) \\
&= -\frac{1}{2} - \frac{3}{4} + \varepsilon^2,
\end{aligned}$$

soit

$$f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) \quad = \quad -\frac{5}{4}+\varepsilon^2$$

soit

$$\begin{aligned} f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= -\frac{5}{4}+\varepsilon^2 \\ &> -\frac{5}{4} = f\left(\frac{\pm\sqrt{3}}{2}, -\frac{1}{2}\right) \end{aligned}$$

pour tout $\varepsilon \neq 0$.

soit

$$\begin{aligned} f(\pm\sqrt{1-y_\varepsilon^2}, y_\varepsilon) &= -\frac{5}{4}+\varepsilon^2 \\ &> -\frac{5}{4} = f\left(\frac{\pm\sqrt{3}}{2}, -\frac{1}{2}\right) \end{aligned}$$

pour tout $\varepsilon \neq 0$. On en déduit donc que les 2 points considérés sont des *minima* locaux stricts de (P_1).

Exercice

On considère le problème d'optimisation

$$(P_2) \quad \left\{ \begin{array}{rl} \text{Max} & f(x,y,z) \\ \text{s.c.} & h_1(x,y,z) = 0 \\ & h_2(x,y,z) = 0 \end{array} \right.$$

avec $f(x,y,z) = x$, $h_1(x,y,z) = 2(x+y)+z-3$, $h_2(x,y,z) = x^2+y^2-z$.

(1) Montrer que le problème (P_2) admet au moins une solution.

(2) Étudier la qualification des contraintes.

(3) En étudiant les conditions nécessaires d'optimalité d'ordre 1 de (P_2), montrer que ce problème admet une unique solution.

(1) Pour montrer l'existence d'une solution au problème (P_2), comme la fonction f est continue, il suffit de montrer que l'ensemble C des points réalisables est fermé et borné.

(1) Pour montrer l'existence d'une solution au problème (P_2), comme la fonction f est continue, il suffit de montrer que l'ensemble C des points réalisables est fermé et borné.

D'une part, on peut écrire $C = h_1^{-1}(\{0\}) \cap h_2^{-1}(\{0\})$.

(1) Pour montrer l'existence d'une solution au problème (P_2), comme la fonction f est continue, il suffit de montrer que l'ensemble C des points réalisables est fermé et borné.

D'une part, on peut écrire $C = h_1^{-1}(\{0\}) \cap h_2^{-1}(\{0\})$. Comme h_1 (resp. h_2) est continue et que $\{0\}$ est un fermé, on obtient immédiatement que $h_1^{-1}(\{0\})$ (resp. $h_2^{-1}(\{0\})$) est un fermé.

(1) Pour montrer l'existence d'une solution au problème (P_2), comme la fonction f est continue, il suffit de montrer que l'ensemble C des points réalisables est fermé et borné.

D'une part, on peut écrire $C = h_1^{-1}(\{0\}) \cap h_2^{-1}(\{0\})$. Comme h_1 (resp. h_2) est continue et que $\{0\}$ est un fermé, on obtient immédiatement que $h_1^{-1}(\{0\})$ (resp. $h_2^{-1}(\{0\})$) est un fermé. Il s'ensuit que C est fermé comme intersection de fermés.

(1) Pour montrer l'existence d'une solution au problème (P_2), comme la fonction f est continue, il suffit de montrer que l'ensemble C des points réalisables est fermé et borné.

D'une part, on peut écrire $C = h_1^{-1}(\{0\}) \cap h_2^{-1}(\{0\})$. Comme h_1 (resp. h_2) est continue et que $\{0\}$ est un fermé, on obtient immédiatement que $h_1^{-1}(\{0\})$ (resp. $h_2^{-1}(\{0\})$) est un fermé. Il s'ensuit que C est fermé comme intersection de fermés. D'autre part,

$$\begin{aligned}
C &= \{(x,y,z) | 3-2(x+y) = x^2+y^2,\ z = 3-2(x+y)\} \\
&= \{(x,y,z) | x^2+y^2+2(x+y)-3 = 0,\ z = 3-2(x+y)\} \\
&= \{(x,y,z) | (x+1)^2+(y+1)^2 = 5,\ z = 3-2(x+y)\}.
\end{aligned}$$

Sous cette forme, il apparaît que les points de C sont les points (x,y,z) tels que (x,y) appartient au cercle de centre $(-1,-1)$ et de rayon $\sqrt{5}$ et $z = 3-2(x+y)$.

Sous cette forme, il apparaît que les points de C sont les points (x,y,z) tels que (x,y) appartient au cercle de centre $(-1,-1)$ et de rayon $\sqrt{5}$ et $z = 3-2(x+y)$. L'ensemble C est donc bien borné.

Sous cette forme, il apparaît que les points de C sont les points (x,y,z) tels que (x,y) appartient au cercle de centre $(-1,-1)$ et de rayon $\sqrt{5}$ et $z = 3-2(x+y)$. L'ensemble C est donc bien borné.

(2) Pour la qualification des contraintes, il faut regarder l'indépendance linéaire des gradients de h_1 et h_2.

Sous cette forme, il apparaît que les points de C sont les points (x,y,z) tels que (x,y) appartient au cercle de centre $(-1,-1)$ et de rayon $\sqrt{5}$ et $z = 3-2(x+y)$. L'ensemble C est donc bien borné.

(2) Pour la qualification des contraintes, il faut regarder l'indépendance linéaire des gradients de h_1 et h_2. On a simplement

$$\nabla h_1(x,y,z) = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \quad \nabla h_2(x,y,z) = \begin{pmatrix} 2x \\ 2y \\ -1 \end{pmatrix}.$$

Il est facile de voir que $\alpha_1 \nabla h_1(x,y,z) + \alpha_2 \nabla h_2(x,y,z) = \mathbf{0}$ si et seulement si

$$\begin{cases} \alpha_2 = \alpha_1 = 0 \ \text{ ou} \\ \alpha_2 = \alpha_1 \ \text{ et } \ x = y = -1 \end{cases}$$

Il est facile de voir que $\alpha_1 \nabla h_1(x,y,z) + \alpha_2 \nabla h_2(x,y,z) = \mathbf{0}$ si et seulement si

$$\begin{cases} \alpha_2 = \alpha_1 = 0 \ \text{ ou} \\ \alpha_2 = \alpha_1 \ \text{ et } \ x = y = -1 \end{cases}$$

Comme les points $(-1,-1,z)$ ne sont clairement pas réalisables, les vecteurs $\nabla h_1(x,y,z)$ et $\nabla h_2(x,y,z)$ sont linéairement indépendants (et donc les contraintes sont qualifiées) en tout point réalisable.

(3) Le lagrangien du problème (P_2) s'écrit

$$\begin{aligned} & L(x,y,z;\lambda) \\ & \quad = \quad -f(x,y,z)+\lambda_1 h_1(x,y,z)+\lambda_2 h_2(x,y,z) \\ & \quad = \quad -x+\lambda_1(2x+2y+z-3)+\lambda_2(x^2+y^2-z). \end{aligned}$$

(3) Le lagrangien du problème (P_2) s'écrit

$$\begin{aligned} & L(x,y,z;\lambda) \\ &= -f(x,y,z) + \lambda_1 h_1(x,y,z) + \lambda_2 h_2(x,y,z) \\ &= -x + \lambda_1(2x+2y+z-3) + \lambda_2(x^2+y^2-z). \end{aligned}$$

En conséquence, les conditions de KKT s'écrivent

$$2(x+y)+z-3 = 0 \tag{4}$$
$$x^2+y^2-z = 0 \tag{5}$$
$$-1+2\lambda_1+2\lambda_2 x = 0 \tag{6}$$
$$2\lambda_1+2\lambda_2 y = 0 \tag{7}$$
$$\lambda_1-\lambda_2 = 0. \tag{8}$$

L'équation (8) est équivalente à $\lambda_2 = \lambda_1$ et les autres équations s'écrivent donc

$$\begin{aligned} 2(x+y)+z-3 &= 0 \qquad (9)\\ x^2+y^2-z &= 0 \qquad (10)\\ -1+2\lambda_1(1+x) &= 0 \qquad (11)\\ \lambda_1(1+y) &= 0. \qquad (12) \end{aligned}$$

Comme $\lambda_1 = 0$ est impossible, l'équation (12) impose $y = -1$.

