Optimisation

Corrigée DS 9 mars 2018

Exercice 1 Problème de modélisation :
Dans un bureau de poste à Lille on n'a que des employés à temps plein. Le nombre d'employés requis par jour dépend de la journée (voir le tableau ci-dessous). Il faut que chaque employé travaille cinq jours consécutifs suivi par deux jours de congé. Le directeur du bureau de poste veut déterminer le nombre minimal d'employés à temps plein qu'il doit recruter afin qu'il y ait suffisamment d'employés disponibles tous les jours.
Quel problème d'optimisation doit il résoudre ? Commencer par définir les variables qui interviennent.

	nombre d'employés requis
lundi	17
mardi	13
mercredi	15
jeudi	19
vendredi	14
samedi	16
dimanche	11

Corrigée :

1. Les variables qui interviennent :
$$x_i, 1 \le i \le 7 \ : \ \text{nombre d'employés qui commencent travailler le } i^{\text{ème}} \text{ jour.}$$

2. Les constantes :
$$b_i, 1 \le i \le 7 \ : \ \text{nombre d'employés requis le } i^{\text{ème}} \text{ jour.}$$

3. L'objectif :
Minimiser le nombre d'employés : $\sum_{i=1}^{7} x_i$.

4. Les contraintes :
Sachant que les employés qui commencent travailler le mardi ou le mercredi ont congé le lundi, le nombre d'employés qui travaillent le lundi vaut $x_1 + x_4 + x_5 + x_6 + x_7$. Il faut donc que
$$x_1 + x_4 + x_5 + x_6 + x_7 \ge b_1.$$

De façon similaire on obtient les contraintes suivantes :

$$
\text{le nombre d'employés qui travaillent le} \quad
\begin{array}{lcl}
\text{mardi} & = & x_1 + x_2 + x_5 + x_6 + x_7 \geq b_2 \\
\text{mercredi} & = & x_1 + x_2 + x_3 + x_6 + x_7 \geq b_3 \\
\text{jeudi} & = & x_1 + x_2 + x_3 + x_4 + x_7 \geq b_4 \\
\text{vendredi} & = & x_1 + x_2 + x_3 + x_4 + x_5 \geq b_5 \\
\text{samedi} & = & x_2 + x_3 + x_4 + x_5 + x_6 \geq b_6 \\
\text{dimanche} & = & x_3 + x_4 + x_5 + x_6 + x_7 \geq b_7
\end{array}
$$

En plus il faut que $x_i \in \mathbb{N}$ (soit $x_i \geq 0$) pour $i \leq i \leq 7$.

5. Le problème d'optimisation :

$$
\begin{array}{ll}
\min & \sum_{i=1}^{7} x_i \\
\text{s.c.} & \begin{cases} b - Ax \leq 0 \\ x_i \in \mathbb{N},\ 1 \leq i \leq 7 \text{ (soit } -x_i \leq 0,\ 1 \leq i \leq 7) \end{cases}
\end{array}
$$

où $b = (b_1, b_2, \ldots, b_7)^T$ et

$$
A = \begin{pmatrix}
1 & 0 & 0 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 0 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 0 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}.
$$

<u>Exercice 2</u> Problème d'optimisation sans contraintes :
On considère sur $\mathbb{R}^2$ la fonction

$$f(x) = x_1^4 - 6x_1^2x_2^2 + 4x_2^3$$

1. Déterminer les candidats à minimum local.

Corrigée :

$$\nabla f(x) = (4x_1^3 - 12x_1x_2^2\ ,\ -12x_1^2x_2 + 12x_2^2),$$

donc

$$\nabla f(x) = 0 \Leftrightarrow \begin{cases} 4x_1(x_1^2 - 3x_2^2) = 0 \\ -12x_2(x_1^2 - x_2) = 0 \end{cases} \Leftrightarrow x = 0 \vee \begin{cases} x_1^2 = x_2 \neq 0 \\ 3x_2^2 = x_2 \end{cases}$$

On a donc 3 points critiques : $x^{(1)} = (0,0)^T$, $x^{(2)} = (1/\sqrt{3}, 1/3)^T$, $x^{(3)} = (-1/\sqrt{3}, 1/3)^T$.

