Contrôle MOM (Avril 2011)

Documents autorisés. Pas de calculette. 1 heure.

Programmation géométrique

On considère le programme géométrique (PG) suivant :

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

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

Méthode du gradient conjugué

Soit $f(x_1, x_2) = x_1^2 - x_1 x_2 + x_2^2 - 3x_2$ avec $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ dans $\mathbb{R}^2$.

On considère le problème : $\min_{(x_1,x_2)\in R^2} f(x_1, x_2)$

Soit $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, la valeur de f en $x^{(0)}$ est $f\left(x^{(0)}\right) = 0$.

1. Calculer ∇f le gradient de f.
2. En partant de $x^{(0)}$, appliquer la méthode du gradient conjugué (2 itérations). Pour calculer $\beta^{(k)}$ à l'itération k, on pourra utiliser $\beta^{(k)} = \frac{\left\|\nabla f\left(x^{(k+1)}\right)\right\|^2}{\left\|\nabla f\left(x^{(k)}\right)\right\|^2}$. Calculer la valeur de f en chaque point parcouru. Vérifier que le point trouvé à la 2$^{\text{ème}}$ itération est bien un point critique de f.

Programmation géométrique

$\min_{\substack{x>0\\y>0}} g(x,y) = x + 2\dfrac{y}{x^2} + \dfrac{2}{y}$

1) calcul hessien. montrer que g n'est pas convexe
2) Résoudre le pb à l'aide du dual géométrique
3) Vérifier que la solution est un point critique de g.

1) $\nabla g = \begin{bmatrix} 1 - \frac{4y}{x^3} \\ \frac{2}{x^2} - \frac{2}{y^2} \end{bmatrix}$ $\quad H_g = \begin{bmatrix} \frac{12y}{x^4} & -\frac{4}{x^3} \\ -\frac{4}{x^3} & \frac{4}{y^3} \end{bmatrix}$

$\Delta_1 = \dfrac{12y}{x^4} > 0$

$\Delta_2 = \dfrac{12\times 4}{x^4 y^2} - \dfrac{4\times 4}{x^6} = \dfrac{16}{x^4}\left[\dfrac{3}{y^2} - \dfrac{1}{x^2}\right]$

$\boxed{2}$

en $\begin{matrix} x=1 \\ y=2 \end{matrix}$ on obtient $\Delta_2 = 16\left[\frac{3}{4} - 1\right] = -4 < 0$

comme Δ_2 est le produit des 2 valeurs propres $\lambda_1 \lambda_2$ il y en a forcément une qui est <0.
donc H_g n'est pas semi-défini positive partout.
donc g n'est pas convexe.

2) g n'étant pas convexe, on décide d'attaquer le pb à l'aide du dual.

$\max v = \left(\dfrac{1}{\delta_1}\right)^{\delta_1} \left(\dfrac{2}{\delta_2}\right)^{\delta_2} \left(\dfrac{2}{\delta_3}\right)^{\delta_3}$ s.c. $\begin{cases} \delta_1 + \delta_2 + \delta_3 = 1 \\ \delta_1 - 2\delta_2 = 0 & \leftarrow (x) \\ \delta_2 - \delta_3 = 0 & \leftarrow (y) \\ \delta_i > 0 & i = 1 \text{ à } 3 \end{cases}$

$\boxed{2}$

$\boxed{1}$ une solution unique : $\delta_2 = \delta_3$, $2\delta_2 + \delta_2 + \delta_2 = 1 \Rightarrow \delta_2 = \frac{1}{4}$, $\delta_3 = \frac{1}{4}$, $\delta_1 = \frac{1}{2}$

$\boxed{1}$ $v = (2)^{1/2} (2\times 4)^{1/4} (2\times 4)^{1/4} = 2^{1/2 + \frac{1}{4} + \frac{1}{4}} \times 4^{1/4 + 1/4} = 2 \times 4^{1/2} = 2\times 2 = 4$

on en déduit $u_i = \delta_i \times v$ $\quad i = 1,2,3$

$\boxed{2+1}$ $x = \frac{1}{2}\times 4 = 2$, $\frac{2y}{x^2} = \frac{1}{4}\times 4 = 1$, $\frac{2}{y} = \frac{1}{4}\times 4 = 1$ $\Rightarrow x = 2, y = 2$

$g(x,y) = 2 + 2\times\frac{2}{4} + \frac{2}{2} = 4$ on a bien $g = v$

3) on reporte dans $\nabla g = \begin{pmatrix} 1 - \frac{4\times 2}{2^3} \\ \frac{2}{2^2} - \frac{2}{2^2} \end{pmatrix}$

$\boxed{1}$

ce qui fait $\nabla g = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Programmation géométrique (Strasbourg mai 2011)

(PG) $\min \ g(x,y) = \frac{1}{2}x + \frac{1}{4}\frac{y}{x^2} + \frac{1}{4}\frac{1}{y}$

1) Ecrire (DPG) le dual de (PG)

$$\max_{\delta} \left(\frac{1/2}{\delta_1}\right)^{\delta_1} \times \left(\frac{1/4}{\delta_2}\right)^{\delta_2} \times \left(\frac{1/4}{\delta_3}\right)^{\delta_3} \quad \text{s.c.} \begin{cases} \delta_1 - 2\delta_2 = 0 & (x) \\ \delta_2 - \delta_3 = 0 & (y) \\ \delta_1 + \delta_2 + \delta_3 = 1 \\ \delta_i > 0 \quad i = 1 \text{ à } 3 \end{cases}$$

2

2) Résoudre (DPG)

$$\left.\begin{array}{l} \delta_1 = 2\delta_2 \\ \delta_2 = \delta_3 \end{array}\right\} \Rightarrow 2\delta_2 + \delta_2 + \delta_2 = 1 \Rightarrow \delta_2 = \frac{1}{4} \Rightarrow \delta_1 = \frac{1}{2} \qquad \text{solution unique}$$

2

$v(\delta) = 1$

3) En déduire la solution de (PG)

on a $\delta_i = \frac{u_i(x^*)}{g(x^*)}$ $\quad i = 1$ à 3 $\quad$ on pose $g(x^*) = v(\delta) = 1$ $\quad$ (pas de saut de dualité)

$\Rightarrow \frac{1}{2}x = \frac{1}{2} \Rightarrow x = 1 \qquad (i=1)$

2 $\quad \Rightarrow \frac{1}{4}\frac{y}{x^2} = \frac{1}{4} \Rightarrow y = 1 \qquad (i=2)$

$\Rightarrow \frac{1}{4}y = \frac{1}{4} \Rightarrow y = 1 \qquad (i=3)$

vérifions que $g(x,y) = 1$: $\frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 1$

4) vérifier que la solution est un point critique de g

$\frac{\partial g}{\partial x} = \frac{1}{2} - \frac{y}{2}x^{-3} = 0$

2

$\frac{\partial g}{\partial y} = \frac{1}{4x^2} - \frac{1}{4} \times y^{-2} = 0$

Montrer l'inégalité : $\frac{x}{4} + \frac{3y}{4} \leq \sqrt{\ln\left(\frac{e^{x^2} + 3e^{y^2}}{4}\right)}$

pour $x, y \geq 0$.

(Utiliser la convexité d'une fction bien choisie)

dém :

convexité de e^{z^2}

$$e^{\left(\frac{x}{4} + \frac{3y}{4}\right)^2} \leq \frac{1}{4} e^{x^2} + \frac{3}{4} e^{y^2}$$

$$\left(\frac{x}{4} + \frac{3y}{4}\right)^2 \leq \ln\left(\frac{1}{4}e^{x^2} + \frac{3}{4}e^{y^2}\right) \qquad \text{car } \ln \nearrow$$

$$\frac{x}{4} + \frac{3y}{4} \leq \sqrt{ 4} \qquad \text{car } \sqrt{\ } \nearrow \text{ et } x, y \geq 0$$

faire la démo en partant du résultat. La dernière inégalité se prouve en utilisant la convexité de e^{z^2} avec $z \in \mathbb{R}$

Inégalité de Hölder

Soient p, q réels, plus gds que 1 t.q. $p^{-1} + q^{-1} = 1$

$(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ vecteurs de $\mathbb{R}^n$, alors :

$$\sum_{i=1}^{n} |x_i y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}$$

dém : préliminaire : on a inégalité de Young : $u \geq 0$ et $v \geq 0$ (réels)

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}$$

$\rightarrow$ AG en posant $\boxed{u^p =: x \text{ et } v^q =: y}$ $\Rightarrow \frac{1}{p}x + \frac{1}{q}y \overset{AG}{\geq} x^{1/p} y^{1/q} = uv$

$\frac{u^p}{p} + \frac{v^q}{q}$ $\left(\frac{1}{p} + \frac{1}{q} = 1\right)$

on note : $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$ $\quad \|y\|_q = \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}$

posons $u = \frac{|x_i|}{\|x\|_p}$ et $v = \frac{|y_i|}{\|y\|_q}$ $\quad$ 1er cas supposons $\|x\|_p \neq 0$ et $\|y\|_q \neq 0$

$$\Rightarrow \frac{|x_i y_i|}{\|x\|_p \|y\|_q} \leq \frac{|x_i|^p}{p\|x\|_p^p} + \frac{|y_i|^q}{q\|y\|_q^q}$$

$$\Rightarrow \frac{1}{\|x\|_p \|y\|_q} \cdot \sum |x_i y_i| \leq \frac{1}{p} \cdot \frac{\sum |x_i|^p}{\|x\|_p^p} + \frac{1}{q} \frac{\sum |y_i|^q}{\|y\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1$$

$$\Rightarrow \sum |x_i y_i| \leq \|x\|_p \|y\|_q =$$

2e cas $\|x\|_p = 0$ ou $\|y\|_q = 0$ alors l'inégalité de Hölder est vérifiée de manière évidente.

$\Rightarrow x = 0$ $\quad \Rightarrow y = 0$

[illegible] Inequality. (Hint: Show that for a fixed p the solution of the maximization problem.)

Maximize $\prod_{k=1}^{n} v_k$ subject to $v_k > 0$, $\sum_{k=1}^{n} v_k = p$ is $v_k = p/n$ for $k = 1, \dots, n$. (Hint: Suppose that $v_1, v_2, \dots, v_n$ are positive and that $\sum_{k=1}^{n} v_k = p$. Show that if [illegible] $= p$ and [illegible]

$(v_1 - v_2)^2 > 0$ si $v_1 \neq v_2$ $\Rightarrow$ $v_1^2 + v_2^2 > 2 v_1 v_2$ $\Rightarrow$ $v_1^2 + v_2^2 + 2 v_1 v_2 > 4 v_1 v_2$

$(v_1 + v_2)^2 > 4 v_1 v_2$

donc $v_1 \neq v_2 \Rightarrow v_1 v_2 < \left(\frac{v_1 + v_2}{2}\right)^2 = \bar{v}_1 \bar{v}_2$

donc on peut supposer que la solution (optimale) de maximiser $\prod_{k=1}^{m} v_k$ s.c. $\sum_{k=1}^{m} v_k = p$, $v_k > 0$

est $v_1^* = v_2^* = \dots = v_m^* = \frac{p}{m}$. Donc $\forall v_k$ t.q. $v_k > 0$, soit $p \overset{\text{déf.}}{=} \sum_{k=1}^{m} v_k$, on a :

$$\prod_{k=1}^{m} v_k \leq \prod_{k=1}^{m} v_k^* = \left(\frac{p}{m}\right)^m \Rightarrow \left(\prod_{k=1}^{m} v_k\right)^{1/m} = \prod_k v_k^{1/m} \leq \frac{p}{m} = \sum_k \frac{v_k}{m}$$

[illegible]

f convexe $\Rightarrow \forall x, y$

$$f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$$
$$f(x) \geq f(y) + \nabla f(y) \cdot (x - y)$$

soit $\lambda \in]0,1[$

$$f((1-\lambda)x + \lambda y) \overset{(1)}{\leq} (1-\lambda) f(x) + \lambda f(y) = f(x) + \lambda (f(y) - f(x)) \leq f(x) + \lambda \nabla f(y) \cdot (y - x)$$

$$f((1-\lambda)x + \lambda y) = f(x + \lambda(y - x)) \geq f(x) + \lambda \underbrace{\nabla f(x) \cdot (y - x) = f(x) + \lambda \nabla f(y)}_{= \text{par hypothèse}} \cdot (y - x)$$

$\Rightarrow$ (1) est une égalité $\Rightarrow$ $x = y$ car f strict. convexe.

<u>Plus simple:</u> on démontre $x \neq y \Rightarrow \nabla f(x) \neq \nabla f(y)$

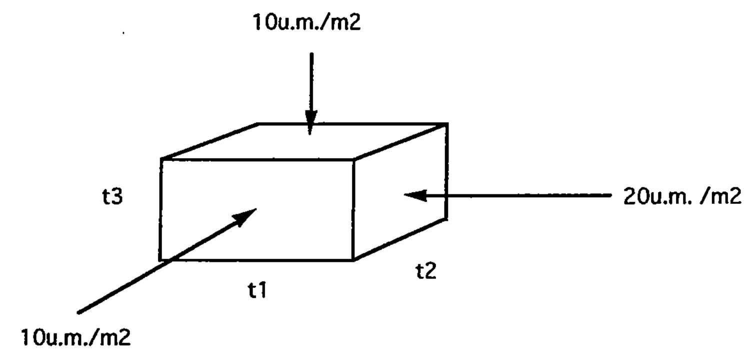
f strict. convexe $\Leftrightarrow \forall x \neq y$
$$\left.\begin{array}{l} f(y) > f(x) + \nabla f(x) \cdot (y - x) \\ f(x) > f(y) + \nabla f(y) \cdot (x - y) \end{array}\right\} \oplus \rightarrow 0 > (\nabla f(x) - \nabla f(y)) \cdot (y - x)$$
$$\Rightarrow \nabla f(x) \neq \nabla f(y)$$

Programmation géométrique

Exercice 1.
400 mètres cube de gravier doivent être transportés d'un bord à l'autre d'une rivière. Les graviers seront transportés dans une boîte parallélépipèdique 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 cotés dans le sens de la largeur coûtent 20 unités monaitaires par mètre carré. La boîte est sans couvercle. Chaque traversée aller-retour de la rivière coûte 0,1 unités monaitaires. Le coût de transport du gravier est la somme du coût de construction de la boîte et du coût de traversée de la rivière. On veut minimiser ce coût de transport.

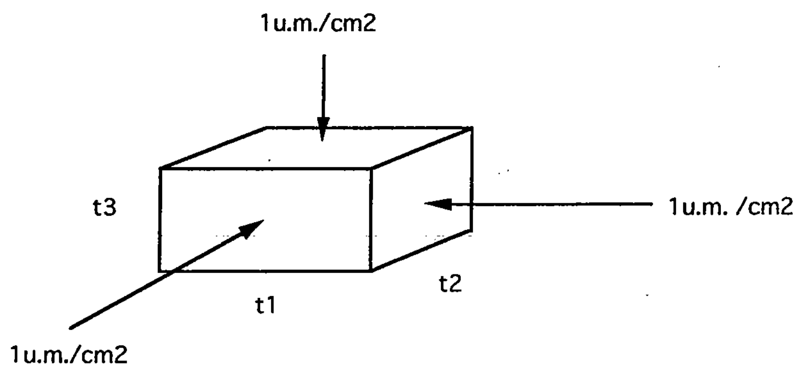
1) On admettra dans un premier temps que le nombre de traversées aller-retour est égal au volume de gravier (400 m^3) divisé par le volume de la boîte.
Modéliser le problème de la recherche du coût de transport minimum comme un programme géométrique.
Ecrire son dual.
Le résoudre et en déduire la solution optimale du programme géométrique primal.

2) On peut vérifier que dans la solution trouvée précédemment, le volume du tas de gravier divisé par le volume de la boite est un nombre entier.
Montrer par un argument simple qu'il en est toujours ainsi.



Exercice 2.

Une boîte parallélépipèdique doit être construite. On désigne la longueur t_1, la largeur t_2 et la hauteur t_3. La boîte est sans couvercle. La boite est faite en un métal coutant 1 unité monaitaire par centimètre carré. On veut construire une boîte de coût minimum et de volume au moins 10 centimètres cube.
1) Donner la formulation mathématique du problème.
2) On a à résoudre un programme géométrique comportant une contrainte d'inégalité. Montrer que nécessairement toute solution optimale "sature" la contrainte d'inégalité c'est-à-dire la vérifie avec égalité. Se ramener alors à un programme géométrique sans contraintes en les variables t_1, t_2 uniquement.
3) Résoudre le dual de ce programme géométrique sans contraintes et en déduire la solution du primal puis donner les dimensions optimales de la boîte.



Contrôle Optimisation Mathématique
(Septembre de l'an 2001)

Tout document autorisé.
Calculette électronique interdite.

Problème 1 (sur 7)

Une boîte rectangulaire doit être construite. On désigne par x_1, x_2, x_3 les dimensions de la boîte: x_3 est la hauteur de la boîte, x_1 et x_2 sont les dimensions du fond de la boîte.
La boîte n'a pas de couvercle. Les côtés et le fond de la boîte seront fait en un métal coûtant 1F au centimètre carré. On veut construire une boîte de coût minimum et de volume au moins 10 centimètres cube.

1) Formuler le problème.
2) On a à résoudre un [redacted] comportant une contrainte. Montrer que nécessairement toute solution optimale vérifie $x_3 = \frac{10}{(x_1 x_2)}$. Se ramener alors à un programme géométrique sans contrainte en les variables x_1 et x_2 que l'on formulera.
3) Résoudre le dual (géométrique) de ce programme géométrique sans contrainte et en déduire la solution du primal puis donner les dimensions optimales de la boîte.

Problème 2 (sur 13)

Un nouvel entrepôt doit être placé de sorte que la somme des carrés des distances à 4 entrepôts déjà existant soit minimisée.
Les 4 entrepôts sont localisés aux points (1,2), (-2,4), (2,6) et (-6,-3).
On note x_1 , x_2 les coordonnées du nouvel entrepôt.
I.1) Formuler le problème.
I.2) On note f la fonction à minimiser. Montrer que f est convexe.
I.3) Résoudre le problème en cherchant un point critique de f.

