TD 3 : Problème géométrique dual et méthode des moindres carrés

lionel.rieg@ensiie.fr

Exercice 1

On considère le programme géométrique suivant :

$$\min_{x>0, y>0} g \quad \text{avec} \quad g(x,y) = x + 2\frac{y}{x^2} + \frac{2}{y} \tag{PG}$$

1. Calculer le hessien de g au point de coordonnées $(x, y) = (1, 2)$. En déduire que g n'est pas convexe.
2. Écrire le dual géométrique (DPG) de (PG).
3. Résoudre (DPG).
4. En déduire la solution de (PG).
5. Vérifier que la solution trouvée est bien un point critique de g.

Exercice 2

On cherche à transporter 400 m^3 de graviers d'un bord à l'autre d'une rivière. Les graviers seront transportés dans une boîte parallélépipédique de longueur t_1, de largeur t_2 et de hauteur t_3. Les côtés dans le sens de la longueur et le fond de la boîte coûtent 10 unités monétaires par mètre carré. Les côtés dans le sens de la largeur coûtent 20 unités monétaires par mètre carré. La boîte est sans couvercle. Chaque traversée aller-retour de la rivière coûte 0,1 unité monétaire. Le coût de transport du gravier est la somme du coût de construction de la boîte et du coût de la traversée de la rivière. On veut minimiser ce coût.

1. On admettra dans un premier temps que le nombre de traversées aller-retour est égal au volume de gravier (400 m^3) divisé par le volume de la boîte.
   (a) Modéliser le problème de la recherche du coût de transport minimum comme un programme géométrique.
   (b) Écrire son dual.
   (c) Le résoudre et en déduire la solution optimale du programme géométrique primal.
2. On peut vérifier que dans la situation trouvée précédemment, le volume du tas de gravier divisé par le volume de la boîte est un nombre entier. Montrer par un argument simple qu'il en est toujours ainsi.

Exercice 3

Soit A une matrice de m lignes et n colonnes avec $m > n$, b un vecteur colonne de m coordonnées. Le système d'équations $Ax = b$ comporte plus d'équations que d'inconnus ($m > n$) et est ouvert sans solution. La méthode des moindres carrés consiste à rechercher x qui minimise $\|Ax - b\|^2$. L'objectif est de trouver un x qui satisfait (au mieux) les équations.

1. Montrer l'égalité $\|Ax - b\|^2 = x.A^TAx - 2A^Tb.x + b.b$.
2. Montrer que cette fonction est convexe.
3. Montrer que le x qui minimise $\|Ax - b\|^2$ vérifie l'équation dite normale $A^TAx = A^Tb$.

4. Application :
   L'ingénieur Émile travaille sous les ordres de son supérieur Roger. Celui-ci lui a demandé les hauteurs h_1, h_2, h_3 au dessus du niveau de la mer de 3 collines numérotées 1, 2 et 3. Il se place alors au niveau de la mer et mesure les hauteurs $h_1 = 36$, $h_2 = 41$, $h_3 = 117$ (en mètres). Pour vérifier ses mesures, Émile monte sur la colline 1 et mesure la hauteur de la colline 2 par rapport à la colline 1. Il trouve 11 mètres. Il mesure de même la hauteur de la colline 3 au dessus de la colline 1 et il trouve 77 mètres. Il constate que ces dernières mesures ne sont pas compatibles avec les précédentes. Il tente alors une autre expérience et monte sur la colline 2 d'où il mesure la hauteur de la colline 3 au dessus de la colline 2. Il trouve 75 mètres. À nouveau, il constate l'inconsistance des résultats avec les précédents. Émile sait que son supérieur Roger ne se satisfera pas de ces résultats contradictoires et il est assez inquiet sur le chemin du retour au bureau. Soudain, il se souvient que son supérieur a l'esprit cartésien et qu'il partage avec lui son goût pour les méthodes mathématiques. Il décide alors de résoudre son problème en estimant les hauteurs au dessus du niveau de la mer h_1, h_2, h_3 à l'aide de la méthode des moindres carrés. Hélas il ne se souvient plus très bien de la méthode. Pouvez-vous lui résoudre son problème pour qu'il puisse rentrer tranquillement au bureau. Lui donner les résultats exacts sous forme de nombres décimaux.