2. Y-a-t-il des minimums locaux ? Justifier votre réponse.

Corrigée :

$$\nabla^2 f(x) = \begin{pmatrix} 12(x_1^2 - x_2^2) & -24x_1x_2 \\ -24x_1x_2 & -12(x_1^2 - 2x_2) \end{pmatrix}.$$

En result

- $$\nabla^2 f(x^{(1)}) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ est SDP mais ne pas DP.}$$

  Donc on ne peut rien dire quant à la nature de $x^{(1)}$ par utilisation des théorèmes du cours. Or pour tout $r > 0$ on a

  $$\begin{array}{l} (0, r/2) \in \ B(x^{(1)}, r) \text{ et } f(0, r/2) = r^3/2 > 0 \\ (0, -r/2) \in \ B(x^{(1)}, r) \text{ et } f(0, -r/2) = -r^3/2 < 0, \end{array}$$

  donc $x^{(1)}$ n'est pas un minimum local ;

- $$\nabla^2 f(x^{(2)}) = \begin{pmatrix} 8/3 & -8/\sqrt{3} \\ -8/\sqrt{3} & 4 \end{pmatrix} \text{ et dét}(\nabla^2 f(x^{(2)}) - \lambda I) = \left(\frac{8}{3} - \lambda\right)(4 - \lambda) - \frac{64}{3},$$

  donc

  $$\begin{aligned} \lambda \in \mathrm{Sp}(\nabla^2 f(x^{(2)}) &\Leftrightarrow (8 - 3\lambda)(4 - \lambda) - 64 = 0 \Leftrightarrow 3\lambda^2 - 20\lambda - 32 = 0 \\ &\Leftrightarrow \lambda = \frac{20 \pm \sqrt{400 - 4 \times 3 \times (-32)}}{2 \times 3} = \frac{20 \pm \sqrt{4^2 \times 25 + 4^2 \times 3 \times 8}}{6} \\ &= \frac{20 \pm 4\sqrt{25 + 24}}{6} = \frac{20 \pm 4 \times 7}{6} = \frac{10 \pm 14}{3}. \end{aligned}$$

  On en déduit que $\nabla^2 f(x^{(2)})$ n'est pas SDP, donc $x^{(2)}$ n'est pas un minimum local ;

- $$\nabla^2 f(x^{(3)}) = \begin{pmatrix} 8/3 & 8/\sqrt{3} \\ 8/\sqrt{3} & 4 \end{pmatrix} \text{ et dét}(\nabla^2 f(x^{(3)}) - \lambda I) = \left(\frac{8}{3} - \lambda\right)(4 - \lambda) - \frac{64}{3},$$

  donc $\mathrm{Sp}(\nabla^2 f(x^{(3)}) = \mathrm{Sp}(\nabla^2 f(x^{(2)})$. On en déduit que $\nabla^2 f(x^{(3)})$ n'est pas SDP, donc $x^{(3)}$ n'est pas un minimum local.

Donc f n'a pas de minimums locaux.

3. La fonction a-t-elle un minimum global ? Pourquoi ?

   Corrigée :
   Comme f n'a pas de minimums locaux, elle n'a pas de minimums globaux non plus. Notez aussi que
   $$\lim_{x_2 \to -\infty} f(0, x_2) = \lim_{x_2 \to -\infty} 4x_2^3 = -\infty,$$
   donc f n'est pas minorée sur $\mathbb{R}^2$.