On suppose maintenant que les coordonnées doivent satisfaire la contrainte $x_1 + x_2 \geq 2$. On note (P) le problème résultant c'est-à-dire minimiser $f(x)$ s.c. $x_1 + x_2 \geq 2$.
II.1) Résoudre le problème à l'aide des conditions de Kuhn-Tucker. Donner la valeur du multiplicateur de Lagrange associé à la contrainte.
II.2) On va résoudre le problème par la méthode barrière logarithmique de Frisch.
Trouver $\min_x q(x) = f(x) - \frac{1}{k} \ln(x_1 + x_2 - 2)$ s.c. $x_1 + x_2 > 2$ en cherchant un point critique de $q(x)$ satisfaisant $x_1 + x_2 > 2$. Faisant tendre k vers l'infini, retrouver la solution de (P).
II.3) On va résoudre le problème par la méthode de pénalité de Beltrami. Pour cela, on rajoute une variable s pour transformer la contrainte d'inégalité de (P) en une contrainte d'égalité c'est-à-dire telle que $x_1 + x_2 = s^2 + 2$.
Trouver $\min_{x,s} q(x,s) = f(x) + k(-x_1 - x_2 + s^2 + 2)^2$ en cherchant un point critique de $q(x,s)$.
Faisant tendre k vers l'infini, retrouver la solution de (P).

problème 1

minimiser $x_1x_2 + 2x_2x_3 + 2x_1x_3$ s.c $x_1x_2x_3 \geqslant 10$, $x_i > 0$ $i = 1, 2, 3$ 1 pt

minimiser $x_1x_2 + 2x_3(x_2 + x_1)$ s.c $x_3 \geqslant \frac{10}{x_1x_2}$, $x_i > 0$ $i = 1, 2, 3$

x_3 positive et $x_2 + x_1$ positif $\Rightarrow$ x_3 doit être la plus petite possible soit $x_3 = \frac{10}{x_1x_2}$ 2 pts

$\Rightarrow$ minimiser $\underset{\delta_1}{\underset{\uparrow}{x_1x_2}} + \underset{\delta_2}{\underset{\uparrow}{\frac{20}{x_1}}} + \underset{\delta_3}{\underset{\uparrow}{\frac{20}{x_2}}}$ s.c $x_i > 0$ $i = 1, 2$

dual : $\begin{cases} \delta_1 - \delta_2 = 0 \\ \delta_1 - \delta_3 = 0 \\ \delta_1 + \delta_2 + \delta_3 = 1 \end{cases}$ $\longrightarrow$ $\begin{matrix} \delta_2 = \delta_1 \\ \delta_3 = \delta_1 \\ \delta_1 = \frac{1}{3} \end{matrix}$

$\delta_i > 0$ $i = 1, 2, 3$

$v^* = \prod_i \left(\frac{c_i}{\delta_i}\right)^{\delta_i} = 3^{1/3} \times 20^{1/3}\, 3^{1/3} \times 20^{1/3}\, 3^{1/3} = 3 \times 20^{2/3}$

$c_i(x) = v^* \times \delta_i \rightarrow \begin{cases} x_1x_2 = \frac{1}{3} \times 3 \times 20^{2/3} \\ \frac{20}{x_1} = \frac{1}{3} \times 3 \times 20^{2/3} \longrightarrow x_1 = 20^{1/3} \\ \frac{20}{x_2} = \frac{1}{3} \times 3 \times 20^{2/3} \longrightarrow x_2 = 20^{1/3} \end{cases}$ 4 pts

solution du pb initial : $x_3 = \frac{10}{20^{2/3}} = \frac{20^{1/3}}{2}$, $x_1 = x_2 = 20^{1/3}$

1)

<u>Problème n° 2</u>

minimiser

$$f(x) = (1 - x_1)^2 + (2 - x_2)^2 + (-2 - x_1)^2 + (4 - x_2)^2 + (2 - x_1)^2 + (6 - x_2)^2 + (-6 - x_1)^2 + (-3 - x_2)^2$$

2)

$\|\cdot\|$ est convexe $\begin{cases} \|u+v\| \leq \|u\| + \|v\| \\ \|\lambda u\| = |\lambda| \|u\| \end{cases}$

$u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$, $\vdots$, $u_4 = \begin{pmatrix} -6 \\ -3 \end{pmatrix}$

une somme de fction convexe est convexe : $f_{u_1}(x) + f_{u_2}(x) + \dots + f_{u_4}(x)$

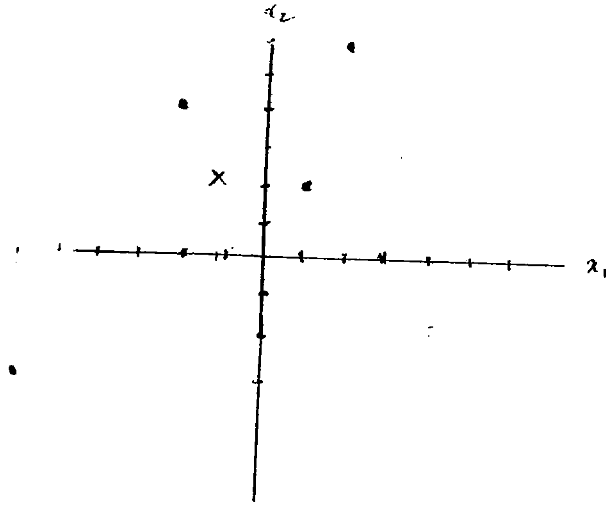
$f_u(x) = \|u - x\|$ est convexe car

$f \circ g$ est convexe si g est affine $(de\ \mathbb{R}^n \to \mathbb{R}^m)$ et f convexe $(\mathbb{R}^m \to \mathbb{R})$: $f\left[g((1-\lambda)x + \lambda y)\right] = f\left[(1-\lambda)g(x) + \lambda g(y)\right] \leq (1-\lambda) f(g(x)) + \lambda f(g(y))$

3)

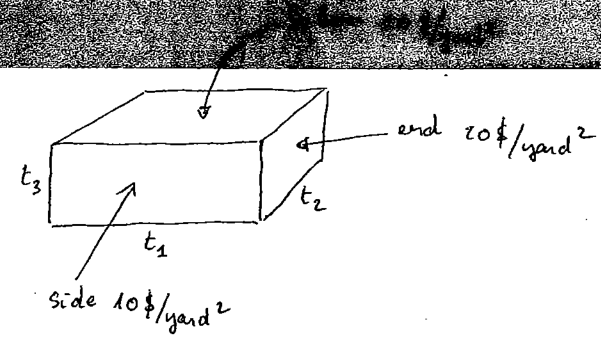
$$\nabla f : \begin{cases} -2(1-x_1) - 2(-2-x_1) - 2(2-x_1) - 2(-6-x_1) = 0 \rightarrow 4x_1 = -5 \\ -2(2-x_2) - 2(4-x_2) - 2(6-x_2) - 2(-3-x_2) = 0 \rightarrow 4x_2 = +9 \end{cases} \quad \begin{cases} x_1 = -5/4 \\ x_2 = 9/4 \end{cases}$$

(solution unique)



$Hf = 4\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ f est strict^t convexe $\rightarrow$ min. unique.

“S[redacted]
Supp[redacted]
height t_3. The sides and bottom of the box cost \$10 per square yard and the ends of the box cost \$20 per square yard. The box will have no salvage value and each round trip of the box on the ferry will cost 10 cents. What is the minimum [redacted]



on remarque la solution optimale sera telle que $\frac{400}{V}$ où V est le volume cherché sera un <u>entier</u> qui représente le nombre de voyages, car pour minimiser le coût de la boite on a jamais intérêt à faire un voyage avec une boite non complètement remplie. En effet dans ce cas mieux vaut diminuer les dimensions de la boite et faire que pour ce voyage (admettons le dernier) la boite soit entièrement remplie.

Le problème se formule alors : minimiser $\frac{40}{t_1t_2t_3} + 40t_2t_3 + 20t_1t_3 + 10t_1t_2$ avec $t_1>0, t_2>0, t_3>0$

dual :
$$\left.\begin{array}{ll} -\delta_1 + \delta_3 + \delta_4 = 0 & \delta_1 + \dots + \delta_4 = 1 \\ -\delta_1 + \delta_2 + \delta_4 = 0 & \delta_1, \delta_2, \delta_3, \delta_4 > 0 \\ -\delta_1 + \delta_2 + \delta_3 = 0 & \end{array}\right\} \Rightarrow \delta^* = \left(\frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$$

$$v(\delta^*) = \left(\frac{5\times 40}{2}\right)^{2/5}(5\times 40)^{1/5}(5\times 20)^{1/5}(5\times 10)^{1/5} = (5\times 20)^{5/5}(2)^{1/5}\left(\frac{1}{2}\right)^{1/5} = 100$$

primal :
$$\frac{40}{t_1t_2t_3} = \frac{2}{5}\times 100 \rightarrow t_1t_2t_3 = 1 \longrightarrow \frac{1}{2t_3} = 1 \rightarrow t_3 = \frac{1}{2}$$
$$40t_2t_3 = \frac{100}{5} = 20 \rightarrow 2t_2t_3 = 1 \rightarrow t_2 = \frac{1}{2t_3} \longrightarrow t_2 = 1$$
$$20t_1t_3 = \frac{100}{5} = 20 \rightarrow t_1t_3 = 1 \rightarrow t_1 = \frac{1}{t_3} \longrightarrow t_1 = 2$$
$$10t_1t_2 = \frac{100}{5} = 20 \rightarrow t_1t_2 = 2$$

Programme géométrique

(PG) minimiser $g(x) = \underset{\delta_1}{\underset{\uparrow}{\frac{x_1}{x_2}}} + \dots + \underset{\delta_{m-1}}{\underset{\uparrow}{\frac{x_{m-1}}{x_m}}} + \underset{\delta_m}{\underset{\uparrow}{\frac{x_m}{x_1}}}$ s.c $x_1 > 0, \dots, x_m > 0$

Le dual de (PG) est : maximiser $v(\delta) = \prod_{i=1}^{m} \left[\frac{1}{\delta_i}\right]^{\delta_i}$ s.c.

$$\begin{cases} \left.\begin{array}{ll} \delta_1 - \delta_m = 0 & (x_1) \\ \delta_2 - \delta_1 = 0 & (x_2) \\ \delta_3 - \delta_2 = 0 & (x_3) \\ \vdots & \\ \delta_m - \delta_{m-1} = 0 & (x_m) \end{array}\right\} (3) \\ \sum_{i=1}^{m} \delta_i = 1 \quad (2) \qquad \delta_i > 0 \quad i = 1 \text{ à } m \quad (1) \end{cases}$$

il y a une solution unique. En effet (3) et (2) $\Rightarrow$ $\delta_i = \frac{1}{m}$ $i = 1$ à m

alors $v(\delta) = \underbrace{m^{1/m} \times \dots \times m^{1/m}}_{m \text{ fois}} = m$

Construire la solution de (PG) : on cherche x t.q. $\delta_i = \frac{u_i(x)}{v(\delta)}$ $i = 1$ à m

c'est-à-dire $\frac{1}{m} = \frac{x_i/x_{i+1}}{m} \Rightarrow 1 = \frac{x_i}{x_{i+1}}$ $i = 1$ à $m-1$ et $1 = \frac{x_m}{x_1}$ $\Rightarrow$ $x_1 = x_2 = \dots = x_m$

alors $g(x) = m = v(\delta)$ on a bien l'optimum de (PG)

ici le primal (PG) a une infinité de solutions

Contrôle MOM (Mai 09)

Documents autorisés. Pas de calculette. 40 minutes.

Programmation géométrique

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 celle solution du problème (PG) suivant :

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

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

Méthodes itératives

Soit $f(x_1,x_2) = x_1^2 + x_1x_2 + \frac{1}{2}x_2^2 - 3x_1 - 2x_2 = \frac{1}{2}x^t\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}x + \begin{pmatrix} -3 \\ -2 \end{pmatrix}^t x$ avec $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

On considère le problème : $\min_{(x_1,x_2)\in R^2} f(x_1,x_2)$

Soit $x^{(0)} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$, la valeur de f en $x^{(0)}$ est $f\left(x^{(0)}\right) = -\frac{3}{2}$.

1. Calculer ∇f le gradient de f.
2. En partant de $x^{(0)}$, appliquer la méthode du gradient conjugué (2 itérations). Calculer la valeur de f en chaque point parcouru. Vérifier que le point trouvé à la 2ème itération est bien un point critique de f.
3. En partant de $x^{(0)}$, appliquer les 2 premières itérations de la méthode de Cauchy (plus forte pente). En remarquant que la 1ère itération est la même que précédemment, on fera seulement la 2ème itération. Calculer la valeur de f en le point obtenu.

Programmation géométrique

On veut positionner une bille dans le fond rectangulaire d'une boîte, 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 orthonormé. Dans ce repère, les coordonnées de la bille sont $\binom{x}{y}$ avec $x>0$ et $y>0$.

On a établi que la position idéale est celle solution du problème suivant :

(PG) $$\underset{x>0,y>0}{\text{minimiser}} \; g(x,y) = x^2 + y^2 + \frac{2}{xy}$$

1) Écrire (DPG) le dual 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

1) contr. de (DPG)

$$\underset{\delta_1}{x^2} + \underset{\delta_2}{y^2} + \underset{\delta_3}{\frac{2}{xy}}$$

$$\begin{cases} 2\delta_1 - \delta_3 = 0 & (x) \\ 2\delta_2 - \delta_3 = 0 & (y) \\ \delta_1 + \delta_2 + \delta_3 = 1 \\ \delta_1, \delta_2, \delta_3 > 0 \end{cases}$$

objectif de (DPG)

$$v(\delta) = \left(\frac{1}{\delta_1}\right)^{\delta_1}\left(\frac{1}{\delta_2}\right)^{\delta_2}\left(\frac{2}{\delta_3}\right)^{\delta_3}$$

2) contr. de (DPG) ⇒ sol. unique.

$(x) \Rightarrow 2\delta_1 = \delta_3$, $(y) \Rightarrow 2\delta_2 = \delta_3$, on reporte dans la 3^e ⇒ $\frac{\delta_3}{2} + \frac{\delta_3}{2} + \delta_3 = 2\delta_3 = 1 \Rightarrow \delta_3 = \frac{1}{2} \Rightarrow \delta_1 = \delta_2 = \frac{1}{4}$

$$v(\delta) = (4)^{1/4}(4)^{1/4}(2\times 2)^{1/2} = 4^{1/4+1/4+1/2} = 4$$

3) on utilise $\delta_i = \frac{u_i}{g}$ $\forall i = 1,2,3$ avec $g = v = 4$

$$\frac{1}{4} = \frac{x^2}{4} \;;\; \frac{1}{4} = \frac{y^2}{4} \;;\; \frac{1}{2} = \frac{2}{xy}\times\frac{1}{4} \Rightarrow x^2 = 1 \;;\; y^2 = 1 \;;\; xy = 1 \Rightarrow x = 1 \;;\; y = 1$$

4) $$\frac{\partial g}{\partial x} = 2x - \frac{2}{x^2 y} = 0$$
$$\frac{\partial g}{\partial y} = 2y - \frac{2}{xy^2} = 0$$

pour $x = 1, y = 1 \Rightarrow \begin{cases} 2 - 2 = 0 \\ 2 - 2 = 0 \end{cases}$

A

OPTIMISATION MATHÉMATIQUE

Durée : **1h30mn**
Documents autorisés : **Cours et notes de TDs**
Calculatrices **interdites**
Nom :
Groupe :

DIRECTIONS CONJUGUÉES

Soit $f(x) = \frac{1}{2}x^T Ax + b^T x$ avec $A = \begin{bmatrix} 4 & -1 \\ -1 & 2 \end{bmatrix}$, $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, et $x \in \Re^2$. On veut résoudre le problème $\min_{x\in\Re^2} f(x)$.

1 (1) On considère $x^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, et $d^{(0)} = -\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ la direction opposée au gradient de f au point $x^{(0)}$. Construire $x^{(1)} = x^{(0)} + t^{(0)}d^{(0)}$ où $t^{(0)}$ minimise $f\left(x^{(0)} + td^{(0)}\right)$ avec $t \in \Re$.

(2) Construction d'une nouvelle direction de descente.

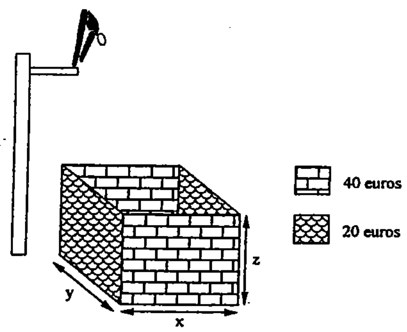
½ (a) Donner la relation que doit vérifier $d^{(1)} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ pour être conjugé à $d^{(0)}$ par rapport à A.

1,5 (b) Donner la relation que doit vérifier $d^{(1)} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$ pour être une direction de descente de f au point $x^{(1)}$. En utilisant la relation trouvée au a) en déduire que d_2 doit être négatif.

½ (c) Construire $d^{(1)}$ conjugué à $d^{(0)}$ par rapport à A tel que $d_2 = -2$.

1,5 (3) Construire $x^{(2)} = x^{(1)} + t^{(1)}d^{(1)}$où $t^{(1)}$ minimise $f\left(x^{(1)} + td^{(1)}\right)$ avec $t \in \Re$. Est-il nécessaire de faire une itération supplémentaire ? Dans la négative, vérifier l'optimalité de $x^{(2)}$.

UN PLONGEON CALCULÉ



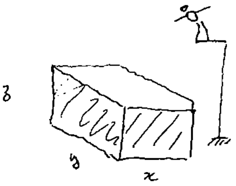
Un plongeur souhaite se fabriquer un bassin dont la forme est un parallèpipède rectangle de longueur x, de largeur y, et de hauteur z. Pour cela il doit fabriquer uniquement les quatre faces verticales (pas de fond car notre ami dispose déjà d'une belle zone carrelée, et pas de haut cela va sans dire...). Les côtés gauche et droit valent 20 euros le mètre carré, et les côtés avant et arrière 40 euros le mètre carré. De nature prudente, il souhaite limiter le risque pris, et après quelques calculs il décide de donner la valeur $\frac{30}{xy} + \frac{90}{z^3}$ à ce dernier.

1,5 (1) Ecrire un programme géométrique (P) décrivant le modèle choisi par le plongeur. On supposera que celui-ci minimise la somme du coût de fabrication et du risque.

2 (2) Donner le programme dual (D) et le résoudre.