Comme $\lambda_1 = 0$ est impossible, l'équation (12) impose $y = -1$. Il reste donc à résoudre

$$\begin{aligned} 3 - 2(x-1) &= z \\ x^2 + 1 - \left(3 - 2(x-1)\right) &= 0 \\ -1 + 2\lambda_1(1+x) &= 0. \end{aligned}$$

Comme $\lambda_1 = 0$ est impossible, l'équation (12) impose $y = -1$. Il reste donc à résoudre

$$\begin{aligned} 3 - 2(x-1) &= z \\ x^2 + 1 - \Big(3 - 2(x-1)\Big) &= 0 \\ -1 + 2\lambda_1(1+x) &= 0. \end{aligned}$$

Ce système équivaut à

$$\begin{aligned} 3 - 2(x-1) &= z \\ x^2 + 2x - 4 &= 0 \\ -1 + 2\lambda_1(1+x) &= 0 \end{aligned}$$

soit:

$$
\begin{aligned}
x &= -1 \pm \sqrt{5} \\
z &= 3 - 2(-2 \pm \sqrt{5}) \\
\lambda_1 &= \frac{1}{2(\pm\sqrt{5})}
\end{aligned}
$$

soit:

$$\begin{aligned} x &= -1 \pm \sqrt{5} \\ z &= 3 - 2(-2 \pm \sqrt{5}) \\ \lambda_1 &= \frac{1}{2(\pm\sqrt{5})} \end{aligned}$$

soit encore

$$\begin{aligned} x &= -1 \pm \sqrt{5} \\ z &= 7 \mp 2\sqrt{5} \\ \lambda_1 &= \pm\frac{1}{2\sqrt{5}}. \end{aligned}$$

Comme les points réalisables (x,y,z) de (P_2) sont tels que (x,y) appartient au cercle de centre$(-1,-1)$ et de rayon $\sqrt{5}$, nécessairement, les voisins réalisables de $(-1+\sqrt{5},-1)$ feront diminuer la fonction objectif, alors que ceux de $(-1-\sqrt{5},-1)$ la feront augmenter.

Comme les points réalisables (x,y,z) de (P_2) sont tels que (x,y) appartient au cercle de centre$(-1,-1)$ et de rayon $\sqrt{5}$, nécessairement, les voisins réalisables de $(-1+\sqrt{5},-1)$ feront diminuer la fonction objectif, alors que ceux de $(-1-\sqrt{5},-1)$ la feront augmenter.

On en déduit que le point $(-1+\sqrt{5},-1,7-2\sqrt{5})$ est l'unique solution de (P_2).

Exercice

On considère le problème d'optimisation

$$(P_3) \quad \begin{cases} \text{Min} & f(x,y) \\ \text{s.c.} & g(x,y) \leq 0 \end{cases}$$

avec $f(x,y) = x + y$ et $g(x,y) = (x-1)^2 + y^2 - 1$.

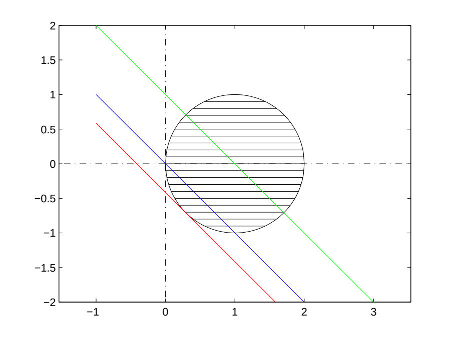
(1) Tracer l'ensemble des points réalisables, ainsi que quelques courbes de niveau de la fonction f.

(2) Montrer que le problème (P_3) admet au moins une solution.

(3) Étudier la qualification des contraintes.

(4) En étudiant les conditions nécessaires d'optimalité d'ordre 1 de (P_3), montrer que ce problème admet une unique solution.

(1) Sur la figure suivante, on a hachuré l'ensemble des points réalisables. On a également représenté en vert la courbe de niveau 1 de f, en bleu la courbe de niveau 0 de f et en rouge la courbe de niveau $1-\sqrt{2}$ de f.



(2) L'ensemble réalisable C est le disque de centre $(1,0)$ et de rayon 1, qui est bien sûr compact. Comme la fonction f est continue, le problème (P_3) admet une solution.

(2) L'ensemble réalisable C est le disque de centre $(1,0)$ et de rayon 1, qui est bien sûr compact. Comme la fonction f est continue, le problème (P_3) admet une solution.

(3) Pour montrer la qualification de la contrainte en tout point réalisable, on doit montrer que le gradient de l'unique contrainte $\nabla g(x,y) = \big(2(x-1), 2y\big)^\top$ forme une famille libre, c'est-à-dire qu'il est non nul, en tout point réalisable en lequel la contrainte est active.

(2) L'ensemble réalisable C est le disque de centre $(1,0)$ et de rayon 1, qui est bien sûr compact. Comme la fonction f est continue, le problème (P_3) admet une solution.

(3) Pour montrer la qualification de la contrainte en tout point réalisable, on doit montrer que le gradient de l'unique contrainte $\nabla g(x,y) = \big(2(x-1), 2y\big)^\top$ forme une famille libre, c'est-à-dire qu'il est non nul, en tout point réalisable en lequel la contrainte est active. Or $\nabla g(x,y)$ s'annule en le seul point $(1,0)$ et ce point est un point réalisable en lequel la contrainte est inactive.

(4) Le lagrangien associé au problème (P_3) s'écrit

$$L(x,y;\mu) = x + y + \mu\left((x-1)^2 + y^2 - 1\right)$$

et les conditions de KKT suivantes en découlent:

$$\begin{aligned}
1 + \mu 2(x-1) &= 0 && (13)\\
1 + \mu 2y &= 0 && (14)\\
(x-1)^2 + y^2 - 1 &\leq 0 && (15)\\
\mu &\geq 0 && (16)\\
\mu\Big((x-1)^2 + y^2 - 1\Big) &= 0.
\end{aligned}$$

(4) Le lagrangien associé au problème (P_3) s'écrit

$$L(x,y;\mu) = x + y + \mu\left((x-1)^2 + y^2 - 1\right)$$

et les conditions de KKT suivantes en découlent:

$$\begin{aligned}
1 + \mu 2(x-1) &= 0 && (13)\\
1 + \mu 2y &= 0 && (14)\\
(x-1)^2 + y^2 - 1 &\leq 0 && (15)\\
\mu &\geq 0 && (16)\\
\mu\left((x-1)^2 + y^2 - 1\right) &= 0. && (17)
\end{aligned}$$

Dans ces conditions, outre l'annulation du gradient du lagrangien par rapport aux variables *primales* (x,y) ((13) et (14)) et la réalisabilité (15), on a une contrainte de signe sur le multiplicateur (16) et la complémentarité (17).

D’après (14), $\mu \neq 0$.

D'après (14), $\mu \neq 0$. On en déduit d'une part que (13) et (14) sont équivalentes à

$$y = -\frac{1}{2\mu} \quad \text{et} \quad x = 1 - \frac{1}{2\mu} = 1 + y,$$

d'autre part que $(x-1)^2 + y^2 - 1 = 0$, c'est-à-dire $2/4\mu^2 = 1$ ou encore $\mu = \pm 1/\sqrt{2}$.

D'après (14), $\mu \neq 0$. On en déduit d'une part que (13) et (14) sont équivalentes à

$$y = -\frac{1}{2\mu} \quad \text{et} \quad x = 1 - \frac{1}{2\mu} = 1 + y,$$

d'autre part que $(x-1)^2 + y^2 - 1 = 0$, c'est-à-dire $2/4\mu^2 = 1$ ou encore $\mu = \pm 1/\sqrt{2}$.
Finalement, la condition (16) permet de dire qu'il n'y a qu'une solution aux conditions de KKT (et donc une seule solution au problème (P_3)):

$$\mu = \frac{1}{\sqrt{2}}, \quad x = 1 - \frac{1}{\sqrt{2}} \quad \text{et} \quad y = -\frac{1}{\sqrt{2}}.$$

Exercice

On considère le problème d'optimisation

$$
(P_4) \qquad \begin{cases} \text{Min} & x \\ \text{s.c.} & |y| \leq 1 - x^2. \end{cases}
$$

(1) Formuler le problème (P_4) comme un problème d'optimisation différentiable.