<u>**Exercice 3**</u> Problème d'optimisation avec contraintes :
On considère le problème (P) $\min_{x \in \mathcal{A}} f(x)$ avec

$$f(x) = \left(\prod_{i=1}^{n} x_i \right)^{-2} \quad \text{et } \mathcal{A} = B(a, r) \cap \mathcal{H}(a, b),$$

où $n > 1$, $a = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$, $r = 1/2$, $B(a, r) = \{x \in \mathbb{R}^n : \|x - a\| \leqslant r\}$ est la boule dans $\mathbb{R}^n$ centrée au point a et de rayon r, $b = n + 1/2$, et $\mathcal{H}(a, b) = \{x \in \mathbb{R}^n : a^T x \leqslant b\}$ est le demi-espace dans $\mathbb{R}^n$.

1. Vérifier que la fonction f est continue sur $\mathcal{A}$. *Remarque : f est continue au point $x_* \in \mathcal{A}$ si $\lim_{x \to x_*} f(x) = \ell \in \mathbb{R}$ et $\ell = f(x_*)$; il suffira donc de démontrer que pour tout $x \in \mathcal{A}$, $f(x) \neq \infty$.*

   Corrigée :
   D'abord on note que pour $x \in \mathbb{R}^n$: $f(x) = \infty$ ssi $\exists i \in \{1, \ldots, n\}$ tq $x_i = 0$. Or soit $x \in \mathbb{R}^n$ tq $x_i = 0$ et $x_j \in \mathbb{R}$ pour tout $j \neq i$. Alors
   $$\|x - a\| = \sqrt{\sum_{j=1}^{n} (x_j - a_j)^2} \geq \sqrt{(x_i - a_i)^2} = |x_i - a_i| = 1 > r, \text{ donc } x \notin B(a, r).$$
   On en déduit que pour tout $x \in \mathcal{A}$, et pour tout $i \in \{1, \ldots, n\}$: $x_i \neq 0$.

D'abord on montrera que (P) est un problème d'optimisation strictement convexe :

2a. Soient $C_1, C_2 \subset \mathbb{R}^n$ deux ensembles convexes. Montrer que l'intersection $C_1 \cap C_2$ est un convexe. En déduire que $\mathcal{A}$ est un convexe. *Remarque : On admet sans démonstration que la boule et le demi-espace sont des convexes.*

   Corrigée :
   $\forall x, y \in C_1 \cap C_2$ et $\forall t \in [0, 1]$ on a :

   - $tx + (1 - t)y \in C_1$ car $x, y \in C_1$ et C_1 est un convexe ;

- $tx + (1-t)y \in C_2$ car $x, y \in C_2$ et C_2 est un convexe.

On en déduit que $tx + (1-t)y \in C_1 \cap C_2$, donc $C_1 \cap C_2$ est un convexe. Comme $\mathcal{A}$ est l'intersection de deux ensembles convexes, et puisque l'intersection de deux ensembles convexes est un convexe, on en déduit que $\mathcal{A}$ est un convexe.

2b. Soient

$$y = \left(\frac{2}{x_1}, \dots, \frac{2}{x_n}\right)^T \in \mathbb{R}^n \text{ et } D = \operatorname{diag}\left(\frac{2}{x_1^2}, \dots, \frac{2}{x_n^2}\right) \in \mathcal{S}_n(\mathbb{R}).$$

Montrer que le gradient de f peut s'ecrire sous la forme $\nabla f(x) = -f(x) \times y^T$. Puis montrer que la matrice hessienne peut s'écrire sous la forme $\nabla^2 f(x) = f(x) \times [yy^T + D]$.