1,5 (3) On note v^* la valeur optimale du dual (D). Montrer que v^*est également la valeur optimale de (P) en donnant z puis x et y en fonction de v^* qui réalisent cette valeur (on ne demande pas de donner les valeurs numériques de x, y et z).

un plongeon calculé



40 \$

20 \$

Risque du plongeon estimé à $\frac{30}{xy} + \frac{90}{z^3}$

but : minimiser le coût de fabrication de la piscine + le coût du risque

$$\min \quad \underset{\delta_1}{40yz} + \underset{\delta_2}{80xz} + \underset{\delta_3}{\frac{30}{xy}} + \underset{\delta_4}{\frac{90}{z^3}}$$

$$\begin{cases} \delta_2 - \delta_3 = 0 & \to x \\ \delta_1 - \delta_3 = 0 & \to y \\ \delta_1 + \delta_2 - 3\delta_4 = 0 & \to z \\ \delta_1 + \delta_2 + \delta_3 + \delta_4 = 1 \end{cases}$$

$\delta_2 = \delta_3$

$\delta_1 = \delta_3$

$\delta_4 = \frac{2}{3}\delta_3$

$\delta_3 + \delta_3 + \delta_3 + \frac{2}{3}\delta_3 = \frac{11}{3}\delta_3 = 1 \Rightarrow \delta_3 = \frac{3}{11}$

$\delta_1 = \delta_2 = \delta_3 = \frac{3}{11} \qquad \delta_4 = \frac{2}{11}$

$$v^* = \left(\frac{40 \times 11}{3}\right)^{3/11} \left(\frac{80 \times 11}{3}\right)^{3/11} \left(\frac{30 \times 11}{3}\right)^{3/11} \left(\frac{90 \times 11}{2}\right)^{2/11} = \frac{11 \times 10 \times 4^{3/11} \times (2\times4)^{3/11} \times 3^{3/11} \times (3\times3 \ldots}{3^{3/11} \times 2^{2/11}}$$

$$= \boxed{11 \times 10 \times 4^{6/11} \times 2^{3/11} \times 3^{-2/11}}$$

$\delta_i = \frac{u_i(x)}{g(x)} = \frac{u_i(x)}{v(\delta)}$ $\quad v^* \times \frac{3}{11} = 40yz, \quad v^* \times \frac{3}{11} = 80xz, \quad v^* \times \frac{3}{11} = \frac{30}{xy}, \quad v^* \times \frac{2}{11} = \frac{90}{z^3}$

$y = 2x$

$v^* \frac{3}{11} = \frac{30}{2x^2}$

$x^2 = \frac{30}{2} \times \frac{11}{3v^*} = \frac{5 \times 11}{v^*}$

$x = \sqrt{\frac{55}{v^*}}, \quad y = 2\sqrt{\frac{55}{v^*}}, \quad z^3 = \frac{90 \times 11}{2v^*} = \frac{45 \times 11}{v^*}$

$z = \sqrt[3]{\frac{45 \times 4}{v^*}}$

vérif. $v^* \times \frac{3}{11} = 40 \times \underbrace{2\sqrt{\frac{55}{v^*}}}_{y} \times \underbrace{\sqrt[3]{\frac{45 \times 11}{v^*}}}_{z} \Rightarrow v^* \, v^{*1/2} \, v^{*1/3} = 40 \times 2 \times 55^{1/2} \times (45 \times 11)^{1/3} \times \frac{11}{3}$

$v^{*1+\frac{1}{2}+\frac{1}{3}}$

$v^{*\frac{6+3+2}{6}}$

$\left[v^{*11/6}\right.$

$\Rightarrow v^*$

$= 40 \times 2 \times 5^{1/2} \times 11^{1/2} \times 9^{1/3} \times 5^{1/3} \times 11^{1/3} \times \frac{11}{3}$

$= 40 \times 2 \times 5^{1/2+1/3} \times 11^{1/2+1/3+1} \times 3^{2/3-1}$

$\left. = 40 \times 2 \times 5^{5/6} \times 11^{11/6} \times 3^{-1/3} \right]^{6/11}$

$= 10^{6/11} \times 4^{6/11} \times 2^{6/11} \times 5^{5/11} \times 11 \times 3^{-2/11}$

$= 10^{6/11} \times 4^{6/11} \times \underbrace{2^{5/11} \times 2^{1/11} \times 5^{5/11}} \times 11 \times 3^{-2/11}$

$= 10 \times 4^{6/11} \times 2^{1/11} \times 11 \times 3^{-2/11} = \boxed{10 \times 2^{13/11} \times 11 \times 3^{-2/11}} = \boxed{10 \times 11 \times 2 \times \left[\frac{2}{3}\right]^{\ldots}}$

$\Rightarrow x = 5^{1/2} \times 11^{1/2} \times 10^{-1/2} \times 2^{-13/22} \times 11^{-1/2} \times 3^{-2/11 \times -1/2} = 2^{-1/2} \times 2^{-13/22} \times 3^{1/11} = \boxed{2^{-24/22} \times 3^{1/11} = 2^{-\frac{22}{22}} \times \left[\frac{3}{2}\right]^{1/11}} = \boxed{\frac{1}{2} \times \left[\frac{3}{2}\right]^{1/11}}$

$= 5^{1/2} \times (5 \times 2)^{-1/2} \times 2^{-13/11 \times 2} \times 3^{1/11} = 2^{-1/2} \times 2 \ldots$

$y = 2 \times 2^{-24/22} \times 3^{1/11} = \boxed{2^{-2/22} \times 3^{1/11} = \left[\frac{3}{2}\right]^{1/11}}$

$z = (9 \times 5 \times 11)^{1/3} \times \left[10 \times 2^{13/11} \times 11 \times 3^{-2/11}\right]^{-1/3} = 3^{2/3} \times 5^{1/3} \times \underbrace{5^{-1/3} \times 2^{-1/3} \times 2^{-13/33} \times 3^{2/33}} = 3^{24/33} \times 2^{-24/33} = \boxed{\left[\frac{3}{2}\right]^{24/33}} = \boxed{\left[\frac{3}{2}\right]^{8/11}}$

vérif : $40 \times \underbrace{2^{-1/11} \times 3^{1/11}}_{y} \times \underbrace{3^{24/33} \times 2^{-24/33}}_{z} + 80 \times \underbrace{2^{-24/22} \times 3^{1/11}}_{x} \times \underbrace{3^{24/33} \times 2^{-24/33}}_{z} + \frac{30}{\underbrace{2^{-24/22} \times 3^{1/11}}_{x} \times \underbrace{2^{-2/22} \times 3^{1/11}}_{y}} + \frac{90}{\left(3^{24/33} \times 2^{-24/33}\right)^3}$

$40 \times 2^{-6/66} \times 2^{-48/66} \times 3^{3/33} \times 3^{24/33} + 80 \times 2^{-72/66 - 48/66} \times 3^{3/33 + 24/33} + \frac{30}{2^{-26/22} \times 3^{2/11}} + \frac{90}{3^{24/11}} \times 2^{24/11}$

$10 \times 2^{13/11} \times 3^{-2/11} \left[4 \times \underbrace{2^{-13/11}} \times 2^{-54/66} \times \underbrace{3^{2/11}} \times 3^{27/33} + 8 \times \underbrace{2^{-13/11}} \times 2^{-120/66} \times \underbrace{3^{2/11}} \times 3^{27/33} + 3 \times \underbrace{2^{-13/11}} \times 2^{26/22} \times \underbrace{3^{2/11}} \times 3^{-2/11} + 9 \times \underbrace{2^{-13/11}} \times 2^{24/11} \times \underbrace{3^{2/11}} \times 3^{-24/11} \right]$

$\left[2^{\frac{132}{66}} \times 2^{-\frac{78}{66}} \times 2^{-\frac{54}{66}} \times 3^{\frac{6}{33}} \times 3^{\frac{27}{33}} + 2^{\frac{198}{66}} \times 2^{-\frac{78}{66}} \times 2^{-\frac{120}{66}} \times 3^{\frac{6}{33}} \times 3^{\frac{27}{33}} + 3 \times 2^{-\frac{26}{22}} \times 2^{\frac{26}{22}} + 9 \times 2^{\frac{11}{11}} \times 3^{-\frac{22}{11}} \right]$

$\left[2^0 \times 3^{\frac{33}{33}} + 2^{\frac{0}{66}} \times 3^{\frac{33}{33}} + 3 + 9 \times 2 \times 3^{-2} \right]$

$\left[3 + 3 + 3 + 2 \right]$

$\left[3 + 3 + 3 \right.$

$10 \times 2^{13/11} \times 3^{-2/11} \left[11 \right] = v^*$

$$\min \frac{x^3}{yz} + y^2z^2 + \frac{1}{xy^5z^5} \qquad \text{s.c. } x,y,z > 0$$

on peut poser $yz = u$

$$\min \frac{x^3}{u} + u^2 + \frac{1}{xu^5} \qquad \begin{array}{l} x>0 \\ u>0 \end{array}$$

$$\begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \delta_1 & \delta_2 & \delta_3 \end{array}$$

Dual
$$\begin{cases} 3\delta_1 - \delta_3 = 0 & \leftarrow x \\ -\delta_1 + 2\delta_2 - 5\delta_3 = 0 & \leftarrow u \quad \rightarrow -\delta_1 + 2\delta_2 - 15\delta_1 = 0 \rightarrow 2\delta_2 = 16\delta_1 \rightarrow \delta_2 = 8\delta_1 \\ \delta_1 + \delta_2 + \delta_3 = 1 & \rightarrow \delta_1 + 8\delta_1 + 3\delta_1 = 1 \end{cases}$$

$$\boxed{\begin{array}{l} \delta_1 = \frac{1}{12} \\ \delta_2 = \frac{8}{12} = \\ \delta_3 = \frac{3}{12} = \end{array}}$$

$$v(\delta) = 12^{1/12} \times \left(\frac{12}{8}\right)^{8/12} \times \left(\frac{12}{3}\right)^{3/12} = 12 \times \frac{1}{(2^3 \times 3)^{8/12}} \times \frac{1}{3^{3/12}} = 2^{2-2} \times 3^{1-3/12} = 3^{9/12} = 3^{3/4}$$

retour au primal avec u

$$\frac{x^3}{u} = \frac{1}{12} \times 3^{3/4}, \quad u^2 = \frac{8}{12} \times 3^{3/4}, \quad \frac{1}{xu^5} = \frac{3}{12} \times 3^{3/4}$$

$$\Rightarrow x^3 = \frac{1}{12} \times 3^{3/4} \times \underbrace{\frac{2}{6^{1/2}} \times 3^{3/8}}_{\boxed{u = 2^{1/2} \times \frac{1}{3^{1/8}}}} = \frac{1}{2^2 \times 3} \times 3^{9/8} \times \frac{2}{3^{1/2} \times 2^{1/2}} = \frac{1}{2^{3/2}} \times 3^{9/8 - \frac{12}{8}} = \frac{1}{2^{3/2}} \times \frac{1}{3^{3/8}} \Rightarrow \boxed{x = \frac{1}{2^{1/2} 3^{1/8}}}$$

$$\Rightarrow \frac{1}{xu^5} = \frac{2^{1/2} \times 3^{1/8}}{3^{5/8}} \times 2^{-5/2} = \frac{1}{4} \times 3^{3/4} \qquad \text{donc c'est bon.}$$

prenons $\boxed{y = 2^{1/2} \text{ et } z = \frac{1}{3^{1/8}}}$

vérifions l'égalité du primal et du dual :

$$\underbrace{\frac{1}{2^{3/2} 3^{3/8}}}_{x^3} \times \underbrace{\frac{3^{+1/8}}{2^{1/2}}}_{1/yz} + \frac{2}{3^{1/4}} + \underbrace{\frac{1}{2^{-1/2} 3^{1/8}}}_{1/x} \times \frac{3^{5/8}}{2^{5/2}} = \frac{1}{4 \times 3^{1/4}} + \frac{2}{3^{1/4}} + \frac{1}{4} \times 3^{3/4} = \frac{1+8+3}{4 \times 3^{1/4}}$$

$$= \frac{12}{4 \times 3^{1/4}} = 3^{3/4} \quad \text{c'est exact}$$

Minimisation d'une fonction quadratique

Soit $q(x) = \frac{1}{2} x \cdot Qx - b \cdot x$ une fonction quadratique définie sur $\mathbb{R}^n$ avec Q symétrique.

1) m.q. $Q \geq 0$ et $b \in \operatorname{Im} Q$ sont des conditions nécessaires et suffisantes pour que $q(x)$ admette un minimum

2) On note b_N la composante de b sur $(\operatorname{Im} Q)^\perp$.
m.q. si $b \notin \operatorname{Im} Q$ (i.e. $b_N \neq 0$) alors q décroît indéfiniment dans la direction ~~opposée à~~ b_N.

3) soit $q(x) = \frac{1}{2}x_1^2 + x_2^2 + \frac{x_3^2}{2} + x_1x_2 + x_1x_3 + x_2x_3 - 3x_1 - 3x_2 - 2x_3$
$q(x)$ admet-elle un minimum ?
m.q. $t\begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}$ est une direction dans laquelle q décroît indéfiniment.

1) $\nabla q = Qx - b = 0 \Leftrightarrow Qx = b$ CN 1^er ordre. solution ssi $b \in \operatorname{Im}(Q)$
$H_q = Q \geq 0$ CN 2^e ordre.
si les 2 conditions sont remplies, il existe un pt critique et q est convexe donc ce pt critique est minimum.

2) $b = b_I + b_N$ avec $b_I \in \operatorname{Im} Q$ et $b_N \in (\operatorname{Im} Q)^\perp = \mathcal{N}(Q^T) = \mathcal{N}(Q)$

$$q(t b_N) = \frac{t^2}{2}\big(b_N \cdot \underbrace{Q b_N}_{0}\big) - t\underbrace{(b_I + b_N)}_{b} \cdot b_N = -t \underbrace{b_I \cdot b_N}_{0} + b_N \cdot b_N = -t\underbrace{\|b_N\|^2}_{\neq 0}$$

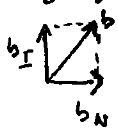
quand $t \to +\infty$ $q(t b_N) \to -\infty$

3) on a ici $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ et $b = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix}$

$Q \geq 0$ mais il n'existe pas d_1, d_2, d_3 t.q. $d_1\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + d_2\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + d_3\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix}$
car on doit ~~aurait~~ $d_1 + d_2 + d_3 = 3$ et $d_1 + d_2 + d_3 = 2$

on vérifie que $\begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix} \in \mathcal{N}(Q) = (\operatorname{Im} Q)^\perp$ $\quad Q\begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

et $\begin{bmatrix} 3 \\ 3 \\ 2 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 5/2 \\ 3 \\ 5/2 \end{bmatrix} \in \operatorname{Im} Q$ $\quad \underset{\substack{\parallel \\ 2}}{d_1}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \underset{\substack{\parallel \\ 1/2}}{d_2}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5/2 \\ 3 \\ 5/2 \end{bmatrix}$ ou encore $\begin{bmatrix} 5/2 \\ 3 \\ 5/2 \end{bmatrix} \cdot \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix} = 0$

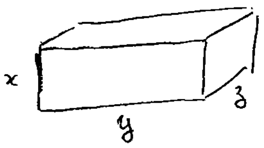


$q\begin{pmatrix} +1/2 \\ 0 \\ -1/2 \end{pmatrix} = -\frac{1}{2}$

~~Positivité~~

Programmation géométrique

~~à faire comme application ... l'exo~~



on ~~peut~~ doit construire une boîte sans couvercle. parallélépipédique.

on veut minimiser la surface ~~tout~~ en maximisant le volume V

maximiser ~~le volume~~ V ~~est~~ équivalent à minimiser $\frac{1}{V}$.

Donc finalement on décide de résoudre le pb suivant :

$x, y, z > 0$

minimiser $g(x,y,z) = \underset{\delta_1}{yz} + \underset{\delta_2}{2xy} + \underset{\delta_3}{2xz} + \underset{\delta_4}{\frac{1}{xyz}}$

$$\begin{cases} \delta_2 + \delta_3 - \delta_4 = 0 & (a) \\ \delta_1 + \delta_2 - \delta_4 = 0 & (b) \\ \delta_1 + \delta_3 - \delta_4 = 0 & (c) \\ \delta_1 + \delta_2 + \delta_3 + \delta_4 = 1 \end{cases}$$

$a - b \rightarrow \delta_3 = \delta_1 \longrightarrow 2\delta_1 = \delta_4$

$a - c \rightarrow \delta_1 = \delta_2$

$\longrightarrow 3\delta_1 + 2\delta_1 = 1 \rightarrow \delta_1 = \frac{1}{5}$

$v = (5)^{1/5} \times (10)^{1/5} \times (10)^{1/5} \times \left(\frac{5}{2}\right)^{2/5} = 5$

$\delta_i = \frac{u_i}{v}$ $\quad yz = \frac{1}{5} \times 5$, $2xy = \frac{1}{5} \times 5$, $2xz = \frac{1}{5} \times 5$, $\frac{1}{\underbrace{xyz}_{1}} = \frac{2}{5} \times 5$

$y = 1$ $\quad z = 1$ $\quad \frac{1}{x} = 2 \rightarrow x = \frac{1}{2}$

$g(x,y,z) = 1 + 1 + 1 + 2 = 5$

$$\nabla g = \begin{cases} 2y + 2z - \frac{yz}{(xyz)^2} = 0 & 2 + 2 - \frac{1}{1/4} = 0 \\ z + 2x - \frac{xz}{(xyz)^2} = 0 & 1 + 1 - \frac{1/2}{1/4} = 0 \\ y + 2x - \frac{xy}{(xyz)^2} = 0 & 1 + 1 - \frac{1/2}{1/4} = 0 \end{cases}$$

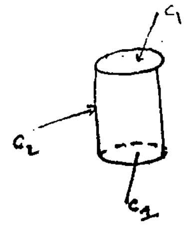
L'Inégalité arithmétique et géométrique

$$\sum_{i=1}^{n} \delta_i x_i \geqslant \prod_{i=1}^{n} (x_i)^{\delta_i}$$

$x_i > 0 \in \mathbb{R}$
$\delta_i > 0 \in \mathbb{R}$
avec $\sum_{i=1}^{n} \delta_i = 1$

on a égalité ssi $x_1 = x_2 = \dots = x_n$

Maximiser le volume d'un cylindre sachant que le cylindre doit coûter c_0 et que le coût du haut et du bas est de c_1 par unité de surface et le coût du coté est de c_2 par unité de surface.



r = rayon
h = hauteur.