(2) Tracer l'ensemble des points réalisables, ainsi que quelques courbes de niveau de la fonction objectif.

(3) Montrer que le problème (P_4) admet au moins une solution.

(4) Étudier la qualification des contraintes.

(5) En étudiant les conditions nécessaires d'optimalité d'ordre 1 de (P_4), montrer que ce problème admet une unique solution.

(1) On a évidemment

$$|y| \leq 1 - x^2 \quad \Leftrightarrow \quad -(1 - x^2) \leq y \leq 1 - x^2$$
$$\Leftrightarrow \quad \left\{ \begin{array}{l} x^2 - y - 1 \leq 0 \\ x^2 + y - 1 \leq 0. \end{array} \right.$$

(1) On a évidemment

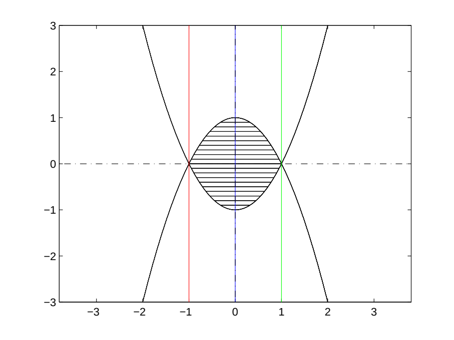
$$\begin{aligned} |y| \leq 1 - x^2 \quad &\Leftrightarrow \quad -(1-x^2) \leq y \leq 1 - x^2 \\ &\Leftrightarrow \quad \left\{ \begin{array}{l} x^2 - y - 1 \leq 0 \\ x^2 + y - 1 \leq 0. \end{array} \right. \end{aligned}$$

On peut donc récrire le problème (P_4) sous la forme

$$\left\{ \begin{array}{ll} \text{Min} & f(x,y) \\ \text{s.c.} & g_1(x,y) \leq 0 \\ & g_2(x,y) \leq 0, \end{array} \right.$$

avec $f(x,y) = x$, $g_1(x,y) = x^2 - y - 1$, $g_2(x,y) = x^2 + y - 1$.

(2) Sur la figure suivante, on a hachuré l'ensemble des points réalisables. On a aussi dessiné en vert la courbe de niveau 1 de f, en bleu la courbe de niveau 0 de f et en rouge la courbe de niveau -1 de f.



(3) Vérifions que l'ensemble admissible C est compact.

(3) Vérifions que l'ensemble admissible C est compact.
L'ensemble C s'écrit aussi

$$\{\, (x,y) \in \mathbb{R}^2 \mid x^2 + |y| \le 1 \,\}\,.$$

(3) Vérifions que l'ensemble admissible C est compact.
L'ensemble C s'écrit aussi

$$\{\, (x,y) \in \mathbb{R}^2 \mid x^2 + |y| \leq 1 \,\}.$$

Un point (x,y) de C vérifie donc

$$x^2 \leq x^2 + |y| \leq 1 \quad \text{et} \quad |y| \leq x^2 + |y| \leq 1.$$

(3) Vérifions que l'ensemble admissible C est compact.
L'ensemble C s'écrit aussi

$$\{\,(x,y) \in \mathbb{R}^2 \mid x^2 + |y| \leq 1\,\}.$$

Un point (x,y) de C vérifie donc

$$x^2 \leq x^2 + |y| \leq 1 \quad \text{et} \quad |y| \leq x^2 + |y| \leq 1.$$

On en déduit que $C \subset [-1,1]^2$ est borné.

(3) Vérifions que l'ensemble admissible C est compact.
L'ensemble C s'écrit aussi

$$\{ (x,y) \in \mathbb{R}^2 \mid x^2 + |y| \leq 1 \} .$$

Un point (x,y) de C vérifie donc

$$x^2 \leq x^2 + |y| \leq 1 \quad \text{et} \quad |y| \leq x^2 + |y| \leq 1.$$

On en déduit que $C \subset [-1,1]^2$ est borné. De plus, $C = g_1^{-1}(\mathbb{R}_-) \cap g_2^{-1}(\mathbb{R}_-)$ est fermé comme intersection de 2 fermés (images réciproques du fermé $\mathbb{R}_-$ par les applications continues g_1 et g_2).

(3) Vérifions que l'ensemble admissible C est compact.
L'ensemble C s'écrit aussi

$$\{\,(x,y)\in\mathbb{R}^2 \mid x^2+|y|\leq 1\,\}\,.$$

Un point (x,y) de C vérifie donc

$$x^2\leq x^2+|y|\leq 1 \quad \text{et} \quad |y|\leq x^2+|y|\leq 1.$$

On en déduit que $C\subset[-1,1]^2$ est borné. De plus, $C=g_1^{-1}(\mathbb{R}_-)\cap g_2^{-1}(\mathbb{R}_-)$ est fermé comme intersection de 2 fermés (images réciproques du fermé $\mathbb{R}_-$ par les applications continues g_1 et g_2). Donc C est bien compact.

(3) Vérifions que l'ensemble admissible C est compact.
L'ensemble C s'écrit aussi

$$\{ (x,y) \in \mathbb{R}^2 \mid x^2 + |y| \leq 1 \} .$$

Un point (x,y) de C vérifie donc

$$x^2 \leq x^2 + |y| \leq 1 \quad \text{et} \quad |y| \leq x^2 + |y| \leq 1.$$

On en déduit que $C \subset [-1,1]^2$ est borné. De plus,
$C = g_1^{-1}(\mathbb{R}_-) \cap g_2^{-1}(\mathbb{R}_-)$ est fermé comme intersection de 2 fermés (images réciproques du fermé $\mathbb{R}_-$ par les applications continues g_1 et g_2). Donc C est bien compact.
Puisque La fonction objectif est continue, nous pouvons conclure à l'existence d'une solution au problème (P_4).

(4) Pour la qualification des contraintes, étudions l'indépendance linéaire des gradients des contraintes

$$\nabla g_1(x,y) = \begin{pmatrix} 2x \\ -1 \end{pmatrix} \quad \text{et} \quad \nabla g_2(x,y) = \begin{pmatrix} 2x \\ 1 \end{pmatrix}.$$

On a:

$$\begin{aligned}
&\mu_1 \nabla g_1(x,y) + \mu_2 \nabla g_2(x,y) = 0 \\
&\qquad \Leftrightarrow \quad 2x(\mu_1+\mu_2) = 0 \quad \text{et} \quad -\mu_1+\mu_2 = 0 \\
&\qquad \Leftrightarrow \quad \mu_2 = \mu_1 \quad \text{et} \quad x\mu_1 = 0 \\
&\qquad \Leftrightarrow \quad (\mu_2 = \mu_1 \quad \text{et} \quad x = 0) \quad \text{ou} \quad \mu_2 = \mu_1 = 0.
\end{aligned}$$

Lorsque $x = 0$, les 2 contraintes ne peuvent pas être actives à la fois.

Lorsque $x = 0$, les 2 contraintes ne peuvent pas être actives à la fois. Le raisonnement est alors le suivant:

- si une seule des 2 contraintes est active, son gradient est nécessairement linéairement indépendant (i.e. différent du vecteur nul);

Lorsque $x = 0$, les 2 contraintes ne peuvent pas être actives à la fois. Le raisonnement est alors le suivant:

- si une seule des 2 contraintes est active, son gradient est nécessairement linéairement indépendant (i.e. différent du vecteur nul);
- si les 2 contraintes sont actives ensemble, ce qui précède montre que les gradients de ces contraintes sont linéairement indépendants.

Lorsque $x = 0$, les 2 contraintes ne peuvent pas être actives à la fois. Le raisonnement est alors le suivant:

- si une seule des 2 contraintes est active, son gradient est nécessairement linéairement indépendant (i.e. différent du vecteur nul);
- si les 2 contraintes sont actives ensemble, ce qui précède montre que les gradients de ces contraintes sont linéairement indépendants.

On en déduit donc que les contraintes sont qualifiées en tout point réalisable.