Corrigée :
Pour tout $j \in \{1, \dots, n\}$ on a :

$$\frac{\partial f}{\partial x_j} = -2\left(\prod_{i=1}^{n} x_i\right)^{-3} \times \prod_{i=1, i\neq j}^{n} x_i = -\frac{2}{x_j}\left(\prod_{i=1}^{n} x_i\right)^{-2} = -\frac{2}{x_j} f(x).$$

On en déduit que $\nabla f(x) = -f(x) \times y^T$. Puis on a :

$$\nabla^2 f(x) = \nabla(\nabla^T f(x)) = \nabla(-f(x) \times y) = -y \times \nabla f(x) - f(x)\nabla y = f(x) \times [yy^T - \nabla y].$$

Or pour tout $j \in \{1, \dots, n\}$ on a :

$$\frac{\partial}{\partial x_j} y = -\frac{2}{x_j^2} e^{(j)},$$

où $e^{(j)}$ est le $j^{\text{ème}}$ vecteur canonique (c-à.d., $e_j^{(j)} = 1$ et $e_i^{(j)} = 0$ si $i \neq j$), donc $\nabla y = -D$.

2c. Montrer que $\nabla^2 f(x)$ est définie positive sur $\mathcal{A}$ et en déduire que f est strictement convexe sur $\mathcal{A}$.

Corrigée :
D'abord on remarque que $f(x) \geq 0$ pour tout $x \in \mathbb{R}^n$ et que $f(x) \neq 0$ pour tout $x \in \mathcal{A}$ car $\mathcal{A} \subset B(a, r)$, donc $\mathcal{A}$ borné ($x_i \neq \infty$ pour tout $i \in \{1, \dots, n\}$). En conséquence $\forall x \in \mathcal{A}$ et $\forall z \in \mathbb{R}^n \setminus \{0\}$ on a :

$$\begin{aligned} z^T \nabla^2 f(x) z = f(x) \times [z^T(yy^T)z + z^T D z] &= f(x) \times [(z^T y)(y^T z) + z^T D z] \\ &= f(x) \times [(y^T z)^T (y^T z) + z^T D z] = f(x) \times [\|y^T z\|^2 + z^T D z] > 0, \end{aligned}$$

car $\|y^T z\|^2 \geq 0$ et D est DP (donc $z^T D z > 0$). Du coup $\nabla^2 f(x)$ est définie positive sur $\mathcal{A}$ et f est strictement convexe d'après Théorème 1b.

2d. Conclure.

Corrigée :
(P) est bien un problème strictement convexe car l'ensemble admissible $\mathcal{A}$ est un convexe d'après question 2a et la fonction objectif f est strictement convexe d'après question 2c.

Puis on étudiera l'existence et l'unicité d'une solution optimale :

3. Montrer que (P) possède une et une seule solution optimale.

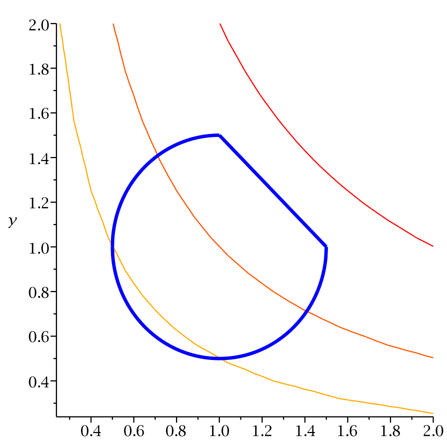
Corrigée :
Comme les fonctions $h_1(x) = \|x - a\|^2 - r^2$ et $h_2(x) = a^T x - b$ sont continues, l'ensemble $\mathcal{A}$ est fermé. De plus $\mathcal{A}$ est borné (car $\mathcal{A} \subset B(a, r)$) et non-vide (car $x = a \in \mathcal{A}$), donc $\mathcal{A}$ est un compact non-vide. Etant donné que f est continue sur $\mathcal{A}$ d'après question 1, on en déduit que (P) possède au moins une solution optimale d'après Corollaire 1. Or d'après question 2 et Théorème 3b, (P) possède au plus une solution optimale, donc (P) possède une et une seule solution optimale.

Finalement on résoudra le problème.

4a. Faire une représentation graphique du problème pour le cas particulier où $n = 2$: l'ensemble admissible $\mathcal{A}$ et les courbes de niveaux $C_k := \{x \in \mathbb{R}^2 : f(x) = k\}$ avec $k \in \{1/4, 1, 4\}$.