volume $V = \pi r^2 h$

coût $c_0 = 2\pi r^2 c_1 + 2\pi r h c_2$

$$c_0 = 2\pi r^2 c_1 + 2\pi r h c_2 = 2\pi r^2 c_1 + \pi r h c_2 + \pi r h c_2 = 3\left(\frac{2\pi r^2 c_1 + \pi r h c_2 + \pi r h c_2}{3}\right)$$

$$\geqslant 3 \left(2\pi \pi^2 r^4 h^2 c_1 c_2^2\right)^{1/3} = 3\left(2\pi V^2 c_1 c_2^2\right)^{1/3}$$

$$\left(a^{1/3} \cdot b^{1/3} = (a \cdot b)^{1/3}\right)$$

$$= 3(2\pi)^{1/3}(c_1 c_2^2)^{1/3} V^{2/3}$$

V est max qd on a égalité c.a.d qd

$$2\pi r^2 c_1 = \pi r h c_2 = \pi r h c_2 = \frac{c_0}{3}$$

$$\Rightarrow h = 2r\frac{c_1}{c_2} \Rightarrow \boxed{h = 2\frac{c_1}{c_2} \times \left(\frac{c_0}{c_1} \times \frac{1}{6\pi}\right)^{1/2}}$$

$$\Rightarrow r^2 = \frac{c_0}{3} \times \frac{1}{c_1 2\pi} \Rightarrow \boxed{r = \left(\frac{c_0}{c_1} \times \frac{1}{6\pi}\right)^{1/2}}$$

Inégalité arithmétique et géométrique (suite)

$\max f(x_1,x_2,x_3) = x_1 x_2^2 x_3$

s.c $x_1+x_2+x_3^2 = k$ $\quad$ $k>0$ fixé, $x_1>0, x_2>0, x_3>0$

$$k = x_1+x_2+x_3^2 = \frac{x_1}{2}+\frac{x_1}{2}+\frac{x_2}{4}+\frac{x_2}{4}+\frac{x_2}{4}+\frac{x_2}{4}+x_3^2 = 7\left(\frac{\frac{x_1}{2}+\frac{x_1}{2}+\cdots\cdots}{7}\right)$$

$$\geq 7\left(\frac{x_1^2}{2^2}\times\frac{x_2^4}{4^4}\times x_3^2\right)^{1/7} = 7\left(\frac{1}{2}\right)^{2/7}\left(\frac{1}{4}\right)^{4/7}\left(x_1 x_2^2 x_3\right)^{2/7}$$

max qd on a égalité c.a.d qd $\frac{x_1}{2} = \frac{x_2}{4} = x_3^2 = \frac{k}{7}$

$\boxed{x_3 = \sqrt{k/7}}$ $\quad$ $\boxed{x_2 = \frac{4}{7}k}$ $\quad$ $\boxed{x_1 = \frac{2}{7}k}$

Pb Dual : $\quad$ $\min\ x_1+x_2+x_3^2$ $\quad$ (contrainte de l'ancien)

s.c $x_1 x_2^2 x_3 = c$ $\quad$ (objectif de l'ancien)

min qd on a égalité cad qd $\frac{x_1}{2} = \frac{x_2}{4} = x_3^2$

$$x_1+x_2+x_3^2 = 7\left(\frac{1}{2}\right)^{2/7}\left(\frac{1}{4}\right)^{4/7} c^{2/7}$$

$\Rightarrow x_1 + 2x_1 + \frac{x_1}{2} = \underline{\qquad\qquad}$

$\Rightarrow x_1 = \frac{2}{\cancel{7}}\times\cancel{7}\left(\frac{1}{2}\right)^{2/7}\left(\frac{1}{4}\right)^{4/7}c^{2/7}$ $\quad$ $\Rightarrow x_2 = 2x_1 = 4\times\left(\frac{1}{2}\right)^{2/7}\left(\frac{1}{4}\right)^{4/7}c^{2/7}$

$\Rightarrow x_3 = \sqrt{\frac{x_1}{2}} =$

<u>trouver ~~les valeurs~~ minimum de :</u>

$$f(x_1,x_2) = 4x_1 + \frac{x_1}{x_2^2} + \frac{4x_2}{x_1} \qquad \text{s.c. } x_1>0,\ x_2>0$$

$$f(x_1,x_2) = 4x_1 + \frac{x_1}{x_2^2} + \frac{2x_2}{x_1} + \frac{2x_2}{x_1} = 4\left(\frac{4x_1 + \frac{x_1}{x_2^2} + \frac{2x_2}{x_1} + \frac{2x_2}{x_1}}{4}\right)$$

$$\geqslant 4\left(16\,\frac{x_1^2x_2^2}{x_1^2x_2^2}\right)^{1/4} = 8$$

minimum atteint qd on a égalité c.a.d qd $4x_1 = \frac{x_1}{x_2^2} = \frac{2x_2}{x_1} = \frac{8}{4} = 2$

$$\Rightarrow x_1 = \frac{1}{2} \qquad x_2 = \frac{1}{2}$$

Remarque : on peut prouver qu'on peut toujours obtenir l'égalité. (voir cours)

<u>trouver minimum de :</u>

$$f(x) = c_1x^3 + \frac{c_2}{x} \qquad \text{pour } x>0$$

$$f(x) = c_1x^3 + \frac{c_2}{3x} + \frac{c_2}{3x} + \frac{c_2}{3x} = 4\left(\frac{c_1x^3 + \frac{c_2}{3x} + \frac{c_2}{3x} + \frac{c_2}{3x}}{4}\right) \geqslant 4\left(\frac{c_1c_2^3}{3^3}\,\frac{x^3}{x^3}\right)^{1/4}$$

égalité obtenue qd $c_1x^3 \underset{(1)}{=} \frac{c_2}{3x} \underset{(2)}{=} \left(\frac{c_1c_2^3}{3^3}\right)^{1/4}$

$$(1) \Rightarrow x^4 = \frac{c_2}{3c_1} \Rightarrow x = \left(\frac{c_2}{3c_1}\right)^{1/4}$$

c'est bien compatible avec (2)

$$(2) \Rightarrow 3^{4/4}x = c_2^{4/4}\left(\frac{3^3}{c_1c_2^3}\right)^{1/4}$$

$$\Rightarrow x = \left(\frac{c_2}{3c_1}\right)^{1/4}$$

<u>par méthode classique :</u>

$$f'(x) = 3c_1x^2 - \frac{c_2}{x^2} = 0 \Rightarrow 3c_1x^2 = \frac{c_2}{x^2} \Rightarrow x^4 = \frac{c_2}{3c_1}$$

$$f''(x) = 6c_1x + \frac{2c_2}{x^3} > 0 \quad \forall x>0 \Rightarrow \text{défini positif} \Rightarrow \text{point critique est un min global.}$$

valeur du minimum : $c_1^{4/4}\left(\frac{c_2}{3c_1}\right)^{3/4} + \left(\frac{c_2^{4}\,3c_1}{c_2}\right)^{1/4} = \left(\frac{c_1c_2}{3}\right)^{3/4} + \left(3c_2^3c_1\right)^{1/4}$

16. (a) Show that if c_1, c_2, c_3, c_4 are positive numbers, then

[illegible]

has no minimum on the set

$$D = \{(x_1, x_2, x_3, x_4): x_i > 0 \text{ for } i = 1, 2, 3, 4\}.$$

(b) Set up the log-linear equations whose solution produces the minimizer of

$$g(t_1, t_2, t_3, t_4) = c_1 t_3^{-1} t_4^{-2} + c_2 t_2^{-6} t_4^2 + c_3 t_1^3 t_4^2 + c_4 t_1^{-1} t_2^4 t_3^2 t_4^2$$

over all $t_1, t_2, t_3, t_4 > 0$. Give the relative contributions of each term to the minimum.

i) $c_1 x_3^{-1} + c_2 x_2^{-6} x_4^2 + c_3 x_1^3 x_4^2 + c_4 x_1^{-1} x_2^4 x_3^2 x_4^2$

$\downarrow \delta_1 \quad \downarrow \delta_2 \quad \downarrow \delta_3 \quad \downarrow \delta_4$

Dual :

$+3\delta_3 - \delta_4 = 0$

$-6\delta_2 + 4\delta_4 = 0$

$-\delta_1 + 2\delta_4 = 0$

$2\delta_2 + 2\delta_3 + 2\delta_4 = 0 \Rightarrow$ impossible avec $\delta_2 > 0, \delta_3 >, \delta_4 > 0$

ii) $3\delta_3 - \delta_4 = 0 \quad \to \delta_4 = 3\delta_3$

$-6\delta_2 + 4\delta_4 = 0 \quad \to \delta_4 = \frac{3}{2}\delta_2$

$-\delta_1 + 2\delta_4 = 0 \quad \to \delta_4 = \frac{\delta_1}{2}$

$-2\delta_1 + 2\delta_2 + 2\delta_3 + 2\delta_4 = 0 \longrightarrow -2\delta_4 + \frac{2}{3}\delta_4 + \frac{\delta_4}{3} + \delta_4 = 0 \Rightarrow$ rien.

$\delta_1 + \delta_2 + \delta_3 + \delta_4 = 1 \to 2\delta_4 + \frac{2}{3}\delta_4 + \frac{\delta_4}{3} + \delta_4 = 1 \to \delta_4 = \frac{1}{4} \to \delta_3 = \frac{1}{12}, \ \delta_2 = \frac{1}{6}, \ \delta_1 = \frac{1}{2}$

$v(\delta^*) = (2c_1)^{1/2} (6c_2)^{1/6} (12c_3)^{1/12} (4c_4)^{1/4} = g(x^*)$

$\Rightarrow \frac{1}{2} \times v(\delta^*) = c_1 x_3^{-1}, \quad \frac{1}{6} \times v(\delta^*) = c_2 x_2^{-6} x_4^2, \quad \frac{1}{12} v(\delta^*) = c_3 x_1^3 x_4^2, \quad \frac{1}{4} v(\delta^*) = c_4 x_1^{-1} x_2^4 x_3^2 x_4^2$

$\Rightarrow y_3 = \log 2 - \log v(\delta^*) + \log c_1, \quad -6y_2 + 2y_4 = -\log 6 + \log v(\delta^*) - \log c_2$, etc....

on a un système linéaire pour les variables $y_i = \log x_i$. $(x_i > 0)$

contribution de chaque terme au minimum est respectivement $\frac{1}{2}, \frac{1}{6}, \frac{1}{12}, \frac{1}{4}$ ceci $\forall c_1, c_2, c_3, c_4$

Montrer que les fonctions suivantes sont convexes ou strictement convexes sur l'ensemble convexe D spécifié : 1/.

* $f(x_1,x_2) = 5x_1^2 + 2x_1x_2 + x_2^2 - x_1 + 2x_2 + 3$ sur $D = \mathbb{R}^2$ est-elle coercive ?

$$\nabla f = \begin{pmatrix} 10x_1 + 2x_2 - 1 \\ 2x_2 + 2x_1 + 2 \end{pmatrix}$$

$$H_f = \begin{pmatrix} 10 & 2 \\ 2 & 2 \end{pmatrix} \qquad \Delta_1 > 0 \qquad \Delta_2 = 20 - 4 > 0$$

H_f est défini positif $\Rightarrow$ f est strictement convexe.

$x_1^* = \frac{3}{8}$ $x_2^* = -\frac{11}{8}$ est le minimum global.

* $f(x_1,x_2) = \frac{x_1^2}{2} + \frac{3x_2^2}{2} + \sqrt{3}\, x_1 x_2$ sur $D = \mathbb{R}^2$ est-elle coercive ?

$$\nabla f = \begin{pmatrix} x_1 + \sqrt{3}\, x_2 \\ 3x_2 + \sqrt{3}\, x_1 \end{pmatrix}$$

$$H_f = \begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix} \qquad \begin{array}{l} \Delta_1 > 0 \\ \Delta_2 = 3 - 3 = 0 \end{array}$$

H_f est semi-défini positif $\Rightarrow$ f est convexe

On peut le voir :

$$\nabla f = 0 \Rightarrow \begin{cases} x_1 + \sqrt{3}\, x_2 = 0 \Rightarrow \sqrt{3}\, x_1 + 3x_2 = 0 \\ 3x_2 + \sqrt{3}\, x_1 = 0 \Rightarrow 3x_2 + \sqrt{3}\, x_1 = 0 \end{cases} \Big\}\ \text{même équation.}$$

ensemble des minimums $\longrightarrow$ $x_2 = -\frac{x_1}{\sqrt{3}}$ ensemble des points critiques.

$$f\left(x_1, -\frac{x_1}{\sqrt{3}}\right) = \frac{x_1^2}{2} + \frac{3}{2}\frac{x_1^2}{3} - \sqrt{3}\,\frac{x_1^2}{\sqrt{3}} = \frac{x_1^2}{2} + \frac{x_1^2}{2} - x_1^2 = 0$$

tous les points critiques donnent la même valeur.

L'ensemble des points critiques est une droite donc un ensemble convexe $\Rightarrow$
toute combinaison linéaire convexe de points critiques est un point critique

$X_1 \neq X_2 \in$ points critiques. $0 < \alpha < 1$

$$\underbrace{f(\alpha X_1 + (1-\alpha) X_2)}_{0} \geq \alpha \underbrace{f(X_1)}_{0} + (1-\alpha) \underbrace{f(X_2)}_{0} \quad \text{car } f \text{ convexe.}$$

donc en fait on a égalité : $\forall X_1 \neq X_2$ points critiques et $0 < \alpha < 1$

$$f(\alpha X_1 + (1-\alpha) X_2) = \alpha f(X_1) + (1-\alpha) f(X_2)$$

donc pas strictement convexe. (on n'a pas inégalité stricte)

plus simplement, on remarque que : $f(x_1,x_2) = \left(\frac{x_1}{\sqrt{2}} + \frac{\sqrt{3}x_2}{\sqrt{2}}\right)^2 = \frac{x_1^2}{2} + \frac{3x_2^2}{2} + \sqrt{3}\frac{x_1x_2}{\sqrt{2}\cdot\sqrt{2}}\cdot 2$

donc min $= 0$ pour $\frac{x_1}{\sqrt{2}} + \frac{\sqrt{3}}{\sqrt{2}} x_2 = 0$

× $f(x_1,x_2) = 4\,e^{3x_1 - x_2} + 5\,e^{x_1^2 + x_2^2}$ sur $D = \mathbb{R}^2$

$= h(x_1,x_2) + g(x_1,x_2)$

$h(x_1,x_2)$ convexe ou strictement convexe ?

$z = 3x_1 - x_2$ est convexe

$e^z = u(z)$ est convexe car $u''(z) = e^z > 0$ (Hessien défini positive)

est croissante —— $u'(z) = e^z > 0$

$\Rightarrow e^{3x_1 - x_2}$ est convexe $\Rightarrow 4\,e^{3x_1 - x_2}$ est convexe.

$g(x_1,x_2)$ convexe ou strictement convexe ?

$z = x_1^2 + x_2^2$ est strictement convexe $H_z = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ défini positif.

$e^z = u(z)$ est convexe | voir plus haut.

est strictement croissante |

$\Rightarrow e^{x_1^2 + x_2^2}$ est strictement convexe $\Rightarrow 5\,e^{x_1^2 + x_2^2}$ est strictement convexe.

Un des membres de la somme $h(x_1,x_2) + g(x_1,x_2)$ est strictement convexe donc la somme est strictement convexe.

× $f(x_1,x_2) = c_1 x_1 + c_2/x_1 + c_3 x_2 + c_4/x_2$ sur $D = \{(x_1,x_2) \in \mathbb{R}^2 \,/\, x_1 > 0 \text{ et } x_2 > 0\}$ où $c_i > 0$ $i = 1,4$.

$$\nabla f = \begin{pmatrix} c_1 - \dfrac{c_2}{x_1^2} \\ c_3 - \dfrac{c_4}{x_2^2} \end{pmatrix}$$

$$Hf = \begin{pmatrix} +2\dfrac{c_2}{x_1^3} & 0 \\ 0 & +2\dfrac{c_4}{x_2^3} \end{pmatrix}$$

matrice diagonale. valeurs propres > 0 sur D.

$\Rightarrow Hf$ est défini positif. $\Rightarrow f$ est strictement convexe.

autre façon de voir :

x est convexe

$1/x \xrightarrow{'} -\dfrac{1}{x^2} \xrightarrow{''} \dfrac{2}{x^3} > 0$ pour $x > 0$ donc $\frac{1}{x}$ strictement convexe.

$f(x_i,x_2)$ est la somme de fonctions convexes $\rightarrow$ elle est convexe

de plus au moins un membre de la somme est ~~strictement~~ <u>strictement convexe</u> $\Rightarrow$ elle est strictement convexe

non car 2 variables.

$\times$ $f(x_1, x_2) = (x_1 + 2x_2 + 1)^8 - \log(x_1 x_2)^2$ sur $D = \{(x_1,x_2) \in \mathbb{R}^2 / x_1 > 0 \text{ et } x_2 > 0\}$

$= h(x_1, x_2) - 2\log(x_1) - 2\log(x_2)$

$= h(x_1, x_2) + g_1(x_1, x_2) + g_2(x_2, x_1)$

$h(x_1, x_2)$ convexe ou strictement convexe ? pour $(x_1, x_2) \in D$

$z = x_1 + 2x_2 + 1$ est convexe

$u(z) = z^8$ $\quad u' = 8z^7 \quad u'' = 56z^6 > 0 \quad \forall z > 0$

$u' > 0$ pour $z > 0$

pour $(x_1, x_2) \in D$ on a $z > 0$

donc u est croissante strictement et convexe.

$\Rightarrow h(x_1, x_2)$ est convexe (seul^t car z est convexe seul^t (pas strict^t))

attention : il faut considérer $\log x_1$ comme fction de 2 variables. x_1 et x_2

$g(x)$ convexe ou strictement convexe ? pour $x > 0$.

$-\log x \xrightarrow{\ '\ } -\frac{1}{x} \xrightarrow{\ ''\ } \frac{1}{x^2} > 0 \Rightarrow -\log x$ est strictement convexe.

$\Rightarrow -2\log x$ est strictement convexe.

~~donc $f(x_1, x_2)$ est la somme de fctions convexe $\Rightarrow$ elle est convexe~~

~~de plus un des membres de la somme est strictement convexe $\Rightarrow$ elle est strictement convexe.~~

$g_1(x_1, x_2) = (2)(-\log(x_1 + 0.x_2))$

$(x_1, x_2) \longmapsto x_1 + 0.x_2$ est convexe.

$-\log z$ est (strict) convexe et ~~croissante~~ non croissante donc on ne peut pas conclure !

~~donc $-\log(x_1 + 0x_2)$ est convexe.~~

~~donc $(2)(-\log(x_1 + 0x_2))$ est convexe.~~

$g(x_1, x_2) = -\log(x_1 + 0.x_2)$ $\quad \nabla g = \begin{pmatrix} -1/x_1 \\ 0 \end{pmatrix} \quad Hg = \begin{pmatrix} 1/x_1^2 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$

* matrice diagonale :

$d_1 > 0 \quad d_2 = 0 \quad \Rightarrow Hg$ semi définie positive. $\Rightarrow g$ convexe.