(4) En notant μ_1 et μ_2 les multiplicateurs de Lagrange associés respectivement à g_1 et g_2, le Lagrangien associé à (P_4) s'écrit

$$L(x,y;\mu_1,\mu_2) = x + \mu_1(x^2 - y - 1) + \mu_2(x^2 + y - 1)$$

et les conditions de KKT en découlent:

$$\left\{\begin{array}{lll} 1 + 2\mu_1 x + 2\mu_2 x = 0 & & \\ -\mu_1 + \mu_2 = 0 & & \\ x^2 - y - 1 \le 0, & \mu_1 \ge 0, & \mu_1(x^2 - y - 1) = 0 \\ x^2 + y - 1 \le 0, & \mu_2 \ge 0, & \mu_2(x^2 + y - 1) = 0. \end{array}\right.$$

On déduit des 2 premières équations que $\mu_2 = \mu_1 \neq 0$.

Les équations des 2 dernières lignes de KKT indiquent donc que

$$\left.\begin{array}{l} x^2 - y - 1 = 0 \\ x^2 + y - 1 = 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} y = x^2 - 1 \\ y = 1 - x^2 \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{l} y = 0 \\ x = \pm 1 \end{array}\right.$$

Pour $x = \pm 1$, les deux premières équations de KKT donnent

$$\mu_2 = \mu_1 = -\frac{1}{4x} = \mp\frac{1}{4}.$$

Pour $x = \pm 1$, les deux premières équations de KKT donnent

$$\mu_2 = \mu_1 = -\frac{1}{4x} = \mp\frac{1}{4}.$$

Finalement, les signes imposés pour μ_1 et μ_2 permettent d'obtenir une unique solution aux conditions de KKT (et donc au problème (P_4)):

$$\mu_2 = \mu_1 = \frac{1}{4}, \quad x = -1, \quad y = 0.$$

Exercice

Soient $A \in \mathbb{R}^{m \times n}$ une matrice dont les lignes sont indépendantes, $D \in \mathbb{R}^{n \times n}$ une matrice diagonale dont les éléments diagonaux sont strictement positifs, $\mathbf{c} \in \mathbb{R}^n$ et $\beta \in \mathbb{R}_+^*$. On cherche à résoudre le problème d'optimisation

$$(\mathscr{P}) \quad \left| \begin{array}{rl} \text{Minimiser} & \langle \mathbf{c}, \mathbf{x} \rangle \\ \text{s.c.} & A\mathbf{x} = \mathbf{0},\ \|D^{-1}\mathbf{x}\|^2 \leq \beta^2. \end{array} \right.$$

(1) Montrer que le problème $(\mathscr{P})$ admet au moins une solution.

(2) Étudier la qualification des contraintes.

(3) Donner le lagrangien du problème $(\mathscr{P})$, puis écrire les conditions nécessaires d'optimalité de Karush-Kuhn-Tucker. Ces conditions sont-elles suffisantes ?

(4) Montrer que, si $\mathbf{c} \notin \operatorname{im} A^\top$, alors

$$\bar{\mathbf{x}} := -\beta \frac{D^2(\mathbf{c} + A^\top \lambda)}{\|D(\mathbf{c} + A^\top \lambda)\|}, \quad \text{où} \quad \lambda := -(AD^2A^\top)^{-1}AD^2\mathbf{c},$$

est solution de $(\mathscr{P})$.

(1) L’ensemble réalisable est compact.

(1) L'ensemble réalisable est compact. En effet, c'est l'intersection d'un fermé, le noyau de A, et d'un compact, l'ellipsoïde

$$E := \{\mathbf{x} \in \mathbb{R}^n \,|\, \|D^{-1}\mathbf{x}\| \leq \beta\} = \varphi^{-1}([0, \beta]).$$

(1) L'ensemble réalisable est compact. En effet, c'est l'intersection d'un fermé, le noyau de A, et d'un compact, l'ellipsoïde

$$E := \{\mathbf{x} \in \mathbb{R}^n \,|\, \|D^{-1}\mathbf{x}\| \leq \beta\} = \varphi^{-1}([0,\beta]).$$

Ici, $\varphi\colon \mathbb{R}^n \to \mathbb{R}_+$, $\mathbf{x} \mapsto \|D^{-1}\mathbf{x}\|$ est clairement continue et $[0,\beta]$ est évidemment fermé, et l'on vérifie sans peine que E est contenu dans la boule fermée de centre l'origine et de rayon $\beta \times \max\{d_1,\ldots,d_n\}$.

(1) L'ensemble réalisable est compact. En effet, c'est l'intersection d'un fermé, le noyau de A, et d'un compact, l'ellipsoïde

$$E := \{\mathbf{x} \in \mathbb{R}^n \,|\, \|D^{-1}\mathbf{x}\| \leq \beta\} = \varphi^{-1}([0,\beta]).$$

Ici, $\varphi\colon \mathbb{R}^n \to \mathbb{R}_+$, $\mathbf{x} \mapsto \|D^{-1}\mathbf{x}\|$ est clairement continue et $[0,\beta]$ est évidemment fermé, et l'on vérifie sans peine que E est contenu dans la boule fermée de centre l'origine et de rayon $\beta \times \max\{d_1,\ldots,d_n\}$. Puisque la fonction objectif est continue, le problème $(\mathscr{P})$ admet au moins une solution.

(2) Notons $\mathbf{a}_i^\top$, $i = 1, \ldots, m$, les lignes de A.

(2) Notons $\mathbf{a}_i^\top$, $i = 1, \ldots, m$, les lignes de A. Si $\mathbf{x} \in \operatorname{int} E$, alors la contrainte d'inégalité est inactive.

(2) Notons $\mathbf{a}_i^\top$, $i = 1, \ldots, m$, les lignes de A. Si $\mathbf{x} \in \operatorname{int} E$, alors la contrainte d'inégalité est inactive. Les contraintes d'égalité $h_i(\mathbf{x}) := \langle \mathbf{a}_i, \mathbf{x} \rangle = 0$ satisfont $\nabla h_i(\mathbf{x}) \equiv \mathbf{a}_i$, de sorte que la famille $\{\nabla h_i(\mathbf{x}) \,|\, i = 1, \ldots, m\}$ est libre.

(2) Notons $\mathbf{a}_i^\top$, $i = 1, \ldots, m$, les lignes de A. Si $\mathbf{x} \in \operatorname{int} E$, alors la contrainte d'inégalité est inactive. Les contraintes d'égalité $h_i(\mathbf{x}) := \langle \mathbf{a}_i, \mathbf{x} \rangle = 0$ satisfont $\nabla h_i(\mathbf{x}) \equiv \mathbf{a}_i$, de sorte que la famille $\{\nabla h_i(\mathbf{x}) \,|\, i = 1, \ldots, m\}$ est libre. Donc la condition (CQ2) est satisfaite en tout point de $\operatorname{int} E$.

(2) Notons $\mathbf{a}_i^\top$, $i = 1, \ldots, m$, les lignes de A. Si $\mathbf{x} \in \operatorname{int} E$, alors la contrainte d'inégalité est inactive. Les contraintes d'égalité $h_i(\mathbf{x}) := \langle \mathbf{a}_i, \mathbf{x} \rangle = 0$ satisfont $\nabla h_i(\mathbf{x}) \equiv \mathbf{a}_i$, de sorte que la famille $\{\nabla h_i(\mathbf{x}) \,|\, i = 1, \ldots, m\}$ est libre. Donc la condition (CQ2) est satisfaite en tout point de $\operatorname{int} E$. Supposons maintenant que $\mathbf{x}$ appartienne au bord ∂E de l'ellispoïde E, c'est-à-dire, que

$$g(\mathbf{x}) := \sum_{k=1}^{n} \left(\frac{x_k}{d_k} \right)^2 - \beta^2 = 0.$$

(2) Notons $\mathbf{a}_i^\top$, $i = 1, \ldots, m$, les lignes de A. Si $\mathbf{x} \in \operatorname{int} E$, alors la contrainte d'inégalité est inactive. Les contraintes d'égalité $h_i(\mathbf{x}) := \langle \mathbf{a}_i, \mathbf{x} \rangle = 0$ satisfont $\nabla h_i(\mathbf{x}) \equiv \mathbf{a}_i$, de sorte que la famille $\{\nabla h_i(\mathbf{x}) \,|\, i = 1, \ldots, m\}$ est libre. Donc la condition (CQ2) est satisfaite en tout point de $\operatorname{int} E$. Supposons maintenant que $\mathbf{x}$ appartienne au bord ∂E de l'ellispoïde E, c'est-à-dire, que

$$g(\mathbf{x}) := \sum_{k=1}^{n} \left(\frac{x_k}{d_k} \right)^2 - \beta^2 = 0.$$