Exercice 4

On doit déterminer les poids x et y de deux objets A et B. On met A sur la balance et on lit 1. On met B sur la balance et on lit 1. On met A et B sur la balance et on lit 1.

1. Écrire le système d'équations que vérifient x et y.
2. Visiblement, la balance n'est pas fiable mais il faut tout de même décider du poids de A et B. Donner les poids de A et B qui minimisent l'erreur de la balance.

Exercice 5

On veut positionner une bille dans le fond d'une boîte rectangulaire, le plus près possible du coin inférieur gauche, mais pas trop près des bords. On a pris pour origine le coin inférieur gauche et les bords incidents à ce coin comme axes d'un repère orthogonal. Dans ce repère, les coordonnées de la bille sont x et y avec $x > 0$ et $y > 0$. On a établi que la position idéale de la bille est la solution du problème (PG) suivant :

$$\min_{x>0,y>0} g(x,y) = x^2 + y^2 + \frac{2}{xy} \tag{PG}$$

1. Écrire (DPG) le dual géométrique de (PG).
2. Résoudre (DPG).
3. En déduire la solution de (PG).
4. Vérifier que la solution trouvée est bien un point critique de g.

Solutions

▶ Exercice 1

1. On a

$$\nabla g(x,y) = \begin{pmatrix} 1 - \frac{4y}{x^3} \\ \frac{2}{x^2} - \frac{2}{y^2} \end{pmatrix} \qquad \text{puis} \qquad \mathcal{H}(g)(x,y) = \begin{pmatrix} \frac{12y}{x^4} & \frac{-4}{x^3} \\ \frac{-4}{x^3} & \frac{4}{y^3} \end{pmatrix} \qquad \text{donc} \qquad \mathcal{H}(g)(1,2) = \begin{pmatrix} 24 & -4 \\ -4 & \frac{1}{2} \end{pmatrix}$$

On a $\Delta_n = -4$ et la dimension $n = 2$ est paire donc $\mathcal{H}(g)(1,2)$ est indéfini. En particulier, g n'est pas convexe.

2. Rappel : le dual géométrique du problème de minimisation d'un posinôme $g = \sum_{i=1}^m u_i(x)$ où $u_i(x) = c_i \prod_{j=1}^n x_j^{\alpha_{ij}}$ avec $c_i > 0$ et α_{ij} réels quelconques est :

$$\max v(\delta) = \prod_{i=1}^m \left(\frac{c_i}{\delta_i}\right)^{\delta_i} \quad \text{sous les contraintes} \quad \begin{cases} \delta_i > 0 & 1 \leqslant i \leqslant m \\ \sum_{i=1}^m \delta_i = 1 & \\ \sum_{i=1}^m \alpha_{ij}\delta_i = 0 & 1 \leqslant j \leqslant n \end{cases} \tag{DPG}$$

Le théorème de dualité donne alors $v(\delta^*) = g(x^*)$ et $u_i(x^*) = \delta_i^* g(x^*)$.

Ici, cela donne : (DPG) $\max v(\delta)$ avec $v(\delta) = \left(\frac{1}{\delta_1}\right)^{\delta_1} \left(\frac{2}{\delta_2}\right)^{\delta_2} \left(\frac{2}{\delta_3}\right)^{\delta_3}$
sous les contraintes $\delta_i > 0$, $\delta_1 + \delta_2 + \delta_3 = 1$, $\delta_1 - 2\delta_2 = 0$, $\delta_2 - \delta_3 = 0$.

3. Les contraintes du dual donnent $\delta_1 = \frac{1}{2}$, $\delta_2 = \delta_3 = \frac{1}{4}$ et $\max v(\delta) = (2)^{\frac{1}{2}}(2^3)^{\frac{1}{4}}(2^3)^{\frac{1}{4}} = 2^{\frac{1}{2}+\frac{3}{4}+\frac{3}{4}} = 2^2 = 4$.