Corrigée :
Dans le cas particulier où $n = 2$ on a:

$$f(x, y) = \frac{1}{x^2 y^2} \quad \text{et} \quad \mathcal{A} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : (x-1)^2 + (y-1)^2 \leq \frac{1}{4},\ x + y \leq \frac{5}{2} \right\}.$$



4b. Montrer que les contraintes sont qualifiées en tout point admissible.

Corrigée en utilisant Slater (le plus efficace) :

* f est convexe sur $\Omega = \{x \in \mathbb{R}^n : x_i > 0, 1 \leq i \leq n\}$;
* il n'y a pas de contraintes d'égalite ;
* les fonctions $h_1(x) = \|x - a\|^2 - r^2$ et $h_2(x) = a^T x - b$ sont convexes et de classe $\mathcal{C}^1$;
* h_2 est affine et pour $x_0 = a$ on a $x_0 \in \mathcal{A} \cap \Omega$ et $h_1(x_0) = -r^2 < 0$.

Les contraintes sont donc qualifiées en tout point admissible d'après Slater.

Corrigée en utilisant Fiacco-McCormick (pas du tout efficace) :
Soient $h_1(x) = \|x - a\|^2 - r^2$ et $h_2(x) = a^T x - b$. Alors

- les contraintes sont bien qualifiées en tout point $x \in \text{Int}(\mathcal{A}) = \{X \in \mathcal{A} : h_1(X) < 0 \wedge h_2(X) < 0\}$;
- d'après Fiacco les contraintes sont bien qualifiées en tout point $x \in \mathcal{A}_1 := \{X \in \mathcal{A} : h_1(X) = 0 \wedge h_2(X) < 0\}$, car le système

$$\left\{ \begin{array}{l} z\nabla h_1(x) = 0 \\ x \in \mathcal{A}_1, \ z \neq 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} 2z(x-a)^T = 0 \\ \|x - a\| = r \\ a^T x < b \\ z \neq 0 \end{array} \right.$$

  n'a pas de solution car $\|x - a\| = r \Rightarrow (x - a)^T \neq 0 \Rightarrow z = 0$;
- d'après Fiacco les contraintes sont bien qualifiées en tout point $x \in \mathcal{A}_2 = \{X \in \mathcal{A} : h_1(X) < 0 \wedge h_2(X) = 0\}$, car le système

$$\left\{ \begin{array}{l} z\nabla h_2(x) = 0 \\ x \in \mathcal{A}_2, \ z \neq 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} za^T = 0 \\ a^T x = b \\ \|x - a\| < r \\ z \neq 0 \end{array} \right.$$

  n'a pas de solution car $a^T \neq 0 \Rightarrow z = 0$;
- d'après Fiacco les contraintes sont bien qualifiées en tout point $x \in \mathcal{A}_3 = \{X \in \mathcal{A} :$

$h_1(X) = 0 \wedge h_2(X) = 0\}$, car le système

$$\left\{\begin{array}{l} z_1 \nabla h_1(x) + z_2 \nabla h_2(x) = 0 \\ x \in \mathcal{A}_3,\ z \neq 0 \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{l} 2z_1(x-a)^T + z_2 a^T = 0 \\ \|x-a\|^2 = r^2 \\ a^T x = b \\ (z_1, z_2) \neq (0,0) \end{array}\right. \qquad \Leftrightarrow \left\{\begin{array}{l} z_1 x = \tilde{z}_2 a, \quad \tilde{z}_2 = (z_2/2 - z_1) \\ \|x-a\|^2 = r^2 \\ a^T x = b \\ (z_1, z_2) \neq (0,0) \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{l} x = \lambda a, \quad \lambda = \frac{\tilde{z}_2}{z_1} \quad \text{(il faut que } z_1 \neq 0, \text{ ainsi que } (z_1, z_2) \neq (0,0) \text{)} \\ r^2 = |\lambda - 1|^2 \|a\|^2 = n|\lambda - 1|^2 \\ b = \lambda \|a\|^2 = n\lambda \end{array}\right.$$