Un calcul immédiat donne $\nabla g(\mathbf{x}) = 2D^{-2}\mathbf{x}$, et l'on voit alors que, pour $\mathbf{x} \in \partial E$, la condition (CQ2) est satisfaite si et seulement si

$$D^{-2}\mathbf{x} \notin \operatorname{vect}\{\mathbf{a}_1, \ldots, \mathbf{a}_m\} = \operatorname{im} A^\top.$$

(3) Le lagrangien du problème $(\mathscr{P})$ est donné par

$$L(\mathbf{x},\lambda,\mu) = \langle \mathbf{c},\mathbf{x}\rangle + \langle \lambda, A\mathbf{x}\rangle + \mu\left(\sum_{k=1}^{n}\left(\frac{x_k}{d_k}\right)^2 - \beta^2\right).$$

(3) Le lagrangien du problème $(\mathscr{P})$ est donné par

$$L(\mathbf{x},\lambda,\mu) = \langle \mathbf{c},\mathbf{x}\rangle + \langle \lambda, A\mathbf{x}\rangle + \mu\left(\sum_{k=1}^{n}\left(\frac{x_k}{d_k}\right)^2 - \beta^2\right).$$

Son gradient (par rapport à $\mathbf{x}$) est donné par:
$\nabla_{\mathbf{x}} L(\mathbf{x},\lambda,\mu) = \mathbf{c} + A^{\top}\lambda + 2\mu D^{-2}\mathbf{x}$.

(3) Le lagrangien du problème $(\mathscr{P})$ est donné par

$$L(\mathbf{x},\lambda,\mu) = \langle \mathbf{c},\mathbf{x}\rangle + \langle \lambda, A\mathbf{x}\rangle + \mu\left(\sum_{k=1}^{n}\left(\frac{x_k}{d_k}\right)^2 - \beta^2\right).$$

Son gradient (par rapport à $\mathbf{x}$) est donné par:
$\nabla_{\mathbf{x}} L(\mathbf{x},\lambda,\mu) = \mathbf{c} + A^\top\lambda + 2\mu D^{-2}\mathbf{x}$. Le théorème de Karush-Kuhn-Tucker affirme que, si $\mathbf{x}$ est une solution qualifiée de $(\mathscr{P})$, alors

$$\begin{aligned}
A\mathbf{x} &= \mathbf{0}, && (1)\\
\|D^{-1}\mathbf{x}\|^2 &\leq \beta^2, && (2)\\
\mathbf{c} + A^\top\lambda + 2\mu D^{-2}\mathbf{x} &= \mathbf{0}, && (3)\\
\mu &\geq 0, && (4)\\
\mu\left(\|D^{-1}\mathbf{x}\|^2 - \beta^2\right) &= 0. && (5)
\end{aligned}$$

(4) Si $\mathbf{c} \notin \operatorname{im} A^\top$, alors ($m < n$ et) $\mathbf{c} + A^\top \lambda \neq \mathbf{0}$, de sorte que $\bar{\mathbf{x}}$ est bien défini et non nul, et que $\mu \neq 0$ (voir l'équation (3)).

(4) Si $\mathbf{c} \notin \operatorname{im} A^\top$, alors ($m < n$ et) $\mathbf{c} + A^\top \lambda \neq \mathbf{0}$, de sorte que $\bar{\mathbf{x}}$ est bien défini et non nul, et que $\mu \neq 0$ (voir l'équation (3)). De plus, $D^{-2}\bar{\mathbf{x}}$ est colinéaire à $\mathbf{c} + A^\top \lambda$, qui n'appartient pas à $\operatorname{im} A^\top$.

(4) Si $\mathbf{c} \notin \operatorname{im} A^\top$, alors ($m < n$ et) $\mathbf{c} + A^\top \lambda \neq \mathbf{0}$, de sorte que $\bar{\mathbf{x}}$ est bien défini et non nul, et que $\mu \neq 0$ (voir l'équation (3)). De plus, $D^{-2}\bar{\mathbf{x}}$ est colinéaire à $\mathbf{c} + A^\top \lambda$, qui n'appartient pas à $\operatorname{im} A^\top$. On en déduit que la condition (CQ2) est satisfaite en $\bar{\mathbf{x}}$.

(4) Si $\mathbf{c} \notin \operatorname{im} A^\top$, alors ($m < n$ et) $\mathbf{c} + A^\top \lambda \neq \mathbf{0}$, de sorte que $\bar{\mathbf{x}}$ est bien défini et non nul, et que $\mu \neq 0$ (voir l'équation (3)). De plus, $D^{-2}\bar{\mathbf{x}}$ est colinéaire à $\mathbf{c} + A^\top \lambda$, qui n'appartient pas à $\operatorname{im} A^\top$. On en déduit que la condition (CQ2) est satisfaite en $\bar{\mathbf{x}}$. Remarquons que le problème est convexe, de sorte que les conditions nécessaires de KKT sont aussi suffisantes.

(4) Si $\mathbf{c} \notin \operatorname{im} A^\top$, alors ($m < n$ et) $\mathbf{c} + A^\top \lambda \neq \mathbf{0}$, de sorte que $\bar{\mathbf{x}}$ est bien défini et non nul, et que $\mu \neq 0$ (voir l'équation (3)). De plus, $D^{-2}\bar{\mathbf{x}}$ est colinéaire à $\mathbf{c} + A^\top \lambda$, qui n'appartient pas à $\operatorname{im} A^\top$. On en déduit que la condition (CQ2) est satisfaite en $\bar{\mathbf{x}}$. Remarquons que le problème est convexe, de sorte que les conditions nécessaires de KKT sont aussi suffisantes.

Par complémentarité (équation (5)), on déduit que la contrainte d'inégalité est active: $\|D^{-1}\mathbf{x}\| = \beta$.

(4) Si $\mathbf{c} \notin \operatorname{im} A^\top$, alors ($m < n$ et) $\mathbf{c} + A^\top \lambda \neq \mathbf{0}$, de sorte que $\bar{\mathbf{x}}$ est bien défini et non nul, et que $\mu \neq 0$ (voir l'équation (3)). De plus, $D^{-2}\bar{\mathbf{x}}$ est colinéaire à $\mathbf{c} + A^\top \lambda$, qui n'appartient pas à $\operatorname{im} A^\top$. On en déduit que la condition (CQ2) est satisfaite en $\bar{\mathbf{x}}$. Remarquons que le problème est convexe, de sorte que les conditions nécessaires de KKT sont aussi suffisantes.

Par complémentarité (équation (5)), on déduit que la contrainte d'inégalité est active: $\|D^{-1}\mathbf{x}\| = \beta$. De l'équation (3), on tire

$$\mathbf{x} = -\frac{D^2(\mathbf{c} + A^\top \lambda)}{2\mu}.$$

Puisque $\|D^{-1}\mathbf{x}\| = \beta$, on déduit que

$$\frac{\|D(\mathbf{c}+A^\top\lambda)\|}{2\mu} = \beta,$$

puis que

$$\mu = \frac{\|D(\mathbf{c}+A^\top\lambda)\|}{2\beta} \quad \text{et} \quad \mathbf{x} = -\beta \frac{D^2(\mathbf{c}+A^\top\lambda)}{\|D(\mathbf{c}+A^\top\lambda)\|}.$$

Puisque $\|D^{-1}\mathbf{x}\| = \beta$, on déduit que

$$\frac{\|D(\mathbf{c}+A^\top\lambda)\|}{2\mu} = \beta,$$

puis que

$$\mu = \frac{\|D(\mathbf{c}+A^\top\lambda)\|}{2\beta} \quad \text{et} \quad \mathbf{x} = -\beta\frac{D^2(\mathbf{c}+A^\top\lambda)}{\|D(\mathbf{c}+A^\top\lambda)\|}.$$

Enfin, l'équation $A\mathbf{x} = \mathbf{0}$ implique que $AD^2(\mathbf{c}+A^\top\lambda) = \mathbf{0}$.