4. On a donc $g(x^*) = v(\delta^*) = 4$. Les relations $u_i(x^*) = \delta_i \cdot g(x^*)$ donnent $x = \frac{1}{2} \cdot 4$, $\frac{2y}{x^2} = \frac{1}{4} \cdot 4$ et $\frac{2}{y} = \frac{1}{4} \cdot 4$. On en tire $x^* = 2$ et $y^* = 2$.

5. Remplaçant x^* et y^* dans ∇g, on a bien $\nabla g(2,2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

▶ Exercice 2

1. (a) La boîte contient deux côtés dans la longueur, deux dans la largeur et un fond. Le coût de construction est donc $2 \cdot 10t_1t_3 + 2 \cdot 20t_2t_3 + 1 \cdot 10t_1t_2 = 20t_1t_3 + 40t_2t_3 + 10t_1t_2$.
Le coût de transport dépend du nombre de voyage qui est égal (par hypothèse) au volume de graviers à déplacer divisé par le volume de la boîte, *i.e.* $\frac{400}{t_1t_2t_3}$.
Le coût total de la traversée est donc $20t_1t_3 + 40t_2t_3 + 10t_1t_2 + \frac{40}{t_1t_2t_3}$, c'est la fonction objectif qu'on cherche à minimiser.

   (b) Le problème dual est

$$\max \left(\frac{20}{\delta_1}\right)^{\delta_1} \left(\frac{40}{\delta_2}\right)^{\delta_2} \left(\frac{10}{\delta_3}\right)^{\delta_3} \left(\frac{40}{\delta_4}\right)^{\delta_4} \quad \text{sous les contraintes} \quad \begin{cases} \delta_i > 0 \\ \delta_1 + \delta_2 + \delta_3 + \delta_4 = 1 \\ \delta_1 + \delta_3 - \delta_4 = 0 \\ \delta_2 + \delta_3 - \delta_4 = 0 \\ \delta_1 + \delta_2 - \delta_4 = 0 \end{cases}$$

   (c) On obtient facilement $\delta_1 = \delta_2 = \delta_3$ puis $\delta_4 = 2\delta_1$. Ceci donne donc $\delta_1 = \delta_2 = \delta_3 = \frac{1}{5}$ et $\delta_4 = \frac{2}{5}$. Le maximum est alors $100^{\frac{1}{5}}200^{\frac{1}{5}}50^{\frac{1}{5}}100^{\frac{2}{5}} = 100^{\frac{5}{5}}(1 \cdot 2 \cdot \frac{1}{2})^{\frac{1}{5}} = 100$. Le minimum du problème primal est donc 100. On en tire aussi le système suivant :

$$\begin{cases} 20t_1t_3 = \frac{1}{5}100 \\ 40t_2t_3 = \frac{1}{5}100 \\ 10t_1t_2 = \frac{1}{5}100 \\ \frac{40}{t_1t_2t_3} = \frac{2}{5}100 \end{cases} \iff \begin{cases} t_1t_3 = 1 \\ t_2t_3 = \frac{1}{2} \\ t_1t_2 = 2 \\ t_1t_2t_3 = 1 \end{cases} \iff \begin{cases} t_3 = \frac{1}{2} \\ t_1 = 2 \\ t_2 = 1 \end{cases}$$

2. Si la boîte n'était pas pleine sur l'un des voyages, on pourrait réduire la taille de la boîte de sorte que l'excédent des autres voyages remplisse la boîte pour ce voyage. Dans ce cas, le nombre de voyage serait le même (donc même coût de traversée) et la boîte serait plus petite (donc le coût de construction serait plus petit). Ainsi, la solution ne serait plus optimale.