$$\Leftrightarrow \left\{\begin{array}{l} x = \lambda a \\ |\lambda - 1| = \frac{r}{\sqrt{n}} = \frac{1}{2\sqrt{n}} \\ \lambda = \frac{b}{n} = 1 + \frac{1}{2n} \end{array}\right.$$

n'a pas de solution car $\lambda = 1 + \frac{1}{2n} \Rightarrow |\lambda - 1| = \frac{1}{2n} \neq \frac{1}{2\sqrt{n}}$.

On en déduit que les contraintes sont bien qualifiées en tout point admissible.

4c. Ecrire les conditions KKT pour le problème (P).

Corrigée :
La fonction de Lagrange est donné par

$$L(x, \mu_1, \mu_2) = f(x) - \mu_1(\|x-a\|^2 - r^2) - \mu_2(a^T x - b).$$

Les conditions KKT deviennent donc

$$\left\{\begin{array}{l} -f(x)y^T - 2\mu_1(x-a)^T - \mu_2 a^T = 0 \\ \mu_1, \mu_2 \leq 0 \\ \mu_1(\|x-a\|^2 - r^2) \leq 0 \\ \mu_2(a^T x - b) \leq 0 \end{array}\right.$$

où $y = (2/x_1, \ldots, 2/x_n)^T$.

4d. Montrer que $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)^T$, avec $\bar{x}_i = 1 + \frac{1}{2n}$, $1 \leq i \leq n$, est un point admissible vérifiant les conditions KKT.

Corrigée :
On a

$$h_1(\bar{x}) = n\left(\frac{1}{2n}\right)^2 - \frac{1}{4} = \frac{1}{4n} - \frac{1}{4} = \frac{1-n}{4n} < 0$$

et

$$h_2(\bar{x}) = n\left(1 + \frac{1}{2n}\right) - \left(n + \frac{1}{2}\right) = 0,$$

donc $\bar{x}$ est bien un point admissible. Puis on remarque que h_1 n'est pas saturée au point $\bar{x}$ (alors on pose $\mu_1 = 0$ ainsi que $\mu_1 h_1(\bar{x}) = 0$) et que h_2 est saturée au point $\bar{x}$ (alors $\mu_2 h_2(\bar{x}) = 0$ quelque soit la valeur de μ_2). Pour que $\bar{x}$ vérifie les conditions KKT, il faut qu'il existe $\mu_2 \leq 0$ tq $\nabla_x L(\bar{x}, 0, \mu_2) = 0$. Soit $\bar{y} = (2/\bar{x}_1, \ldots, 2/\bar{x}_n)^T$. Alors

$$\begin{aligned} \nabla_x L(\bar{x}, 0, \mu_2) = -f(\bar{x})\bar{y}^T - \mu_2 a^T = 0 \Leftrightarrow \mu_2 = -\frac{f(\bar{x})}{a_i}\bar{y}_i, \quad 1 \leq i \leq n \\ \Leftrightarrow \mu_2 = -f(\bar{x})\frac{4n}{2n+1} < 0 \end{aligned}$$

car $f(\bar{x}) > 0$. On en déduit que $\bar{x}$ vérifie les conditions KKT.

4e. Conclure.

Corrigée :
Comme (P) et un problème strictement convexe d'après question 2, et puisque $\bar{x}$ est un point admissible vérifiant les conditions KKT d'après question 4d, on a que $\bar{x}$ est l'unique solution de (P) d'après Théorème 8 et question 3. La valeur optimale vaut

$$f(\bar{x}) = \left(\frac{1}{1 + (1/2n)}\right)^{2n} = \left(\frac{2n}{2n+1}\right)^{2n}.$$