$\Rightarrow g_1(x_1, x_2) = 2 \cdot g(x_1, x_2)$ convexe.

(matrice diagonale : semi définie positive $\Leftrightarrow d_i \geq 0 \quad i = 1, \dots, n$) <u>on peut faire mieux</u> $\Rightarrow$

$$g = -\mathcal{L}og\, x_1 x_2 = -\mathcal{L}og\, x_1 - \mathcal{L}og\, x_2$$

$$\nabla g = \begin{pmatrix} -\frac{1}{x_1} \\ -\frac{1}{x_2} \end{pmatrix} \qquad H_g = \begin{pmatrix} 1/x_1^2 & 0 \\ 0 & 1/x_2^2 \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$$

$d_i > 0 \quad i = 1,2 \iff H_g$ définie positive. $\Rightarrow$ g strict^t convexe

remarque : on a équivalence pour convexe et semi-déf. positif mais pas pour strict. convexe et déf. positif. ex : $f(x) = x^4$ $f''(x) = 12x^2$ n'est pas déf. positif pour $x = 0$

et donc f est en fait strict^t convexe.

Convexité

Soit f une fonction convexe de $\mathbb{R}$ dans $\mathbb{R}$. Soit $a \in \mathbb{R}^m$, $\alpha \in \mathbb{R}$ fixés.
On définit la fonction g de $\mathbb{R}^m$ dans $\mathbb{R}$ par $g(x) = f(a\cdot x + \alpha)$ pour tout $x \in \mathbb{R}^m$
Montrer que g est convexe.
Application : en déduire que $h(x,y,z) = (4x + 5y - 8z + 7)^8$ est convexe.

Remarque : par rapport au th. sur la composition des fonctions convexes, ici on peut se passer de l'hypothèse de croissance de f.

$g((1-\lambda)x + \lambda y) =$

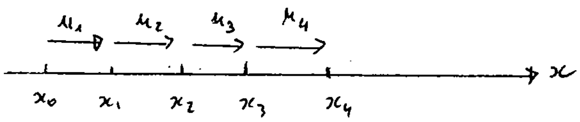
$f(a\cdot[(1-\lambda)x + \lambda y] + \alpha) = f((1-\lambda)a\cdot x + (1-\lambda)\alpha + \lambda a\cdot y + \lambda\alpha)$

$f[(1-\lambda)(a\cdot x+\alpha) + \lambda(a\cdot y + \alpha)] \leq (1-\lambda) f(a\cdot x + \alpha) + \lambda f(a\cdot y + \alpha) = (1-\lambda) g(x) + \lambda g(y)$

application : $a = \begin{bmatrix} 4 \\ 5 \\ -8 \end{bmatrix}$ $\alpha = 7$ $f(z) = z^8$ (élévation à la puissance 8)

Un chariot sur un axe est à la position initiale x_0. Pour rapprocher le chariot de la position $x = 0$ on applique successivement m déplacements, chacun eux ayant un coût égal au carré du déplacement.
On note u_k le k^e déplacement et x_k la k^e position du chariot.
Pour rapprocher le chariot de la position $x=0$ en tenant compte du coût induit par les déplacements on minimise $\sum_{k=1}^{m} x_k^2 + \sum_{k=1}^{m} u_k^2$

Montrer que l'objectif à minimiser est strict^t convexe en fonction des u_k
Résoudre le pb pour $m = 3$.



$x_k = x_0 + \sum_{i=1}^{k} u_i$ est de la forme $\alpha + a\cdot u$ $\alpha = x_0$ et $a = \begin{array}{c} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \}\, k \\ \}\, m-k \end{array}$

$f(z) = z^2$ est convexe.

d'après l'exercice précédent x_k^2 est <u>convexe</u>.

$\sum_{k=1}^{m} u_k^2$ a pour Hessien $2 \times I_m \Rightarrow$ défini positif $\Rightarrow$ <u>strict^t convexe</u>.

L'objectif à minimiser est la <u>somme de $m+1$ fonctions convexes dont la dernière l'est strict^t</u>
par théorème l'objectif est strict^t convexe.

Pour $n=3$

$$f(u_1,u_2,u_3) = x_1^2+x_2^2+x_3^2+u_1^2+u_2^2+u_3^2 = (x_0+u_1)^2+(x_0+u_1+u_2)^2+(x_0+u_1+u_2+u_3)^2+u_1^2+u_2^2+u_3^2$$

on cherche point critique (f convexe : x^* point critique $\Leftrightarrow$ x^* minimise f sur $\mathbb{R}^n$)

$$\frac{\partial f}{\partial u_1} = \cancel{2}(x_0+u_1) + \cancel{2}(x_0+u_1+u_2) + \cancel{2}(x_0+u_1+u_2+u_3) + \cancel{2}u_1 = 0$$

$$\frac{\partial f}{\partial u_2} = \cancel{2}(x_0+u_1+u_2) + \cancel{2}(x_0+u_1+u_2+u_3) + \cancel{2}u_2 = 0$$

$$\frac{\partial f}{\partial u_3} = \cancel{2}(x_0+u_1+u_2+u_3) + \cancel{2}u_3 = 0$$

$$\begin{cases} 4u_1+2u_2+u_3+3x_0 = 0 \\ 2u_1+3u_2+u_3+2x_0 = 0 \\ u_1+u_2+2u_3+x_0 = 0 \end{cases} \Rightarrow \begin{cases} u_1^* = -\frac{8}{13}x_0 \\ u_2^* = -\frac{3}{13}x_0 \\ u_3^* = -\frac{x_0}{13} \end{cases}$$

u^* est minimum global $\Rightarrow$ (par convexité de f)

2 minimums.

le minimum est unique, c'est normal f est strict. convexe. $\quad u^* \neq v^* \quad 0<\lambda<1$

$$f(u^*) \leqslant f((1-\lambda)u^*+\lambda v^*) \underset{\substack{\uparrow \\ \text{strict. convexe.}}}{<} (1-\lambda)f(u^*)+\lambda f(v^*) = (1-\lambda)f(u^*)+\lambda f(u^*) = f(u^*)$$

3.4 Infimum of a quadratic function

Technically, the essential content of this paper consists in applying Lagrangian duality to optimization problems involving quadratic functions. As it happens, linear matrix inequalities and their accompanying SDP duality appear naturally in this context, and the aim of this section is to demonstrate the connection. Given $Q \in \mathcal{S}_p$, $b \in \mathbb{R}^p$ and $c \in \mathbb{R}$, consider the function

$$\mathbb{R}^p \ni y \mapsto q(y) := y^\top Q y + b^\top y + c\,.$$

The following result is elementary.

Lemma 3.6 *A necessary and sufficient condition for q to have a finite lower bound over $\mathbb{R}^n$ $(\inf q > -\infty)$ is*

condition nécessaire du 2e ordre — du 1er ordre

$$Q \succcurlyeq 0 \quad \textit{and} \quad b \in \mathcal{R}(Q)\,. \tag{19}$$

Proof. [Necessity] If $Q \not\succcurlyeq 0$, *i.e.*, q is not convex, there exists $y \in \mathbb{R}^p$ such that $y^\top Q y < 0$. Then $q(ty) \to -\infty$ when $t \to +\infty$.

On the other hand, $b \notin \mathcal{R}(Q)$ means that b has a nonzero component $b_\mathcal{N}$ on $(\mathcal{R}(Q))^\perp = (\mathcal{R}(Q^\top))^\perp = \mathcal{N}(Q)$. Then

$$q(-tb_\mathcal{N}) = b^\top(-tb_\mathcal{N}) + c = -t\|b_\mathcal{N}\|^2 + c\,,$$

which tends to $-\infty$ when $t \to +\infty$.

[Sufficiency] Conversely, the property $b \in \mathcal{R}(Q)$ means that the linear system $2Qy = -b$ (*i.e.*, $\nabla q(y) = 0$) has at least one solution, and this *characterizes* the minimum points of the convex function q. ☐

When (19) holds, the minimization of q can be carried out explicitly in terms of the spectral decomposition $Q = \sum_{i=1}^{p-k} \lambda_i q_i q_i^\top$. Here $0 \leqslant k \leqslant p$ is the rank of Q, $\lambda_1 > \ldots > \lambda_{p-k} > 0$ and $q_1, \ldots, q_{p-k}$ are orthonormal (eigen)vectors forming a basis of $\mathcal{R}(Q)$. The *pseudo-inverse* of Q is then

$$Q^\dagger := \sum_{i=1}^{p-k} \lambda_i^{-1} q_i q_i^\top\,.$$

Lemma 3.7 *When* (19) *holds and using the above notation,* $\bar{y} := -\frac{1}{2}Q^\dagger b$ *minimizes* q *and the corresponding infimal value is*

$$\inf_{y\in\mathbb{R}^p} q(y) = q(\bar{y}) = c - \tfrac{1}{4}b^\top Q^\dagger b\,.$$

Proof. We have

$$QQ^\dagger = \sum_{i,j=1}^{p-k} \lambda_i\lambda_j^{-1} q_i(q_i^\top q_j)q_j^\top = \sum_{i=1}^{p-k} q_iq_i^\top\,.$$

Because the q_i's form an orthonormal basis of $\mathcal{R}(Q)$, we can write $b = \sum \beta_i q_i$, with $\beta_i = q_i^\top b$. This gives

$$QQ^\dagger b = \sum_{i,j=1}^{p-k} q_iq_i^\top q_jq_j^\top b = \sum_{i=1}^{p-k} q_iq_i^\top b = b$$

(actually, $QQ^\dagger$ characterizes the orthogonal projection onto $\mathcal{R}(Q)$). Then $Q\bar{y} = -\frac{1}{2}QQ^\dagger b = -\frac{1}{2}b$, *i.e.*, $\bar{y}$ minimizes q. The computation of $q(\bar{y})$ is then straightforward. □

These elementary results can be used to prove a lemma very useful in the analysis of linear matrix inequalities. It is everyday bread in control theory, [2, Chap. 2], and will be useful for us as well. Here we give a simple proof.

Lemma 3.8 (Schur's Lemma) *For* $Q \in \mathcal{S}_p$, $P \in \mathcal{S}_n$ *and* $S = [s_1, \dots, s_p] \in \mathbb{R}^{n\times p}$, *the following two statements are equivalent:*

(i) The matrix $\begin{bmatrix} P & S^\top \\ S & Q \end{bmatrix}$ *is positive [resp. semi-]definite.*

(ii) The matrices Q *and* $P - S^\top Q^\dagger S$ *are both positive [resp. semi-]definite, and* $s_i \in \mathcal{R}(Q)$, *for* $i = 1, \dots, p$; *these last range conditions are automatically satisfied when* $Q \succ 0$ *(then* $\mathcal{R}(Q) = \mathbb{R}^p$*).*

Proof. For $(x, y) \in \mathbb{R}^n \times \mathbb{R}^p$, define the function

$$q_x(y) := \begin{pmatrix} x^\top & y^\top \end{pmatrix} \begin{bmatrix} P & S^\top \\ S & Q \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x^\top P x + 2\,(Sx)^\top y + y^\top Q y$$

and consider Lemma 3.6 with $b = 2Sx$, $c = x^\top Px$; observe that (i) means exactly $q_x(y) > 0$ [resp. $\geqslant 0$] for all nonzero $(x, y) \in \mathbb{R}^n \times \mathbb{R}^p$.

[$(i) \Rightarrow (ii)$] Under (i), we always have $Sx \in \mathcal{R}(Q)$.

– Take $x = 0$ to see that (i) implies $Q \succ 0$ [resp. $\succcurlyeq 0$].

– Then take successively $x = e_i$, $i = 1, \dots, p$ so that $Sx = s_i$, which lies in $\mathcal{R}(Q)$.

Finally use Lemma 3.7 to write

$$\inf_{y \in \mathbb{R}^p} q_x(y) = x^\top Px - (Sx)^\top Q^\dagger Sx = x^\top (P - S^\top Q^\dagger S)x .$$

This must be positive [resp. nonnegative] for all nonzero $x \in \mathbb{R}^n$, which was the last property we wanted.

[$(ii) \Rightarrow (i)$] Under (ii), we can use again Lemma 3.7 to write for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^p$,

$$q_x(y) \geqslant \inf_{y \in \mathbb{R}^p} q_x(y) = x^\top (P - S^\top Q^\dagger S)x .$$

To say that this number is positive [resp. nonnegative] is just (i). □

We give now two applications of Schur's Lemma.

Lemma 3.9 (Rank-one constraints) *The set of $(X, x) \in \mathcal{S}_n \times \mathbb{R}^n$ satisfying $X \succcurlyeq xx^\top$ is closed and convex. Actually,*

$$X - xx^\top \succcurlyeq 0 \text{ [resp. } \succ 0] \quad \Longleftrightarrow \quad \begin{bmatrix} X & x \\ x^\top & 1 \end{bmatrix} \succcurlyeq 0 \text{ [resp. } \succ 0].$$

Proof. The set of (X, x)'s described by the right part of the equivalence is the pre-image of the closed [resp. open] convex set $\mathcal{S}^+_{n+1}$ [resp. $\mathcal{S}^{++}_{n+1}$] by an affine operator; as such, it is closed [resp. open] and convex. To prove the equivalence, apply Lemma 3.8 with $p = 1$, $P = X$, $S = x^\top$ and $Q = 1$. □

Lemma 3.10 (Debreu's Lemma) *Let $A \in \mathbb{R}^{n \times m}$ and $H \in \mathcal{S}_n$ be positive definite on $\mathcal{N}(A)$ ($x^\top Hx > 0$ whenever $Ax = 0$ and $x \neq 0$). Then $H + \pi AA^\top$ is positive definite for all $\pi > 0$ large enough.*

Proof. Calling m' the rank of A, let $Z_{\mathcal{N}}$ and $Z_{\mathcal{R}}$ be respectively $n \times (n-m')$ and $n \times m'$ matrices whose columns form orthonormal bases of $\mathcal{N}(A)$ and $\mathcal{N}(A)^\perp$. The columns of the $n \times n$ matrix $B := [Z_{\mathcal{R}}\, Z_{\mathcal{N}}]$ are orthonormal: B is nonsingular, $B^{-1} = B^\top$ and a matrix M is positive definite if and only if $B^\top M B$ is. Using the block structure of B, we have

$$B^\top M B = \begin{bmatrix} Z_{\mathcal{R}}^\top \\ Z_{\mathcal{N}}^\top \end{bmatrix} M[Z_{\mathcal{R}}\, Z_{\mathcal{N}}] = \begin{bmatrix} Z_{\mathcal{R}}^\top M Z_{\mathcal{R}} & Z_{\mathcal{R}}^\top M Z_{\mathcal{N}} \\ Z_{\mathcal{N}}^\top M Z_{\mathcal{R}} & Z_{\mathcal{N}}^\top M Z_{\mathcal{N}} \end{bmatrix}$$

Since $AZ_{\mathcal{N}} = 0$, we obtain in particular

$$B^\top[H + \pi A^\top A]B = \begin{bmatrix} Z_{\mathcal{R}}^\top H Z_{\mathcal{R}} + \pi Z_{\mathcal{R}}^\top A^\top A Z_{\mathcal{R}} & Z_{\mathcal{R}}^\top H Z_{\mathcal{N}} \\ Z_{\mathcal{N}}^\top H Z_{\mathcal{R}} & Z_{\mathcal{N}}^\top H Z_{\mathcal{N}} \end{bmatrix} =: \begin{bmatrix} P & S^\top \\ S & Q \end{bmatrix}.$$

Now apply Lemma 3.8. Observe that $Q = Z_{\mathcal{N}}^\top H Z_{\mathcal{N}}$ is positive definite (by assumption); it is therefore “onto” and the only thing we need is $P - S^\top Q^\dagger S \succ 0$, *i.e.*,

$$Z_{\mathcal{R}}^\top H Z_{\mathcal{R}} + \pi Z_{\mathcal{R}}^\top A^\top A Z_{\mathcal{R}} \succ Z_{\mathcal{R}}^\top H Z_{\mathcal{N}} (Z_{\mathcal{N}}^\top H Z_{\mathcal{N}})^{-1} Z_{\mathcal{N}}^\top H Z_{\mathcal{R}}. \tag{20}$$

We claim that $Z_{\mathcal{R}}^\top A^\top A Z_{\mathcal{R}}$, which obviously lies in $\mathcal{S}_{m'}^+$, is actually positive definite: indeed, for all $v \in \mathbb{R}^{m'}$, the property $v^\top Z_{\mathcal{R}}^\top A^\top A Z_{\mathcal{R}} v = \|AZ_{\mathcal{R}} v\|^2 = 0$ implies

$$Z_{\mathcal{R}} v \in \mathcal{N}(A) \cap \mathcal{R}(A) = \mathcal{N}(A) \cap \mathcal{N}(A)^\perp = \{0\},$$

and hence $v = 0$.

Thus, with simplified notation, we have to prove that a certain matrix $\pi M - N$ is positive definite. Calling R (a symmetric positive definite matrix) the square root of $M = Z_{\mathcal{R}}^\top A^\top A Z_{\mathcal{R}} \succ 0$, we have for all nonzero $v \in \mathbb{R}^{m'}$:

$$v^\top(\pi M - N)v = \pi v^\top R R v - v^\top N v = \pi \|Rv\|^2 - v^\top R R^{-1} N R^{-1} R v.$$

Setting $w = Rv$ ($\neq 0$), this is equal to $\pi\|w\|^2 - w^\top R^{-1} N R^{-1} w$, which is a positive number for π large enough, namely for $\pi > \lambda_1(R^{-1}NR^{-1})$. □

4 Lagrangian duality for quadratic programs

The material introduced in Sections 2 and 3 is just what is needed to apply Lagrange duality to quadratic optimization problems.