Puisque $\|D^{-1}\mathbf{x}\| = \beta$, on déduit que

$$\frac{\|D(\mathbf{c}+A^\top\lambda)\|}{2\mu} = \beta,$$

puis que

$$\mu = \frac{\|D(\mathbf{c}+A^\top\lambda)\|}{2\beta} \quad \text{et} \quad \mathbf{x} = -\beta\frac{D^2(\mathbf{c}+A^\top\lambda)}{\|D(\mathbf{c}+A^\top\lambda)\|}.$$

Enfin, l'équation $A\mathbf{x} = \mathbf{0}$ implique que $AD^2(\mathbf{c}+A^\top\lambda) = \mathbf{0}$. Il s'ensuit que $AD^2A^\top\lambda = -AD^2\mathbf{c}$, et puisque les lignes de A sont indépendantes, la matrice $AD^2A^\top$ est inversible, de sorte que

$$\lambda = -(AD^2A^\top)^{-1}AD^2\mathbf{c}.$$

Exercice

On considère le problème d'optimisation

$$(\mathscr{P}) \quad \left| \begin{array}{rl} \text{Minimiser} & f(\mathbf{x}) := \frac{1}{2}\left\langle \mathbf{x}, A\mathbf{x} \right\rangle - \left\langle \mathbf{b}, \mathbf{x} \right\rangle \\ \text{s.c.} & \left\langle \mathbf{c}, \mathbf{x} \right\rangle = \gamma, \end{array} \right.$$

où $A \in \mathbb{R}^{n\times n}$ est une matrice symétrique définie positive, $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{c} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ et $\gamma \in \mathbb{R}$.

(1) Montrer que $(\mathscr{P})$ admet une unique solution. Déterminer cette solution.

(2) Donner le lagrangien du problème $(\mathscr{P})$, puis écrire l'algorithme d'Uzawa. Le multiplicateur de Lagrange associé à la contrainte sera initialisé à une valeur λ_0, et sa mise à jour sera effectuée avec un pas constant $\rho > 0$.

(3) Montrer que, si

$$\rho = \rho_\circ := \frac{1}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle},$$

alors les suites (λ_k) et $(\mathbf{x}_k)$ sont constantes à partir du rang $k = 1$.

(4) Montrer que, si $\rho \in (0, \rho_\circ)$, alors les suites (λ_k) et $(\mathbf{x}_k)$ convergent vers des limites à déterminer.

(1) La fonction f est évidemment continue.

(1) La fonction f est évidemment continue. Elle est aussi coercive: en notant $\lambda_{\min}$ la plus petite valeur propre de A, qui est strictement positive puisque A est symétrique définie positive (SDP), on montre aisément que

$$f(\mathbf{x}) \geq \frac{\lambda_{\min}}{2} \left\| \mathbf{x} \right\|^2 - \left\| \mathbf{b} \right\| \left\| \mathbf{x} \right\| \to \infty \quad \text{lorsque} \quad \left\| \mathbf{x} \right\| \to \infty.$$

(1) La fonction f est évidemment continue. Elle est aussi coercive: en notant $\lambda_{\min}$ la plus petite valeur propre de A, qui est strictement positive puisque A est symétrique définie positive (SDP), on montre aisément que

$$f(\mathbf{x}) \geq \frac{\lambda_{\min}}{2} \|\mathbf{x}\|^2 - \|\mathbf{b}\| \|\mathbf{x}\| \to \infty \quad \text{lorsque} \quad \|\mathbf{x}\| \to \infty.$$

Elle atteint donc son *infimum* sur l'ensemble fermé $C := \{\mathbf{x} \in \mathbb{R}^n \,|\, \langle \mathbf{c}, \mathbf{x} \rangle = \gamma\}$.

(1) La fonction f est évidemment continue. Elle est aussi coercive: en notant $\lambda_{\min}$ la plus petite valeur propre de A, qui est strictement positive puisque A est symétrique définie positive (SDP), on montre aisément que

$$f(\mathbf{x}) \geq \frac{\lambda_{\min}}{2} \|\mathbf{x}\|^2 - \|\mathbf{b}\| \|\mathbf{x}\| \to \infty \quad \text{lorsque} \quad \|\mathbf{x}\| \to \infty.$$

Elle atteint donc son *infimum* sur l'ensemble fermé $C := \{\mathbf{x} \in \mathbb{R}^n \,|\, \langle \mathbf{c}, \mathbf{x} \rangle = \gamma\}$. De plus, $\nabla^2 f(\mathbf{x}) = A$ pour tout $\mathbf{x} \in \mathbb{R}^n$, et comme A est SDP, on voit que f est strictement convexe, ce qui entraîne l'unicité de la solution.

Soit $\bar{\mathbf{x}}$ la solution de $(\mathscr{P})$, et soit $h(\mathbf{x}) := \langle \mathbf{c}, \mathbf{x} \rangle - \gamma$.

Soit $\bar{\mathbf{x}}$ la solution de $(\mathscr{P})$, et soit $h(\mathbf{x}) := \langle \mathbf{c} , \mathbf{x} \rangle - \gamma$. Puisque $\nabla h(\mathbf{x}) = \mathbf{c} \neq \mathbf{0}$ pour tout $\mathbf{x}$, la famille $\{\nabla h(\mathbf{x})\}$ est toujours libre, de sorte que la contrainte est partout qualifiée.

Soit $\bar{\mathbf{x}}$ la solution de $(\mathscr{P})$, et soit $h(\mathbf{x}) := \langle \mathbf{c}, \mathbf{x} \rangle - \gamma$. Puisque $\nabla h(\mathbf{x}) = \mathbf{c} \neq \mathbf{0}$ pour tout $\mathbf{x}$, la famille $\{\nabla h(\mathbf{x})\}$ est toujours libre, de sorte que la contrainte est partout qualifiée. D'après le théorème de Karusch-Kuhn-Tucker, il existe $\bar{\lambda}$ tel que

$$\nabla f(\bar{\mathbf{x}}) + \bar{\lambda} \nabla h(\bar{\mathbf{x}}) = \mathbf{0},$$

c'est-à-dire, $A\bar{\mathbf{x}} - \mathbf{b} + \bar{\lambda}\mathbf{c} = \mathbf{0}$.

Soit $\bar{\mathbf{x}}$ la solution de $(\mathscr{P})$, et soit $h(\mathbf{x}) := \langle \mathbf{c}, \mathbf{x} \rangle - \gamma$. Puisque $\nabla h(\mathbf{x}) = \mathbf{c} \neq \mathbf{0}$ pour tout $\mathbf{x}$, la famille $\{\nabla h(\mathbf{x})\}$ est toujours libre, de sorte que la contrainte est partout qualifiée. D'après le théorème de Karusch-Kuhn-Tucker, il existe $\bar{\lambda}$ tel que

$$\nabla f(\bar{\mathbf{x}}) + \bar{\lambda} \nabla h(\bar{\mathbf{x}}) = \mathbf{0},$$

c'est-à-dire, $A\bar{\mathbf{x}} - \mathbf{b} + \bar{\lambda}\mathbf{c} = \mathbf{0}$. Le système KKT s'écrit donc

$$\begin{aligned} \bar{\mathbf{x}} &= A^{-1}(\mathbf{b} - \bar{\lambda}\mathbf{c}), \\ \langle \mathbf{c}, \bar{\mathbf{x}} \rangle &= \gamma. \end{aligned}$$

En combinant ces deux équations, on obtient:

$$\gamma = \langle \mathbf{c}, A^{-1}(\mathbf{b} - \bar{\lambda}\mathbf{c}) \rangle = \langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \bar{\lambda} \langle \mathbf{c}, A^{-1}\mathbf{c} \rangle.$$