▶ **Exercice 3**

1.

$$\begin{aligned} \|Ax-b\|^2 &= (Ax-b).(Ax-b) \\ &= Ax.Ax - Ax.b - b.Ax + b.b \\ &= Ax.Ax - 2b.Ax + b.b \\ &= x.A^TAx - 2A^Tb.x + b.b \end{aligned}$$

2. Puisqu'on cherche à minimiser, on peut ignorer le terme constant $b.b$ et se restreindre à étudier $x.A^TAx - 2A^Tb.x$. Pour montrer que cette fonction est convexe, il suffit de montrer que son hessien est semi-défini positif puisqu'elle est quadratique (car A^TA est symétrique). Comme c'est une fonction quadratique, on peut directement lire son hessien : $\mathcal{H}(\|Ax-b\|^2)(x) = 2A^TA$. On cherche donc a montrer que A^TA est semi-défini positif. Par définition, Q semi-défini positif veut dire que pour tout vecteur x, $x.Qx \geq 0$. Enfin, $x.A^TAx = Ax.Ax = \|Ax\|$ qui est clairement positif car c'est une norme.
3. Pour une fonction quadratique de la forme $x.Qx + c.x$, gradient vaut $Ax - b$. Or un minimum doit annuler le gradient, ce qui donne ici : $2A^TAx - 2A^Tb = 0$, *i.e.* $A^TAx = A^Tb$.
4. Notons h_1, h_2 et h_3 les hauteurs des trois collines et h leur vecteur. On regarde les mesures comme une opération sur ces hauteurs : Ah. Ou bien, on écrit le système sans solution qui devrait satisfaire ces hauteurs et on en tire A et b Donnons donc A et b et calculons A^TA et A^Tb :

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 36 \\ 41 \\ 117 \\ 11 \\ 77 \\ 75 \end{pmatrix} \qquad \text{puis} \qquad A^TA = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix} \quad \text{et} \quad A^Tb = \begin{pmatrix} -52 \\ -23 \\ 269 \end{pmatrix}$$

La résolution de l'équation normale donne :

$$\begin{cases} 3h_1 - h_2 - h_3 = -52 \\ -h_1 + 3h_2 - h_3 = -23 \\ -h_1 - h_2 + 3h_3 = 269 \end{cases} \implies h_1 + h_2 + h_3 = 194 \implies \begin{cases} 4h_1 = 142 \\ 4h_2 = 171 \\ 4h_3 = 463 \end{cases} \iff \begin{cases} h_1 = \frac{142}{4} = 35,5 \\ h_2 = \frac{171}{4} = 42,75 \\ h_3 = \frac{463}{4} = 115,75 \end{cases}$$

▶ **Exercice 4** 1. Le système d'équation est :

$$\begin{cases} x = 1 \\ y = 1x + y = 1 \end{cases}$$

2. En reprenant les notation de l'exercice précédent, on a

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \text{d'où} \qquad A^T A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{et} \quad A^T b = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

La résolution du système $A^T A(x^*, y^*)^T = A^T b$ donne $x^* = y^* = \frac{2}{3}$.

▶ **Exercice 5**

1. Sous les contraintes

$$\begin{aligned} \delta_1 + \delta_2 + \delta_3 &= 1 \\ 2\delta_1 - \delta_3 &= 0 \\ 2\delta_2 - \delta_3 &= 0 \\ \delta_i &> 0 \end{aligned}$$

le problème dual est

$$\max \prod \left(\frac{2}{\delta_1}\right)^{\delta_1} \left(\frac{2}{\delta_2}\right)^{\delta_2} \left(\frac{-1}{\delta_3}\right)^{\delta_3} \tag{DPG}$$

2. Les contraintes donnent $\delta_1 = \delta_2 = \frac{1}{4}$ et $\delta_3 = \frac{1}{2}$.
   On obtient donc la solution $4^{\frac{1}{4}} \cdot 4^{\frac{1}{4}} \cdot 4^{\frac{1}{2}} = 4$.
3. On a $g(x^*, y^*) = 4$ et les relation par terme $u_i(x^*, y^*) = \delta_i g(x^*, y^*)$ donnent $x^2 = 1$, $y^2 = 1$ et $\frac{2}{xy} = 2$.
   On en tire $x = y = -1$ ou $x = y = 1$.
4. Le gradient de g est

$$\begin{pmatrix} 2x - \frac{2}{x^2 y} \\ 2y - \frac{2}{xy^2} \end{pmatrix}$$

On vérifie facilement qu'il s'annule en $(1, 1)$ et $(-1, -1)$.