Exemple AG m.q. $\frac{1}{2}x_1 + \frac{x_2}{x_1^2} + \frac{1}{4}\frac{1}{x_2} \geqslant \sqrt{2}$ $\forall x_1>0, x_2>0$ et qu'il y a égalité ssi $x_1=\sqrt{2}$ et $x_2=\frac{1}{\sqrt{2}}$

$\forall x_1>0$, $x_2>0$

$$\frac{1}{2}\boxed{x_1} + \frac{1}{4}\times\boxed{4\frac{x_2}{x_1^2}} + \frac{1}{4}\boxed{\frac{1}{x_2}} \underset{\uparrow \text{AG}}{\geqslant} [x_1]^{1/2}\left[4\frac{x_2}{x_1^2}\right]^{1/4}\left[\frac{1}{x_2}\right]^{1/4} = \underbrace{x_1^{1/2-1/2}}_{1}\ \underbrace{x_2^{1/4-1/4}}_{1}\ 4^{1/4} = 4^{1/2\times\frac{1}{2}} = 2^{1/2} = \sqrt{2}$$

égalité ssi $x_1 = 4\frac{x_2}{x_1^2} = \frac{1}{x_2} = A \rightarrow$

$$\frac{1}{2}A + \frac{1}{4}A + \frac{1}{4}A = \sqrt{2} \Rightarrow A=\sqrt{2} \Rightarrow x_2 = \frac{1}{\sqrt{2}}$$

$$x_1 = \sqrt{2}$$

vérifions : $\frac{1}{2}\boxed{\sqrt{2}} + \frac{1}{4}\boxed{\frac{4}{\sqrt{2}}\times\frac{1}{2}} + \frac{1}{4}\times\boxed{\sqrt{2}} = \sqrt{2}\times\frac{3}{4} + \frac{1}{\sqrt{2}}\times\frac{1}{2} = \sqrt{2}\left[\frac{3}{4}+\frac{1}{4}\right] = \sqrt{2}$

Inégalité AG

On considère le pb : maximiser $z = x_1 x_2^2 x_3$ s.c $x_1 + x_2 + x_3^2 = k$, $x_1 > 0$, $x_2 > 0$, $x_3 > 0$

où k est une constante positive.

En utilisant l'inégalité AG, montrer que $7 \times \left(\frac{1}{2}\right)^{2/7} \times \left(\frac{1}{4}\right)^{4/7} z^{2/7} \leq k$

et que la borne k peut être atteinte.

En déduire la solution du pb.

$$k = x_1 + x_2 + x_3^2 = 7 \times \frac{1}{7} \left(\frac{x_1}{2} + \frac{x_1}{2} + \frac{x_2}{4} + \frac{x_2}{4} + \frac{x_2}{4} + \frac{x_2}{4} + x_3^2 \right) \underset{\underset{AG}{\uparrow}}{\geq} 7 \times \left(\frac{1}{2}\right)^{2/7} \times \left(\frac{1}{4}\right)^{4/7} \times \big(\underbrace{x_1^2 x_2^4 x_3^2}_{z^2}\big)^{1/7}$$

on a égalité dans AG ssi $\frac{x_1}{2} = \frac{x_2}{4} = x_3^2$ et dans ce cas, la borne k (constante) étant atteinte, z est maximisé.

Pour satisfaire la contrainte $x_1 + x_2 + x_3^2 = k$ (solution réalisable du pb)

on a, en exprimant tout en fonction de x_1 : $x_1 + 2x_1 + \frac{x_1}{2} = k$

$\Rightarrow \frac{7}{2} x_1 = k \Rightarrow x_1 = \frac{2}{7} k \Rightarrow x_2 = \frac{4}{7} k$ et $x_3^2 = \frac{1}{7} k \Rightarrow x_3 = \sqrt{\frac{k}{7}} > 0$

Programmation géométrique.

Exemple : minimiser $4x_1 + \frac{x_1}{x_2^2} + 4\frac{x_2}{x_1}$ sous les contraintes $x_1 > 0,\ x_2 > 0$

$$g(x) = \delta_1\left(\frac{4x_1}{\delta_1}\right) + \delta_2\left(\frac{x_1}{\delta_2 x_2^2}\right) + \delta_3\left(\frac{4x_2}{\delta_3 x_1}\right) \overset{AG}{\geqslant} \left(\frac{4x_1}{\delta_1}\right)^{\delta_1}\left(\frac{x_1}{\delta_2 x_2^2}\right)^{\delta_2}\left(\frac{4x_2}{\delta_3 x_1}\right)^{\delta_3} = \left(\frac{4}{\delta_1}\right)^{\delta_1}\left(\frac{1}{\delta_2}\right)^{\delta_2}\left(\frac{4}{\delta_3}\right)^{\delta_3} x_1^{\delta_1+\delta_2-\delta_3}\, x_2^{-2\delta_2+\delta_3}$$

où $\delta_1 > 0,\ \delta_2 > 0,\ \delta_3 > 0$ $\quad \delta_1+\delta_2+\delta_3 = 1 \quad (2)$

on rajoute $\begin{cases} \delta_1+\delta_2-\delta_3 = 0 \\ -2\delta_2+\delta_3 = 0 \end{cases} \quad (3)$

(2) et $(3) \Rightarrow \delta_1^* = \frac{1}{4} \quad \delta_2^* = \frac{1}{4} \quad \delta_3^* = \frac{1}{2}$ (ici le dual a un seul point réalisable)

on en déduit les solution du problème (primal)

$$\frac{4x_1^*}{g(x^*)} = \frac{1}{4} \qquad \frac{x_1^*}{x_2^{2*}\, g(x^*)} = \frac{1}{4} \qquad \frac{4x_2^*}{x_1^*\, g(x^*)} = \frac{1}{2}$$

$$g(x^*) = v(\delta^*) = 8$$

$$\Rightarrow x_1^* = \frac{1}{2} \qquad x_2^* = \frac{1}{2}$$

a/ $f(x)$ définie sur $\mathbb{R}^3$ par

$f(x_1, x_2, x_3) = e^{x_1^2 + x_2^2 + x_3^2}$ est-elle convexe ? strictement convexe ?

$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)$
$\forall x, y \quad x \neq y$ et $0 < \lambda < 1$

$h(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$ est strictement convexe car son Hessien

$$H_h = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

est défini positif.

$g(t) = e^t \quad t \in \mathbb{R}$ est $\nearrow$ car $g'(t) = e^t > 0$ et même strictement $\nearrow$.

$g(t)$ est convexe (le strictement n'est pas nécessaire) car $g''(t) = e^t > 0$

Donc : $f(x) = g(h(x))$ est strictement convexe

déterminer le minimum de f : $\nabla f = \begin{bmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{bmatrix} e^{x_1^2 + x_2^2 + x_3^2} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow x_1 = x_2 = x_3 = 0$ unicité

b/ $\begin{cases} a^{(i)} \in \mathbb{R}^n & \forall i = 1, \dots, K \\ c_i > 0 \in \mathbb{R} & \forall i = 1, \dots, K \end{cases}$ $\quad f(x) = \sum_{i=1}^{K} c_i e^{a^{(i)} \cdot x} \quad x \in \mathbb{R}^n$ est-elle convexe ?

$g_i(x) = a^{(i)} \cdot x \quad i = 1, \dots, K$ est linéaire donc convexe sur $\mathbb{R}^n$.

$h(t) = e^t$ est croissante et convexe $\Rightarrow h(g_i(x)) = e^{a^{(i)}x}$ est convexe $\forall i = 1, \dots, K$

$c_i \, h(g_i(x))$ est convexe car $c_i > 0$

$\sum_i c_i \, h(g_i(x))$ est convexe car somme de fonctions convexes.

$f = e^{x_1^2+x_2^2+x_3^2}$ Montrer que f est convexe par la méthode "directe" (à l'aide du hessien).

$$\nabla f = \begin{bmatrix} 2x_1 e^A \\ 2x_2 e^A \\ 2x_3 e^A \end{bmatrix} \qquad H_f = \begin{bmatrix} 2e^A + 4x_1^2 e^A & 4x_1x_2 e^A & 4x_1x_3 e^A \\ 4x_1x_2 e^A & 2e^A + 4x_2^2 e^A & 4x_2x_3 e^A \\ 4x_1x_3 e^A & 4x_2x_3 e^A & 2e^A + 4x_3^2 e^A \end{bmatrix}$$

en posant $A = x_1^2 + x_2^2 + x_3^2$

$$= 2e^A \begin{bmatrix} 1+2x_1^2 & 2x_1x_2 & 2x_1x_3 \\ 2x_1x_2 & 1+2x_2^2 & 2x_2x_3 \\ 2x_1x_3 & 2x_2x_3 & 1+2x_3^2 \end{bmatrix}$$

Etude du hessien :

$$\Delta_1 = 1+2x_1^2 > 0$$

$$\Delta_2 = (1+2x_1^2)(1+2x_2^2) - 4x_1^2x_2^2 = 1+2x_1^2+2x_2^2 > 0$$

$$\Delta_3 = (1+2x_1^2)\left[(1+2x_2^2)(1+2x_3^2) - 4x_2x_3\right]$$

$$- 2x_1x_2\left[2x_1x_2(1+2x_3^2) - 4x_1x_2x_3^2\right]$$

$$+ 2x_1x_3\left[4x_1x_3x_2^2 - 2x_1x_3(1+2x_2^2)\right]$$

$$= (1+2x_1^2)\left[1+2x_2^2+2x_3^2\right]$$

$$-4x_1^2x_2^2 - 4x_1^2x_3^3 \quad = 1+2x_2^2+2x_3^2+2x_1^2 > 0$$

H_f défini positif $\Rightarrow$ f strict^t convexe.

On arrive à démontrer le résultat mais c'est tout de même plus long qu'en utilisant le th. sur les compositions de fonctions.

Convexité

Soit $g(x)$ fonction concave sur S (S convexe de $\mathbb{R}^n$)

$h(y)$ fonction décroissante et convexe sur $g(S)$

1) Montrer que $f = h \circ g$ est convexe sur S

2) Application :

soit g concave sur S et $g(x) > 0 \quad \forall x \in S$

Montrer que $f(x) = \dfrac{1}{g(x)}$ est convexe sur S

---

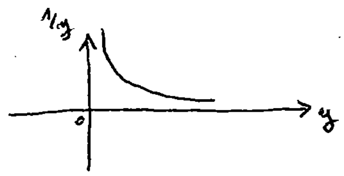
1) $$g((1-\vartheta)x + \vartheta y) \underset{\substack{\uparrow \\ g \text{ concave}}}{\geq} (1-\vartheta)\,g(x) + \vartheta\, g(y) \qquad x, y \in S \; , \; 0 \leq \vartheta \leq 1$$

$$h[\qquad\qquad] \underset{\substack{\uparrow \\ h \text{ décroissante}}}{\leq} h[\qquad\qquad] \underset{\substack{\uparrow \\ h \text{ convexe}}}{\leq} (1-\vartheta)\, h(g(x)) + \vartheta\, h(g(y))$$

emarque : on voit que si g est strict^t concave et h strict^t décroissante, alors $h \circ g$ est strict^t convexe.

2) $h(y) = \dfrac{1}{y}$ pour $y > 0$ est décroissante et convexe

$$h'(y) = -\frac{1}{y^2} < 0 \qquad h''(y) = \frac{2}{y^3} > 0 \quad \text{pour } y > 0$$



on applique la question 1) : $f = h \circ g = \dfrac{1}{g(x)}$ est convexe sur S (on a $\underbrace{g(x) > 0}_{\text{par hypothèse}}$ sur S)

M.q. $f(x_1, x_2) = \dfrac{1}{-x_1^2 - x_2^2 + 2x_1 + x_2 + \frac{15}{4}}$ est convexe sur la boule $\|x\| < \frac{\sqrt{5}}{2}$

où $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ et $\|x\| = \sqrt{x_1^2 + x_2^2}$

posons $g(x_1, x_2) = -x_1^2 - x_2^2 + 2x_1 + x_2 + \frac{15}{4}$

- montrons que $g(x_1, x_2) > 0$ sur la boule $\|x\| < \sqrt{5}/2$

$g(x_1, x_2) = -\|x\|^2 + b \cdot x + \frac{15}{4}$ en posant $b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

Cauchy-Schwarz $\Rightarrow$ $b \cdot x \geqslant -\|b\| \|x\|$

$\Rightarrow g(x_1, x_2) \geqslant -\|x\|^2 - \sqrt{5}\,\|x\| + \frac{15}{4} = -\left(\|x\|^2 + \sqrt{5}\,\|x\|\right) + \frac{15}{4} = -\left(\|x\| + \frac{\sqrt{5}}{2}\right)^2 + \frac{5}{4} + \frac{15}{4}$

$\|x\| < \sqrt{5}/2 \Rightarrow \|x\| + \frac{\sqrt{5}}{2} < \sqrt{5} \Rightarrow \left(\|x\| + \frac{\sqrt{5}}{2}\right)^2 < 5 \Rightarrow -\left(\|x\| + \frac{\sqrt{5}}{2}\right)^2 > -5$

(x dans la boule)

$\Rightarrow g(x_1, x_2) > -5 + \frac{20}{4} = 0$

- montrons que $g(x_1, x_2)$ est concave.

$H_g = 2\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ matrice définie négative $\Rightarrow$ g concave (i.e $(-g)$ convexe)

on conclut en appliquant le résultat précédent.

Contrôle MOM
(Septembre 2007)

Durée 1h30.
Documents autorisés. Calculettes interdites.

<u>Rappels</u> :
$u \bullet v$ désigne le produit scalaire euclidien des 2 vecteurs u et v. $\|u\|$ désigne la norme (norme2) du vecteur u.

Exemples : $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, u \bullet v = u_1 v_1 + u_2 v_2, \|u\| = \sqrt{u \bullet u} = \sqrt{u_1^2 + u_2^2}$.

La transposée d'une matrice A est notée A^T. Exemple : $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, A^T = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$.

Noter que si u et v sont deux vecteurs colonnes alors $u \bullet v = u^T v$.
Une matrice carrée symétrique A est semi-définie positive (resp. définie positive) si $x \bullet Ax \geq 0 \ \forall x$ (resp. >0 $\forall x \neq 0$).
Soit f la fonction quadratique définie sur $\Re^n$ par $f(x) = \frac{1}{2} x \bullet Qx + q \bullet x$ avec Q une matrice carrée symétrique nxn, q un vecteur colonne de n coordonnées, x (la variable) un vecteur colonne de n coordonnées. Le gradient de f est donné par $\nabla f(x) = Qx + q$ et le hessien de f par $H_f = Q$.

<u>Problème 1</u> : La méthode des moindres carrés.
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'inconnues ($m>n$) et est souvent 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. On peut vérifier facilement que $\|Ax-b\|^2 = x \bullet A^T Ax - 2A^T b \bullet x + b \bullet b$.

1) Montrer que cette fonction est convexe.

2) Montrer que le x minimiseur de $\|Ax-b\|^2$ vérifie l'équation dite normale $A^T Ax = A^T b$.

3) On considère le problème suivant.

L'ingénieur Emile 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, Emile 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. A nouveau, il constate l'inconsistence des résultats avec les précédents. Emile 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.

Méthode des moindres carrés.

système $Ax=b$ avec m lignes et n inconnues et $m>n$

$$\|Ax-b\|^2 = x\cdot A^TAx - 2A^Tb\cdot x + b\cdot b$$

1) M.q. la fonction est convexe.

2 façons : a) hessien de la fonction - on remarque déjà que A^TA est symétrique $(A^TA)^T = A^T(A^T)^T = A^TA$

$H_f = 2A^TA \Rightarrow x\cdot H_f x = 2x^TA^TAx \Rightarrow x\cdot H_f x = 2\|Ax\|^2 \geq 0 \quad \forall x \Rightarrow$ hessien semi-défini positif $\Rightarrow f$ convexe

b) $\|A[(1-\lambda)x+\lambda y]-b\| = \|A[1-\lambda]x + A\lambda y - [[1-\lambda]b+\lambda b]\| = \|(1-\lambda)[Ax-b]+\lambda[Ay-b]\| \leq \|(1-\lambda)[Ax-b]\| + \|\lambda(Ay-b)\|$

prop de la norme (inégalité triangulaire)

$= (1-\lambda)\|Ax-b\| + \lambda\|Ay-b\|$ donc $\|Ax-b\|$ est convexe.

$(1-\lambda)\geq 0$

$\lambda \geq 0$

maintenant z^2 est convexe et croissante pour $z\geq 0$ donc par th. sur les fonctions composées $\|Ax-b\|^2$ est convexe.

2) M.q. le minimiseur de f vérifie l'équation dite normale.

f est convexe donc son minimiseur est donné par un point critique (un point qui annule le gradient)

ici $\nabla f = 2A^TAx - 2A^Tb \Rightarrow \nabla f = 0 \Leftrightarrow A^TAx - A^Tb = 0 \Leftrightarrow A^TAx = A^Tb$

3) Résolution du pb d'Emile

Emile fait 6 mesures de h_1, h_2, h_3 t.q. h_1, h_2, h_3 doivent vérifier le système (inconsistent) suivant :

$$\begin{array}{ll} h_1 = 36 & h_2-h_1 = 11 \\ h_2 = 41 & h_3-h_1 = 77 \\ h_3 = 117 & h_3-h_2 = 75 \end{array} \Rightarrow A = \begin{bmatrix} 1&0&0\\0&1&0\\0&0&1\\-1&1&0\\-1&0&1\\0&-1&1 \end{bmatrix} \text{ et } A^T = \begin{bmatrix} 1&0&0&-1&-1&0\\0&1&0&1&0&-1\\0&0&1&0&1&1 \end{bmatrix} \Rightarrow A^TA = \begin{bmatrix} 3&-1&-1\\-1&3&-1\\-1&-1&3 \end{bmatrix}$$

$$b = \begin{bmatrix}36\\41\\117\\11\\77\\75\end{bmatrix} \quad \text{et} \quad A^Tb = \begin{bmatrix} 36-11-77\\41+11-75\\117+77+75 \end{bmatrix} = \begin{bmatrix}-52\\-23\\269\end{bmatrix}$$

équation normale : $\begin{cases} 3h_1-h_2-h_3 = -52 \\ -h_1+3h_2-h_3 = -23 \\ -h_1-h_2+3h_3 = 269 \end{cases} \Rightarrow h_1+h_2+h_3 = -52-23+269 = 194 \Rightarrow \begin{cases} 4h_1 = -52+194 = 142 \\ 4h_2 = -23+194 = 171 \\ 4h_3 = 269+194 = 463 \end{cases} \Rightarrow \boxed{\begin{array}{l} h_1 = \frac{142}{4} = 35{,}5 \\ h_2 = \frac{171}{4} = 42{,}75 \\ h_3 = \frac{463}{4} = 115{,}75 \end{array}}$

Pb quadratique sous contraintes linéaires.

P) $\min \frac{1}{2}x\cdot Hx$ s.c. $Ax=b$ avec H déf >0, A matrice de m lignes, b vecteur de m lignes

1) Kühn - Tucker.

on met $Ax=b$ sous la forme $Ax-b=0$.