Soit $\bar{\mathbf{x}}$ la solution de $(\mathscr{P})$, et soit $h(\mathbf{x}) := \left\langle \mathbf{c}, \mathbf{x} \right\rangle - \gamma$. Puisque $\nabla h(\mathbf{x}) = \mathbf{c} \neq \mathbf{0}$ pour tout $\mathbf{x}$, la famille $\{\nabla h(\mathbf{x})\}$ est toujours libre, de sorte que la contrainte est partout qualifiée. D'après le théorème de Karusch-Kuhn-Tucker, il existe $\bar{\lambda}$ tel que

$$\nabla f(\bar{\mathbf{x}}) + \bar{\lambda} \nabla h(\bar{\mathbf{x}}) = \mathbf{0},$$

c'est-à-dire, $A\bar{\mathbf{x}} - \mathbf{b} + \bar{\lambda}\mathbf{c} = \mathbf{0}$. Le système KKT s'écrit donc

$$\begin{aligned} \bar{\mathbf{x}} &= A^{-1}(\mathbf{b} - \bar{\lambda}\mathbf{c}), \\ \langle \mathbf{c}, \bar{\mathbf{x}} \rangle &= \gamma. \end{aligned}$$

En combinant ces deux équations, on obtient:

$$\gamma = \left\langle \mathbf{c}, A^{-1}(\mathbf{b} - \bar{\lambda}\mathbf{c}) \right\rangle = \left\langle \mathbf{c}, A^{-1}\mathbf{b} \right\rangle - \bar{\lambda} \left\langle \mathbf{c}, A^{-1}\mathbf{c} \right\rangle.$$

Il vient:

$$\bar{\lambda} = \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle}, \quad \bar{\mathbf{x}} = A^{-1}\mathbf{b} - \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} A^{-1}\mathbf{c}.$$

(2) Le lagrangien du problème $(\mathscr{P})$ est donné par

$$\begin{array}{rccl} L\colon & \mathbb{R}^n \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ & (\mathbf{x}, \lambda) & \longmapsto & f(\mathbf{x}) + \lambda\big(\langle \mathbf{c}, \mathbf{x}\rangle - \gamma\big). \end{array}$$

(2) Le lagrangien du problème $(\mathscr{P})$ est donné par

$$\begin{array}{rccl} L: & \mathbb{R}^n \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ & (\mathbf{x}, \lambda) & \longmapsto & f(\mathbf{x}) + \lambda\left(\langle \mathbf{c}, \mathbf{x}\rangle - \gamma\right). \end{array}$$

On a:

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda) = \nabla f(\mathbf{x}) + \lambda \mathbf{c} = A\mathbf{x} + \mathbf{b} + \lambda \mathbf{c} \quad \text{et} \quad \nabla_{\lambda} L(\mathbf{x}, \lambda) = \langle \mathbf{c}, \mathbf{x}\rangle - \gamma.$$

(2) Le lagrangien du problème $(\mathscr{P})$ est donné par

$$L: \begin{array}{ccc} \mathbb{R}^n \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (\mathbf{x}, \lambda) & \longmapsto & f(\mathbf{x}) + \lambda\left(\langle \mathbf{c}, \mathbf{x}\rangle - \gamma\right). \end{array}$$

On a:

$$\nabla_{\mathbf{x}} L(\mathbf{x},\lambda) = \nabla f(\mathbf{x}) + \lambda \mathbf{c} = A\mathbf{x} + \mathbf{b} + \lambda \mathbf{c} \quad \text{et} \quad \nabla_{\lambda} L(\mathbf{x},\lambda) = \langle \mathbf{c}, \mathbf{x}\rangle - \gamma.$$

Rappelons que l'algorithme d'Uzawa, s'écrit, dans le cas général avec contraintes d'égalités et d'inégalités:

Initialisation: Choisir $(\mu_0, \lambda_0) \in \mathbb{R}_+^p \times \mathbb{R}^m$.
Itération: Tant qu'une *condition d'arrêt* n'est pas satisfaite,

1. calculer $\mathbf{x}_k$, minimiseur de $L(\cdot, \mu_k, \lambda_k)$;
2. calculer μ_{k+1} et λ_{k+1}, via:
   ▷ $\lambda_{k+1} = \lambda_k + \alpha_k \mathbf{h}(\mathbf{x}_k)$ [gradient],
   ▷ $\mu_{k+1} = \max\{0, \mu_k + \alpha_k \mathbf{g}(\mathbf{x}_k)\}$ [gradient projeté].

Dans le cas présent, nous avons une seule contrainte d'égalité et aucune contrainte d'inégalité.

Dans le cas présent, nous avons une seule contrainte d'égalité et aucune contrainte d'inégalité. De plus, la minimisation de $L(\cdot, \lambda_k)$ est explicite.

Dans le cas présent, nous avons une seule contrainte d'égalité et aucune contrainte d'inégalité. De plus, la minimisation de $L(\cdot, \lambda_k)$ est explicite. En effet, cette dernière fonction est convexe et différentiable, et le principe de Fermat nous donne l'équation $A\mathbf{x}_k - \mathbf{b} + \lambda_k \mathbf{c} = \mathbf{0}$, soit

$$\mathbf{x}_k = A^{-1}\left(\mathbf{b} - \lambda_k \mathbf{c}\right).$$

Dans le cas présent, nous avons une seule contrainte d'égalité et aucune contrainte d'inégalité. De plus, la minimisation de $L(\cdot, \lambda_k)$ est explicite. En effet, cette dernière fonction est convexe et différentiable, et le principe de Fermat nous donne l'équation $A\mathbf{x}_k - \mathbf{b} + \lambda_k \mathbf{c} = \mathbf{0}$, soit

$$\mathbf{x}_k = A^{-1}\left(\mathbf{b} - \lambda_k \mathbf{c}\right).$$

Enfin, la maximisation en la variable duale λ se fait ici par la méthode du gradient à pas constant ρ, ce qui s'écrit: $\lambda_{k+1} = \lambda_k + \rho(\langle \mathbf{c}, \mathbf{x}_k \rangle - \gamma)$. Nous obtenons l'algorithme:

Initialisation: Choisir $\lambda_0 \in \mathbb{R}$.
Itération: Tant qu'une *condition d'arrêt* n'est pas satisfaite,

1. calculer $\mathbf{x}_k = A^{-1}(\mathbf{b} - \lambda_k \mathbf{c})$;
2. calculer $\lambda_{k+1} = \lambda_k + \rho(\langle \mathbf{c}, \mathbf{x}_k \rangle - \gamma)$.

On pourra prendre comme condition d'arrêt:

$$\left\| \mathbf{x}_{k+1} - \mathbf{x}_k \right\| + \left| \lambda_{k+1} - \lambda_k \right| \leq \varepsilon,$$

où $\varepsilon > 0$ est une tolérance.

On pourra prendre comme condition d'arrêt:

$$\|\mathbf{x}_{k+1}-\mathbf{x}_k\|+|\lambda_{k+1}-\lambda_k|\leq\varepsilon,$$

où $\varepsilon>0$ est une tolérance.

(3) Déroulons les premières itérations de l'algorithme:

$$\mathbf{x}_0 \quad = \quad A^{-1}(\mathbf{b}-\lambda_0\mathbf{c})$$

On pourra prendre comme condition d'arrêt:

$$\left\| \mathbf{x}_{k+1} - \mathbf{x}_k \right\| + \left| \lambda_{k+1} - \lambda_k \right| \leq \varepsilon,$$

où $\varepsilon > 0$ est une tolérance.

(3) Déroulons les premières itérations de l'algorithme:

$$\begin{aligned} \mathbf{x}_0 &= A^{-1}(\mathbf{b} - \lambda_0 \mathbf{c}), \\ \lambda_1 &= \lambda_0 + \frac{\langle \mathbf{c}, A^{-1}(\mathbf{b} - \lambda_0 \mathbf{c}) \rangle - \gamma}{\langle \mathbf{c}, A^{-1} \mathbf{c} \rangle} \end{aligned}$$

On pourra prendre comme condition d'arrêt:

$$\left\| \mathbf{x}_{k+1} - \mathbf{x}_k \right\| + \left| \lambda_{k+1} - \lambda_k \right| \leq \varepsilon,$$

où $\varepsilon > 0$ est une tolérance.

(3) Déroulons les premières itérations de l'algorithme:

$$\begin{aligned}
\mathbf{x}_0 &= A^{-1}(\mathbf{b} - \lambda_0 \mathbf{c}), \\
\lambda_1 &= \lambda_0 + \frac{\langle \mathbf{c}, A^{-1}(\mathbf{b} - \lambda_0 \mathbf{c}) \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_0 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \lambda_0 \langle \mathbf{c}, A^{-1}\mathbf{c} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle}
\end{aligned}$$

On pourra prendre comme condition d'arrêt:

$$\left\| \mathbf{x}_{k+1} - \mathbf{x}_k \right\| + \left| \lambda_{k+1} - \lambda_k \right| \leq \varepsilon,$$

où $\varepsilon > 0$ est une tolérance.