K.T $\begin{cases} Hx + A^T\mu = 0 \\ Ax=b \end{cases}$ gradient de la fonction + comb. lin. des gradients des contraintes d'égalité. μ vecteur de m coordonnées

x doit satisfaire les contraintes

H est inversible $\Rightarrow \begin{cases} x = -H^{-1}A^T\mu \\ -AH^{-1}A^T\mu = b \end{cases}$ en posant $w=-\mu$ on obtient $\begin{cases} x = H^{-1}A^Tw \\ AH^{-1}A^Tw = b \end{cases}$

le pb (P) étant convexe (fonction convexe, contraintes égalité affines)

K.T. est suffisant pour l'optimum de (P)

2) Application : trouver la solution de norme minimum du système $\begin{cases} 2x_1+x_2+x_3+5x_4 = 8 \\ -x_1-x_2+3x_3+2x_4 = 0 \end{cases}$

on minimisera $\frac{1}{2}x\cdot x$ donc $H=I$

et $A = \begin{bmatrix} 2&1&1&5\\-1&-1&3&2 \end{bmatrix}$, $b = \begin{bmatrix}8\\0\end{bmatrix}$

on applique les formules de la question 1)

$A^T = \begin{bmatrix} 2&-1\\1&-1\\1&3\\5&2 \end{bmatrix}$ $\quad AA^T = \begin{bmatrix} 4+1+1+25 & -2-1+3+10 \\ \text{〃} & 1+1+9+4 \end{bmatrix} = \begin{bmatrix} 31&10\\10&15 \end{bmatrix}$, on cherche w $\begin{cases} 31w_1+10w_2 = 8 \\ 10w_1+15w_2 = 0 \end{cases} \Rightarrow \boxed{\begin{array}{l} w_1 = \frac{24}{73} \\ w_2 = -\frac{16}{73} \end{array}}$

vérifions $\begin{cases} \frac{8}{73}[31\times 3 - 10\times 2] = \frac{8}{73}\times 73 = 8 \\ \frac{8}{73}[10\times 3 - 15\times 2] = \frac{8}{73}\times 0 = 0 \end{cases}$

on obtient x : $x = A^Tw = \frac{24}{73}\begin{bmatrix}2\\1\\1\\5\end{bmatrix} - \frac{16}{73}\begin{bmatrix}-1\\-1\\3\\2\end{bmatrix} = \frac{8}{73}\left[\begin{bmatrix}6\\3\\3\\15\end{bmatrix} + \begin{bmatrix}2\\2\\-6\\-4\end{bmatrix}\right] = \boxed{\frac{8}{73}\begin{bmatrix}8\\5\\-3\\11\end{bmatrix}}$

vérifions que x satisfait le système

$\left.\begin{array}{l} 2\times 8+5-3+5\times 11 = 16+2+55 = 73 \\ -8-5-3\times 3+2\times 11 = -22+22 = 0 \end{array}\right\}$ en multipliant par $\frac{8}{73}$ on obtient bien $\begin{pmatrix}8\\0\end{pmatrix}$

Problème de moindres carrés

Soit le système linéaire $Ax = b$ $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ tel que $m > n$

c'est-à-dire un système linéaire avec plus d'équations que d'inconnues.

On cherche une solution x^* telle que Ax^* soit le plus "proche" possible de b au sens de la norme euclidienne c'est-à-dire on cherche à résoudre le problème

minimiser $\|Ax - b\|^2$ x parcourant $\mathbb{R}^n$

0) mettre ce pb sous la forme de la minimisation d'une fonction quadratique $q(x)$ que l'on explicitera

1) Montrer que x^* est solution du problème ssi x^* est solution de l'équation dite normale $A^TAx = A^Tb$

2) Montrer que si A est de rang n alors x^* est unique

Remarque : la matrice $(A^TA)^{-1}A^T$ est appelée inverse généralisé de A

0) $\|Ax - b\|^2 = (Ax - b) \bullet (Ax - b) = (Ax) \bullet (Ax) - 2\, b \bullet (Ax) + b \bullet b$

$= x^TA^TAx - 2b^TAx + b \bullet b$ rappel : $a \bullet b = a^T b$

$= x \bullet (A^TAx) - 2(A^Tb) \bullet x + b \bullet b$

on a donc une forme quadratique $q(x)$ avec A^TA symétrique

1) condition nécessaire d'optimalité à l'ordre 1 :

$\nabla q(x) = +2A^TAx - 2A^Tb = 0 \Rightarrow A^TAx = A^Tb$

hessien de $q(x)$: $H_q = 2A^TA$

$y \bullet H_q y = y^T H_q y = 2\, y^TA^TAy = 2\,(Ay) \bullet (Ay) = 2\,\|Ay\|^2 \geqslant 0$

donc H_q est semi-défini positif (partout puisque constant)

et donc q est convexe et donc la condition nécessaire du 1^er^ ordre est suffisante.

2) Si A est de rang n alors n colonnes lin^t^ ind^t^

alors $\|Ay\| = 0 \Rightarrow Ay = 0 \Rightarrow y = 0$

et donc H_q est défini positif et donc q est stric^t^ convexe et donc le minimiseur de q est unique

On a un système représenté par une "boîte noire" admettant une entrée e et une sortie s

On a effectué 3 relevés de s pour 3 valeurs de e. Les couples (e, s) mesurés sont $(1,2)$, $(2,3)$, $(3,3)$. On voudrait établir une relation entre e et s de la forme $s = f(e) = x_0 + x_1 e$ qui tienne compte au mieux de ces 3 mesures.

Modéliser ce problème comme un problème de moindres carrés et le résoudre.

On a le système à 3 équations et 2 inconnues à résoudre :

$$x_0 + x_1 = 2$$
$$x_0 + 2x_1 = 3$$
$$x_0 + 3x_1 = 3$$

Approcher au mieux le second membre est un problème de moindres carrés

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \quad b = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} \quad A^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \quad A^TA = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \quad A^Tb = \begin{bmatrix} 8 \\ 17 \end{bmatrix}$$

$$A^TAx = A^Tb \iff \begin{cases} 3x_0 + 6x_1 = 8 \\ 6x_0 + 14x_1 = 17 \end{cases} \Rightarrow \begin{cases} x_0 = 5/3 \\ x_1 = \frac{1}{2} \end{cases}$$

Par curiosité regardons les écarts par rapport à b de la solution :

$$5/3 + 1/2 = \frac{13}{6} \quad \text{au lieu de } 2 = 12/6$$
$$5/3 + 1 = 8/3 \quad \text{"} \quad \text{"} \quad \text{"} \quad 3 = 9/3$$
$$5/3 + 3/2 = 19/6 \quad \text{"} \quad \text{"} \quad \text{"} \quad 3 = 18/6$$

<u>Projection d'un vecteur sur un sous-espace.</u>

On désire trouver le vecteur du sous-espace de $\mathbb{R}^3$ engendré par $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ le plus proche de $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ au sens de la norme euclidienne.

Montrer que ce problème revient à un problème de moindres carrés et le résoudre.

Posons $u = x_1\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + x_2\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \quad (x_1, x_2) \in \mathbb{R}^2$

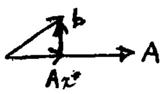
$$= \begin{bmatrix} 1 & 2 \\ 1 & -1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = Ax \quad , \text{ posons } b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

on veut minimiser $\|u - b\|$ ce qui est équivalent à minimiser $\|u - b\|^2$ car x^2 est croissante pour $x \geq 0$ soit minimiser $\|Ax - b\|^2$ donc un problème de moindres carrés

$$A^TA = \begin{bmatrix} 1 & 1 & 2 \\ 2 & -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ 3 & 6 \end{bmatrix} \quad A^Tb = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \quad \text{équation normale} = \begin{cases} 6x_1 + 3x_2 = 4 \\ 3x_1 + 6x_2 = 2 \end{cases} \Rightarrow \begin{cases} x_1 = 2/3 \\ x_2 = 0 \end{cases}$$

on peut vérifier que $r = b - Ax = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2/3 \\ 2/3 \\ 4/3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \\ -1/3 \end{bmatrix}$ est $\perp$ aux vecteurs colonnes de A

<u>Interprétation géométrique</u> : chercher le vecteur du sous espace engendré par les colonnes de A le plus proche de b revient à faire la projection de b sur ce sous espace :



on vérifie que $r = b - Ax^*$ est $\perp$ au sous-espace engendré par les colonnes de A c'est-à-dire vérifie $A^Tr = 0$

Localisation d'équipements dans un réseau avec fonctions de coûts concaves

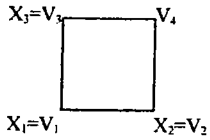
Soit un graphe G=(V,E) connexe, non orienté, avec les arêtes munies de coûts positifs. On peut voir G comme le modèle d'un réseau routier où les sommets sont les croisements, les arêtes les routes, le coût d'une arête représentant le coût de parcours d'une route. Les sommets v_i de G sont considérés comme des points de $\mathbb{R}^2$. En certains des sommets de G sont localisés des clients. Chaque client x_i est muni d'un poids a_i positif. On note N l'ensemble des clients.
On veut placer un équipement dans le graphe (le réseau) de sorte que la somme pondérée des distances des clients à l'équipement soit minimum. L'équipement n'est pas nécessairement placé en un sommet du graphe, il peut être placé sur une arête. Plus formellement, soit z le point de $\mathbb{R}^2$ représentant la position de l'équipement dans G, on veut résoudre le problème :

$$(P) \qquad \min_{z\in G} f(z) = \sum_{x_i\in N} a_i \mathrm{d}(x_i, z)$$

où $\mathrm{d}(x_i, z)$ est la distance, au sens des coûts des arêtes, du client x_i à l'équipement z. La distance de x_i à z est la valeur minimum du coût d'un chemin de x_i à z.

Considérons une arête quelconque (v_1,v_2) et $z\in(v_1,v_2)$. On peut écrire $z(\lambda)=(1-\lambda)v_1+\lambda v_2$ pour $0\leq\lambda\leq 1$. On note $c(v_i,z)$ le coût pour aller de v_i à z (i=1,2) en empruntant l'arête (v_1,v_2).

Exemple : considérons le graphe suivant



Avec chaque arête munie d'un coût égal à 1. Chaque client x_1, x_2, x_3 est muni d'un poids égal à 1. Considérons que $c(v_1,z(\lambda))=\lambda^2$, $c(v_2,z(\lambda))=(1-\lambda)^2$.
Si l'équipement z est placé au milieu de l'arête (v_1,v_2) c'est-à-dire si $\lambda=1/2$ alors f(z)=0,25+0,25+1,25=1,75
Si l'équipement z est placé en v_1 c'est-à-dire si $\lambda=0$ alors f(z)=0+1+1=2
Si l'équipement z est placé en v_2 c'est-à-dire si $\lambda=1$ alors f(z)=1+0+2=3

On suppose maintenant que les fonctions $c(v_i,z(\lambda))$ sont des fonctions de λ concaves pour i=1,2.

1) Propriété générale : soit h_1 et h_2 deux fonctions concaves définies sur un convexe C de $\mathbb{R}^n$, montrer que $h(x)=\min\{h_1(x),h_2(x)\}$ est une fonction concave.

2) a) Soit x_i un client, montrer que $d(x_i,z(\lambda))$ est une fonction (de λ) concave.
b) Montrer que $f(z(\lambda))$est une fonction (de λ) concave.
c) Montrer qu'une solution optimale de (P) (si l'on met l'équipement sur l'arête (v_1,v_2)) est obtenue en $z=v_1$ ou $z=v_2$.

3) Conclusion : sous l'hypothèse de concavité des fonctions $c(v_i,z)$, $c(v_j,z)$ pour toute arête (v_i,v_j) de G et $z\in(v_i,v_j)$, déduire du 2.c qu'une solution optimale de (P) est atteinte en plaçant l'équipement sur l'un des sommet de G.

<u>Fonctions concaves</u>

Soit un graphe $G=(V,E)$ connexe, non orienté, avec les arêtes munies de coûts positifs. On peut voir G comme le modèle d'un réseau routier où les sommets sont les croisements, les arêtes les routes, le coût d'une arête représentant le coût de parcours d'une route. Les sommets v_i de G sont considérés comme des points de $\mathbb{R}^2$. En certains des sommets de G sont localisés des clients. Chaque client x_i est muni d'un poids a_i positif. On note N l'ensemble des clients.

On veut placer un équipement dans le graphe (le réseau) de sorte que la somme pondérée des distances des clients à l'équipement soit minimum. L'équipement n'est pas nécessairement placé en un sommet du graphe, il peut être placé sur une arête. Plus formellement, soit z le point de $\mathbb{R}^2$ représentant la position de l'équipement dans G, on veut résoudre le problème :

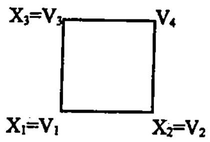
$$(P) \qquad \min_{z \in G} f(z) = \sum_{x_i \in N} a_i \mathrm{d}(x_i, z)$$

où $\mathrm{d}(x_i, z)$ est la distance, au sens des coûts des arêtes, du client x_i à l'équipement z.

la distance de x_i à z est la valeur minimale du coût d'un chemin de x_i à z

Considérons une arête quelconque (v_1,v_2) et $z \in (v_1,v_2)$. On peut écrire $z(\lambda)=(1-\lambda)v_1+\lambda v_2$ pour $0 \leq \lambda \leq 1$. On note $c(v_i,z)$ le coût pour aller de v_i à z ($i=1,2$) en empruntant l'arête (v_1,v_2).

<u>Exemple</u> : considérons le graphe suivant



Avec chaque arête munie d'un coût égal à 1. Chaque client x_1, x_2, x_3 est muni d'un poids égal à 1. Considérons que $c(v_1,z(\lambda))=\lambda^2$, $c(v_2,z(\lambda))=(1-\lambda)^2$.

Si l'équipement z est placé au milieu de l'arête (v_1,v_2) c'est-à-dire si $\lambda=1/2$ alors $f(z)=0{,}25+0{,}25+1{,}25=1{,}75$

Si l'équipement z est placé en v_1 c'est-à-dire si $\lambda=0$ alors $f(z)=0+1+1=2$

Si l'équipement z est placé en v_2 c'est-à-dire si $\lambda=1$ alors $f(z)=1+0+2=3$

On suppose maintenant que les fonctions $c(v_i,z(\lambda))$ sont des fonctions de λ concaves pour $i=1,2$.

1) <u>Propriété générale</u> : soit h_1 et h_2 deux fonctions concaves définies sur un convexe C de $\mathbb{R}^n$, montrer que $h(x)=\min\{h_1(x),h_2(x)\}$ est une fonction concave.

2) a) Soit x_i un client, montrer que $d(x_i,z(\lambda))$ est une fonction (de λ) concave.

b) Montrer que $f(z(\lambda))$est une fonction (de λ) concave.

c) Montrer qu'une solution optimale de (P) (si l'on met l'équipement sur l'arête (v_1,v_2)) est obtenue en $z=v_1$ ou $z=v_2$.

3) <u>Conclusion</u> : sous l'hypothèse de concavité des fonctions $c(v_i,z)$, $c(v_j,z)$ pour toute arête (v_i,v_j) de G et $z\in(v_i,v_j)$, déduire du 2.c qu'une solution optimale de (P) est atteinte en plaçant l'équipement sur l'un des sommet de G.

Fonctions concaves

1) $h((1-\lambda)x+\lambda y) = \min\{h_1[(1-\lambda)x+\lambda y],\ h_2[(1-\lambda)x+\lambda y]\}$

avec $0 \le \lambda \le 1$

$$h_1[(1-\lambda)x+\lambda y] \overset{\text{concavité}}{\geq} (1-\lambda)h_1(x)+\lambda h_1(y)$$
$$h_2[\quad] \geq (1-\lambda)h_2(x)+\lambda h_2(y)$$

$\left.\begin{array}{l} A \geq C \\ B \geq D \end{array}\right\} \Rightarrow \min(A,B) \geq \min(C,D)$

$\Rightarrow \min\{h_1[(1-\lambda)x+\lambda y], h_2[\quad]\} \geq \min\left[\{(1-\lambda)h_1(x)+\lambda h_1(y)\}, \{(1-\lambda)h_2(x)+\lambda h_2(y)\}\right]$

supposons que le min $= (1-\lambda)h_1(x)+\lambda h_1(y)$ (le premier membre)

alors $\min \geq \underbrace{(1-\lambda)}_{\geq 0}\min(h_1(x),h_2(x)) + \underbrace{\lambda}_{\geq 0}\min(h_1(y),h_2(y)) = (1-\lambda)h(x)+\lambda h(y)$

$0 \le \lambda \le 1$

<u>dém + courte :</u>

$h_1[(1-\lambda)x+\lambda y] \overset{\text{concavité } h_i}{\geq} (1-\lambda)h_1(x)+\lambda h_1(y) \overset{\text{positivité des } (1-\lambda) \text{ et } \lambda}{\geq} (1-\lambda)\min[h_1(x),h_2(x)] + \lambda\min[h_1(y),h_2(y)]$

$h_2[(1-\lambda)x+\lambda y] \geq (1-\lambda)h_2(x)+\lambda h_2(y) \geq (1-\lambda)\min[h_1(x),h_2(x)] + \lambda\min[h_1(y),h_2(y)]$

$= (1-\lambda)h(x)+\lambda h(y)$

2) a)

$v_1 \longrightarrow z \longrightarrow v_2$

pour aller de x_i à z en passant par v_1, le coût est $d(x_i,v_1)+c(v_1,z(\lambda))$

v_2, —— $d(x_i,v_2)+c(v_2,z(\lambda))$

pour aller de x_i à z, on prend le min des 2 (le chemin le moins coûteux)

$$\Rightarrow d(x_i,z) = \min\{\underbrace{d(x_i,v_1)}_{A}+c(v_1,z(\lambda)),\ \underbrace{d(x_i,v_2)}_{B}+c(v_2,z(\lambda))\}$$

A et B sont des constantes, donc $A+c(v_1,z(\lambda))$ est concave

$B+c(v_2,z(\lambda))$ est concave

on applique alors la propriété générale de la question 1.

b) $f(z) = \sum_{x_i \in N} a_i d(x_i,z)$ $\quad a_i > 0$ et $d(x_i,z)$ concave

une combinaison linéaire positive de fonctions concaves est une fonction concave.

c'est une propriété vue en cours

c) $z = (1-\lambda)v_1 + \lambda v_2$ avec $0 \le \lambda \le 1$

$$\underset{\uparrow\, f \text{ concave}}{f(z) \geq} (1-\lambda)f(v_1)+\lambda f(v_2) \geq \underbrace{(1-\lambda)}_{\geq 0}\min\{f(v_1),f(v_2)\} + \underbrace{\lambda}_{\geq 0}\min\{f(v_1),f(v_2)\} = \min\{f(v_1),f(v_2)\}$$

3) Sous l'hyp. de concavité de $c(v_i,z)$ pour tout v_i de G

on peut refaire la question 2 en plaçant z sur n'importe qu'elle arête (v_i,v_j) de G

et en déduire qu'une solution optimale de (P) est obtenue en plaçant z sur v_i ou sur v_j.

Quasiconvexité

On dit que f, définie sur un convexe C de $\mathbb{R}^n$, est <u>quasi-convexe</u> lorsque :

$$\forall\, x,y \in C\ ,\ 0 \le \lambda \le 1\ ,\ \Rightarrow f[(1-\lambda)x + \lambda y] \le \max\{f(x), f(y)\}$$

si de plus $\quad \forall\, x \neq y \in C\ ,\ 0 < \lambda < 1 \Rightarrow f[(1-\lambda)x + \lambda y] < \max\{f(x), f(y)\}$

on dit que f est <u>fortement</u> <u>quasi-convexe</u>.

Soit x^* min. local strict de f quasi-convexe sur C
alors x^* est un min. global strict de f sur C

$\forall y \in C,\ y \neq x^*$, $0 < \lambda \le 1$ suffisamment petit alors $f(x^*) \underset{\substack{\uparrow \\ x^* \text{ min local} \\ \text{strict}}}{<} f[x^* + \lambda(y - x^*)] = f[(1-\lambda)x^* + \lambda y] \underset{\substack{\uparrow \\ \text{quasi-convexité}}}{\le} \max\{f(x^*), f(y)\}$

si $f(y) \le f(x^*)$ alors on obtient $f(x^*) < f(x^*)$ ce qui est impossible
donc $f(y) > f(x^*)$

Soit x^* min. local de f fortement quasi-convexe sur C
alors x^* est un min. global strict de f sur C

$\forall y \neq x^* \in C$, $0 < \lambda < 1$ suffisamment petit alors $f(x^*) \underset{\substack{\uparrow \\ x^* \text{ min. local}}}{\le} f[x^* + \lambda(y - x^*)] = f[(1-\lambda)x^* + \lambda y] \underset{\substack{\uparrow \\ \text{forte} \\ \text{quasi-convexité}}}{<} \max\{f(x^*), f(y)\}$

On conclut alors de la même manière que précédemment.

Quasiconvexité.

Soit $g : C \to \mathbb{R}$ $\quad C$ convexe de $\mathbb{R}^n$, $\quad g$ convexe et $g(x) \geq 0$ sur C

$\quad h : C \to \mathbb{R}$ $\quad\quad h$ concave et $h(x) > 0$ sur C

m.q. $f(x) = g(x)/h(x)$ est quasi-convexe sur C

$0 \leq \vartheta \leq 1$, $x, y \in C$

$$f[(1-\vartheta)x + \vartheta y] = \frac{g[(1-\vartheta)x+\vartheta y]}{h[(1-\vartheta)x+\vartheta y]} \underset{\substack{\text{dénominateur} > 0 \\ \text{convexité de } g}}{\leq} \frac{(1-\vartheta)g(x)+\vartheta g(y)}{h[(1-\vartheta)x+\vartheta y]} \underset{\text{numérateur} \geq 0}{\leq} \frac{(1-\vartheta)g(x)+\vartheta g(y)}{(1-\vartheta)h(x)+\vartheta h(y)}$$

$$h[(1-\vartheta)x+\vartheta y] \underset{h \text{ concave}}{\geq} (1-\vartheta)h(x)+\vartheta h(y) \Rightarrow \frac{1}{h[(1-\vartheta)x+\vartheta y]} \underset{\substack{h>0 \\ \text{sur } C}}{\leq} \frac{1}{(1-\vartheta)h(x)+\vartheta h(y)}$$

étudions $k(\vartheta) = \dfrac{(1-\vartheta)a+\vartheta b}{(1-\vartheta)c+\vartheta d}$ sur $0 \leq \vartheta \leq 1$

$$k'(\vartheta) = \frac{(-a+b)[(1-\vartheta)c+\vartheta d] - [(1-\vartheta)a+\vartheta b](-c+d)}{D^2} = \frac{-\cancel{ac} + \cancel{ac\vartheta} - \cancel{ad\vartheta} + bc - \cancel{bc\vartheta} + \cancel{bd\vartheta} - (-\cancel{ac} + \cancel{ac\vartheta} - \cancel{bc\vartheta} + ad - \cancel{ad\vartheta} + \cancel{bd\vartheta})}{D^2}$$

$$= \frac{bc - ad}{D^2}$$

le signe de $k'(\vartheta)$ ne dépend pas de ϑ

donc soit k est croissante, soit k est décroissante

$k(0) = \frac{a}{c}$ et $k(1) = \frac{b}{d}$

si k est croissante $k(1) \geq k(\vartheta)$ pour $0 \leq \vartheta \leq 1$

si k est décroissante $k(0) \geq k(\vartheta)$

dans les 2 cas $k(\vartheta) \leq$ ~~max~~ $\max\{k(0), k(1)\}$

Finalement $f[(1-\vartheta)x+\vartheta y] \leq \max\left\{\dfrac{g(x)}{h(x)}, \dfrac{g(y)}{h(y)}\right\} = \max\{f(x), f(y)\}$

Soit $g : C \to \mathbb{R}$ $\quad C$ convexe de $\mathbb{R}^n$, $\quad g$ convexe et $g(x) \leq 0$ sur C

$\quad h : C \to \mathbb{R}$ $\quad\quad h$ convexe et $h(x) > 0$ sur C

m.q. $f(x) = \dfrac{g(x)}{h(x)}$ est quasi-convexe sur C

$0 \leq \vartheta \leq 1$, $x, y \in C$

$$f[(1-\vartheta)x+\vartheta y] = \frac{g[(1-\vartheta)x+\vartheta y]}{h[(1-\vartheta)x+\vartheta y]} \underset{\substack{g \text{ convexe} \\ \text{dénominateur} > 0}}{\leq} \frac{(1-\vartheta)g(x)+\vartheta g(y)}{h[(1-\vartheta)x+\vartheta y]} \underset{\text{numérateur} \leq 0}{\leq} \frac{(1-\vartheta)g(x)+\vartheta g(y)}{(1-\vartheta)h(x)+\vartheta h(y)}$$

$$h[(1-\vartheta)x+\vartheta y] \underset{h \text{ convexe.}}{\leq} (1-\vartheta)h(x)+\vartheta h(y) \Rightarrow \frac{1}{h[(1-\vartheta)x+\vartheta y]} \underset{\substack{h>0 \\ \text{sur } C}}{\geq} \frac{1}{(1-\vartheta)h(x)+\vartheta h(y)} \Rightarrow \frac{a}{h[(1-\vartheta)x+\vartheta y]} \underset{\text{lorsque } a \leq 0}{\leq} \frac{a}{(1-\vartheta)h(x)+\vartheta h(y)}$$

sachant, cf ex. précédent, que $\dfrac{(1-\vartheta)a+\vartheta b}{(1-\vartheta)c+\vartheta d} \leq \max\left\{\dfrac{a}{c}, \dfrac{b}{d}\right\}$

il résulte que $f[(1-\vartheta)x+\vartheta y] \leq \max\left\{\dfrac{g(x)}{h(x)}, \dfrac{g(y)}{h(y)}\right\} = \max\{f(x), f(y)\}$

Quasiconvexité.

Rappel exo précédent : soit g concave sur S convexe de $\mathbb{R}^n$ et $g(x) > 0 \quad \forall x \in S$

alors $f(x) = \dfrac{1}{g(x)}$ est convexe sur S

Soit $g : C \to \mathbb{R}$ $\quad$ C convexe de $\mathbb{R}^n$ et g convexe et $g(x) \leq 0$ sur C

$\quad h : C \to \mathbb{R}$ $\quad$ h concave et $h(x) > 0$ sur C

m.q. $f(x) = g(x)\,h(x)$ est quasiconvexe sur C

$h(x) > 0 \Rightarrow g(x) = \dfrac{f(x)}{h(x)} \qquad \forall x \in C$

$0 \leq \vartheta \leq 1$, $x, y \in C$

$$g[(1-\vartheta)x + \vartheta y] \underset{g \text{ convexe}}{\leq} (1-\vartheta) g(x) + \vartheta g(y) = (1-\vartheta)\frac{f(x)}{h(x)} + \vartheta \frac{f(y)}{h(y)} \leq \left[\frac{1-\vartheta}{h(x)} + \frac{\vartheta}{h(y)}\right] \max\{f(x), f(y)\}$$

$$g[(1-\vartheta)x + \vartheta y] = \frac{f[(1-\vartheta)x + \vartheta y]}{h[(1-\vartheta)x + \vartheta y]}$$

et $\dfrac{1-\vartheta}{h(x)} \geq 0$, $\dfrac{\vartheta}{h(y)} \geq 0$

par ailleurs, cf. rappel, $k(x) = \dfrac{1}{h(x)}$ est convexe

$$\Rightarrow k[(1-\vartheta)x + \vartheta y] \leq (1-\vartheta) k(x) + \vartheta k(y) \Leftrightarrow \frac{1}{h[(1-\vartheta)x + \vartheta y]} \leq \frac{(1-\vartheta)}{h(x)} + \frac{\vartheta}{h(y)} \Rightarrow \frac{a}{h[(1-\vartheta)x + \vartheta y]} \geq a\left[\frac{1-\vartheta}{h(x)} + \frac{\vartheta}{h(y)}\right]$$ avec $a \leq 0$

finalement

$$\frac{f[(1-\vartheta)x + \vartheta y]}{h[(1-\vartheta)x + \vartheta y]} \underset{\text{car } \max\{f(x), f(y)\} \leq 0}{\leq} \frac{\max\{f(x), f(y)\}}{h[(1-\vartheta)x + \vartheta y]} \Rightarrow f\{(1-\vartheta)x + \vartheta y\} \leq \max\{f(x), f(y)\}$$

maintenant il ne reste plus qu'à simplifier en enlevant les dénominateurs des 2 membres

Fonctions convexes

Soit $f : \mathbb{R}^m \to \mathbb{R}$ définie par $f(u) = \max\{f_1(u), \dots, f_p(u)\}$ où $f_1, \dots, f_p$ sont des fonctions convexes définies sur $\mathbb{R}^m$

1) Montrer que f est convexe

2) Application : soit $f : \mathbb{R} \to \mathbb{R}$ définie par $f(u) = \max\limits_{(x_1,x_2)\in S} \{3x_1 + 2x_2 + u(2x_1 + 3x_2 - 4)\}$ où $S = \{0,1\}^2$

Montrer, en appliquant le résultat précédent, que $f(u)$ est convexe.

1) $0 \le \lambda \le 1$, $u, v \in \mathbb{R}^m$ $\quad f[(1-\lambda)u + \lambda v] = \max\limits_i \{f_i[(1-\lambda)u + \lambda v]\} = f_{i^*}[(1-\lambda)u + \lambda v]$ où i^* réalise le max

f_{i^*} étant convexe, on a $f_{i^*}[(1-\lambda)u + \lambda v] \le (1-\lambda) f_{i^*}(u) + \lambda f_{i^*}(v)$.

$f_{i^*}(u) \le \max\limits_i \{f_i(u)\}$ et $f_{i^*}(v) \le \max\limits_i \{f_i(v)\} \Rightarrow$

$(1-\lambda) f_{i^*}(u) + \lambda f_{i^*}(v) \le (1-\lambda) \max\limits_i f_i(u) + \lambda \max\limits_i f_i(v) = (1-\lambda) f(u) + \lambda f(v)$

2) $u \in \mathbb{R}$, $f(u) = \max\limits_{(x_1,x_2)\in S = \{0,1\}^2} \{3x_1 + 2x_2 + u(2x_1 + 3x_2 - 4)\} = \max\{\underbrace{-4u}_{f_1(u)}, \underbrace{3-2u}_{f_2(u)}, \underbrace{2-u}_{f_3(u)}, \underbrace{5+u}_{f_4(u)}\}$

les 4 fonctions sont linéaires donc convexes et on applique 1)

Soit $f_\theta(x_1,x_2,x_3) = 2x_1^2 + 3x_2^2 + \theta x_3^2 + 3x_1x_2 + 4x_1x_3 + 3x_2x_3$ définie sur $\mathbb{R}^3$

1) Pour quelles valeurs de θ la fonction f_θ est-elle (strict. convexe ?) convexe ?

Soit $g(x_1,x_2,x_3) = 2x_1^2 + 3x_2^2 + 3x_3^2 + 3x_1x_2 + 4x_1x_3 + 3x_2x_3 - 5x_1 - 2x_2 - x_3$ définie sur $\mathbb{R}^3$

2) (Montrer que) $x_1 = 18/5$, $x_2 = -7/15$, $x_3 = -2$ est minimum global de g sur $\mathbb{R}^3$? (bien justifier la réponse)

1) On calcule le hessien de f_θ $\quad \nabla f = \begin{bmatrix} 4x_1 + 3x_2 + 4x_3 \\ 6x_2 + 3x_1 + 3x_3 \\ 2\theta x_3 + 4x_1 + 3x_2 \end{bmatrix}$ $\quad Hf = \begin{bmatrix} 4 & 3 & 4 \\ 3 & 6 & 3 \\ 4 & 3 & 2\theta \end{bmatrix}$

Hf est déf. positif ssi $\Delta_1, \Delta_2, \Delta_3 > 0$ $\quad \Delta_1 = 4$, $\Delta_2 = 24 - 9 > 0$

$\Delta_3 = 4 \times (12\theta - 9) - 3(6\theta - 12) + 4(9 - 24) = \theta(4\times 12 - 3\times 6) - \cancel{4}\times 9 + 3\times 12 + \cancel{4}\times 9 - 4\times 24 = 6\theta(4\times 2 - 3) + 12(3 - 4\times 2)$

$= 6\theta\times 5 + 12(-5) = 5\times 6(\theta - 2) > 0$ pour $\theta > 2$

donc si $\theta > 2$ Hf est déf positif $\Rightarrow f_\theta$ est strict. convexe.

si $\theta = 2$ Hf est semi-défini positif $\Rightarrow f_\theta$ est convexe. car Hf est semi-définie positive

si $\theta < 2$ Δ_3 est < 0 or $\Delta_3 = \lambda_1 \times \lambda_2 \times \lambda_3$ les 3 valeurs propres donc l'une au moins est < 0 donc Hf n'est pas semi-définie positive.

donc f_θ n'est pas convexe.

2) $g = f_\theta(x_1, x_2, x_3) + (-5x_1 - 2x_2 - x_3)$ avec $\theta = 3$ donc f_θ est strict. convexe, par ailleurs $-5x_1 - 2x_2 - x_3$ est linéaire donc convexe

donc g est strict. convexe alors x^* est minimum global de g ssi x^* est un point critique de g

$\nabla g = \nabla f_{\theta=3} + \begin{bmatrix} -5 \\ -2 \\ -1 \end{bmatrix} = \begin{cases} 4x_1 + 3x_2 + 4x_3 - 5 = 0 \\ 6x_2 + 3x_1 + 3x_3 - 2 = 0 \\ 6x_3 + 4x_1 + 3x_2 - 1 = 0 \end{cases}$ $\quad$ le point donné vérifie bien ces équations

Soit $g : \mathbb{R}^m \to \mathbb{R}$ fonction convexe

$h : \mathbb{R}^n \to \mathbb{R}^m$ définie par $h(x) = Ax + b$ avec A matrice $m \times n$, b vecteur $m \times 1$

1) Montrer que $f = g \circ h$ est convexe.

2) Application : soit $f(x_1, x_2) = 2(2x_1 + 3x_2 + 1)^2 + (4x_1 + x_2 + 3)^2 + 2(2x_1 + 3x_2 + 1)(4x_1 + x_2 + 3)$ définie sur $\mathbb{R}^2$

montrer que f est convexe. On identifiera A et b, g et on appliquera le résultat précédent

(On ne calculera pas H_f)

1) $0 \leq \theta \leq 1 \quad x, y \in \mathbb{R}^n \quad h[(1-\theta)x + \theta y] = A[(1-\theta)x + \theta y] + b = (1-\theta)Ax + \theta Ay + (1-\theta)b + \theta b = (1-\theta)h(x) + \theta h(y)$

donc $g[h[(1-\theta)x + \theta y]] = g[(1-\theta)h(x) + \theta h(y)] \underset{\substack{\uparrow \\ g \text{ convexe}}}{\leq} (1-\theta)g(h(x)) + \theta g(h(y))$

2) on pose $A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$ $b = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ alors $h(x_1, x_2) = Ax + b = \begin{bmatrix} 2x_1 + 3x_2 + 1 \\ 4x_1 + x_2 + 3 \end{bmatrix}$

on pose $g(y_1, y_2) = y \cdot Qy$ avec $Q = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ alors $y \cdot Qy = 2y_1^2 + y_2^2 + 2y_1y_2$

et $H_g = 2Q$ matrice déf. positive. $\left[\Delta_1 = 2 > 0, \ \Delta_2 = 2 - 1 = 1 > 0\right]$

donc g est convexe. et $f(x_1, x_2) = g \circ h(x_1, x_2)$

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x_1, x_2) = e^{x_1^2 + x_2^2} - 5x_1 + 10x_2$

On montrera que f est convexe et même strictement convexe de 2 façons différentes.

1) En calculant son hessien

2) Sans calculer son hessien.

1) $\nabla f = \begin{bmatrix} 2x_1 e^{x_1^2 + x_2^2} - 5 \\ 2x_2 e^{x_1^2 + x_2^2} - 10 \end{bmatrix}$ $\quad H_f = \begin{bmatrix} 2e^A + 2x_1 \times 2x_1 e^A & 2x_1 \times 2x_2 e^A \\ 2x_2 \times 2x_1 e^A & 2e^A + 2x_2 \times 2x_2 e^A \end{bmatrix}$ avec $A = x_1^2 + x_2^2$

$= e^A \begin{bmatrix} 2 + 4x_1^2 & 4x_1x_2 \\ 4x_1x_2 & 2 + 4x_2^2 \end{bmatrix}$ $\quad \Delta_1 = 2 + 4x_1^2 > 0$

$\Delta_2 = (2 + 4x_1^2)(2 + 4x_2^2) - 16x_1^2x_2^2$

$= 4 + 8x_1^2 + 8x_2^2 > 0$

H_f déf. positif $\forall x_1, x_2 \Rightarrow f$ est strict^t convexe.

2) $f = f_1 + f_2$ avec $f_1 = e^{x_1^2 + x_2^2}$ $\quad f_2 = -5x_1 + 10x_2$

$f_1 = g \circ h$ avec $g(y) = e^y$ et $h(x_1, x_2) = x_1^2 + x_2^2$ $\quad h$ est strict^t convexe car $H_h = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ déf. positif

g est convexe et strict^t croissante

donc f_1 est strict^t convexe.

f_2 est linéaire donc convexe.

donc f est strict^t convexe.