(3) Déroulons les premières itérations de l'algorithme:

$$\begin{aligned}
\mathbf{x}_0 &= A^{-1}(\mathbf{b} - \lambda_0 \mathbf{c}), \\
\lambda_1 &= \lambda_0 + \frac{\langle \mathbf{c}, A^{-1}(\mathbf{b} - \lambda_0 \mathbf{c}) \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_0 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \lambda_0 \langle \mathbf{c}, A^{-1}\mathbf{c} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle}
\end{aligned}$$

On pourra prendre comme condition d'arrêt:

$$\| \mathbf{x}_{k+1} - \mathbf{x}_k \| + | \lambda_{k+1} - \lambda_k | \leq \varepsilon,$$

où $\varepsilon > 0$ est une tolérance.

(3) Déroulons les premières itérations de l'algorithme:

$$\begin{aligned}
\mathbf{x}_0 &= A^{-1}(\mathbf{b} - \lambda_0 \mathbf{c}), \\
\lambda_1 &= \lambda_0 + \frac{\langle \mathbf{c}, A^{-1}(\mathbf{b} - \lambda_0 \mathbf{c}) \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_0 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \lambda_0 \langle \mathbf{c}, A^{-1}\mathbf{c} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle}, \\
\mathbf{x}_1 &= A^{-1}(\mathbf{b} - \lambda_1 \mathbf{c})
\end{aligned}$$

puis

$$\lambda_2 = \lambda_1 + \frac{\langle \mathbf{c}, \mathbf{x}_1 \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle}$$

puis

$$\begin{aligned}
\lambda_2 &= \lambda_1 + \frac{\langle \mathbf{c}, \mathbf{x}_1 \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_1 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} - \lambda_1 A^{-1}\mathbf{c} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle}
\end{aligned}$$

puis

$$\begin{aligned}
\lambda_2 &= \lambda_1 + \frac{\langle \mathbf{c}, \mathbf{x}_1 \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_1 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} - \lambda_1 A^{-1}\mathbf{c} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_1 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} - \lambda_1
\end{aligned}$$

puis

$$
\begin{aligned}
\lambda_2 &= \lambda_1 + \frac{\langle \mathbf{c}, \mathbf{x}_1 \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_1 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} - \lambda_1 A^{-1}\mathbf{c} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_1 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} - \lambda_1 \\
&= \lambda_1
\end{aligned}
$$

puis

$$
\begin{aligned}
\lambda_2 &= \lambda_1 + \frac{\langle \mathbf{c}, \mathbf{x}_1 \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_1 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} - \lambda_1 A^{-1}\mathbf{c} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} \\
&= \lambda_1 + \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} - \lambda_1 \\
&= \lambda_1 ,
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{x}_2 &= A^{-1}(\mathbf{b} - \lambda_2 \mathbf{c}) \\
&= A^{-1}(\mathbf{b} - \lambda_1 \mathbf{c}) \\
&= \mathbf{x}_1 ,
\end{aligned}
$$

et ainsi de suite. On constate donc que les suites (λ_k) et $(\mathbf{x}_k)$ sont constantes à partir du rang $k = 1$.

(4) On a:

$$\lambda_{k+1} = \lambda_k + \rho(\langle \mathbf{c}, \mathbf{x}_k \rangle - \gamma)$$

(4) On a:

$$\begin{aligned}
\lambda_{k+1} &= \lambda_k + \rho(\langle \mathbf{c}, \mathbf{x}_k \rangle - \gamma) \\
&= \lambda_k + \rho(\langle \mathbf{c}, A^{-1}\mathbf{b} - \lambda_k A^{-1}\mathbf{c} \rangle - \gamma)
\end{aligned}$$

(4) On a:

$$\begin{aligned}
\lambda_{k+1} &= \lambda_k + \rho(\langle \mathbf{c}, \mathbf{x}_k \rangle - \gamma) \\
&= \lambda_k + \rho(\langle \mathbf{c}, A^{-1}\mathbf{b} - \lambda_k A^{-1}\mathbf{c} \rangle - \gamma) \\
&= \lambda_k \left(1 - \rho \langle \mathbf{c}, A^{-1}\mathbf{c} \rangle\right) + \rho\left(\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma\right)
\end{aligned}$$

(4) On a:

$$\begin{aligned}
\lambda_{k+1} &= \lambda_k + \rho(\langle \mathbf{c}, \mathbf{x}_k\rangle - \gamma) \\
&= \lambda_k + \rho(\langle \mathbf{c}, A^{-1}\mathbf{b} - \lambda_k A^{-1}\mathbf{c}\rangle - \gamma) \\
&= \lambda_k\left(1 - \rho\langle \mathbf{c}, A^{-1}\mathbf{c}\rangle\right) + \rho\left(\langle \mathbf{c}, A^{-1}\mathbf{b}\rangle - \gamma\right) \\
&= \varphi(\lambda_k),
\end{aligned}$$

où $\varphi(\lambda) = \alpha\lambda + \beta$, avec

$$\alpha = 1 - \rho\langle \mathbf{c}, A^{-1}\mathbf{c}\rangle \quad \text{et} \quad \beta = \rho\left(\langle \mathbf{c}, A^{-1}\mathbf{b}\rangle - \gamma\right)$$

(4) On a:

$$\begin{aligned}
\lambda_{k+1} &= \lambda_k + \rho(\langle \mathbf{c}, \mathbf{x}_k \rangle - \gamma) \\
&= \lambda_k + \rho(\langle \mathbf{c}, A^{-1}\mathbf{b} - \lambda_k A^{-1}\mathbf{c} \rangle - \gamma) \\
&= \lambda_k \left(1 - \rho \langle \mathbf{c}, A^{-1}\mathbf{c} \rangle\right) + \rho\left(\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma\right) \\
&= \varphi(\lambda_k),
\end{aligned}$$

où $\varphi(\lambda) = \alpha\lambda + \beta$, avec

$$\alpha = 1 - \rho \langle \mathbf{c}, A^{-1}\mathbf{c} \rangle \quad \text{et} \quad \beta = \rho\left(\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma\right)$$

Puisque $\alpha \in (0,1)$, la fonction φ est une contraction, et le théorème du point fixe de Banach-Picard montre que la suite (λ_k) converge vers l'unique point fixe $\bar{\lambda}$ de la fonction φ, à savoir

$$\bar{\lambda} = \frac{\beta}{1-\alpha} = \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle}.$$

(4) On a:

$$\begin{aligned}
\lambda_{k+1} &= \lambda_k + \rho(\langle \mathbf{c}, \mathbf{x}_k \rangle - \gamma) \\
&= \lambda_k + \rho(\langle \mathbf{c}, A^{-1}\mathbf{b} - \lambda_k A^{-1}\mathbf{c} \rangle - \gamma) \\
&= \lambda_k \left(1 - \rho \langle \mathbf{c}, A^{-1}\mathbf{c} \rangle\right) + \rho\left(\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma\right) \\
&= \varphi(\lambda_k),
\end{aligned}$$

où $\varphi(\lambda) = \alpha\lambda + \beta$, avec

$$\alpha = 1 - \rho \langle \mathbf{c}, A^{-1}\mathbf{c} \rangle \quad \text{et} \quad \beta = \rho\left(\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma\right)$$

Puisque $\alpha \in (0,1)$, la fonction φ est une contraction, et le théorème du point fixe de Banach-Picard montre que la suite (λ_k) converge vers l'unique point fixe $\bar{\lambda}$ de la fonction φ, à savoir

$$\bar{\lambda} = \frac{\beta}{1-\alpha} = \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle}.$$

Il s'ensuit que $(\mathbf{x}_k)$ converge vers

$$A^{-1}\mathbf{b} - \bar{\lambda} A^{-1}\mathbf{c} = A^{-1}\mathbf{b} - \frac{\langle \mathbf{c}, A^{-1}\mathbf{b} \rangle - \gamma}{\langle \mathbf{c}, A^{-1}\mathbf{c} \rangle} A^{-1}\mathbf{c} = \bar{\mathbf{x}}.$$