Dualité Lagrangienne

Ici, nous nous intéressons au problème d'optimisation suivant :

$$\min_{x \in X} f(x)$$

s.c:

$$g_i(x) \leq 0, \quad i=1, \ldots, q$$
$$h_j(x)=0, \quad j= 1, \ldots, p$$

$g_i(x)$ et $h_j(x)$ sont supposées à valeurs réelles et de classe C^1

1. Problème primal, problème dual

Problème primal (P): minimiser $f(x)$ sous les contraintes $g(x) \leq 0$, $h(x) = 0$, $x \in X \subseteq R^n$

où $h = \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix}$, $g = \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix}$

X est une partie quelconque de R^n. Par exemple: R^n, un ensemble discret (nombres entiers naturels par exemple), ou encore un ensemble défini par des contraintes d'inégalité ou d'égalité autres que g et h.

A partir de (P) on définit un autre problème: le problème dual.

Problème dual (D): maximiser $\theta(u,v)=\inf\{f(x)+u\bullet g(x)+v\bullet h(x): x\in X\}$ sous les contraintes $u\geq 0$

$L(x,u,v)=f(x)+u\bullet g(x)+v\bullet h(x)$ est appelée fonction de Lagrange

$\theta(u,v)=\inf\limits_{x\in X} L(x,u,v)$ est appelée fonction duale

2. Relation entre primal et dual

Théorème de dualité (faible)
Si x est une solution réalisable de (P) et si (u,v) est une solution réalisable de (D) alors $f(x)\geq \theta(u,v)$.

Corollaire
Si x est une solution réalisable de (P) et si (u,v) est une solution réalisable de (D) telles que $f(x)=\theta(u,v)$ alors x est une solution (optimale) de (P) et (u,v) est une solution (optimale) de (D)

Remarque: la fonction $\theta(u, v)$ est une fonction concave.

Dualité en programmation linéaire

En programmation linéaire le primal et le dual sont reliés selon le tableau suivant:

primal	dual
min c : coût b : terme de droite	max b : coût c : terme de droite
contrainte i du type $\geqslant$ contrainte i du type $=$ contrainte i du type $\leqslant$	variable i du type $\geqslant 0$ variable i du type *libre* variable i du type $\leqslant 0$
variable i du type $\geqslant$ variable i du type *libre* variable i du type $\leqslant$	contrainte i du type $\leqslant 0$ contrainte i du type $=$ contrainte i du type $\geqslant 0$

3. Points selles

3.1. définition

Soient $x^* \in X, u^* \geq 0$. $\left(x^*, u^*, v^*\right)$ est un point selle ssi:

$$L(x^*, u, v) \leq L(x^*, u^*, v^*) \leq L(x, u^*, v^*) \quad \forall x \in X,\ \forall u \geq 0$$

3.2. Caractérisation des points selles

Soient $x^* \in X, u^* \geq 0$. $\left(x^*, u^*, v^*\right)$ est un point selle ssi
$$\begin{cases} 1)\ L(x^*, u^*, v^*) = \inf\limits_{x \in X} L(x, u^*, v^*) \\ 2)\ g(x^*) \leq 0,\ h(x^*) = 0 \\ 3)\ u_i^* g_i(x^*) = 0 \quad i = 1, \ldots, m \end{cases}$$

3.3. théorème de dualité (fort)

Si $\left(x^*, u^*, v^*\right)$ est un point selle alors x^* est une solution (optimale) de (P) et (u^*, v^*) est une solution (optimale) de (D)

En effet x^* est une solution réalisable de (P) en raison de la caractéristique 2). De plus les caractéristiques 2), 3) et 1) permettent d'établir les égalités suivantes:

$$f(x^*) = f(x^*) + u^* \bullet g(x^*) + v^* \bullet h(x^*) = L(x^*, u^*, v^*) = \inf_{x \in X} L(x, u^*, v^*) = \theta\left(u^*, v^*\right)$$

La conclusion découle du corollaire du théorème de la dualité faible.

Max (D)=θ(u*, ν*) = L(x*, u*,ν*) = f(x*)=Min (P)

• Pour les problèmes de type (p) qui n'admettent pas de point-col (c'est le cas général, pour les programmes discrets où S est un ensemble discret (programmation linéaire en nombres entiers)) on a:

$$\theta(u^*, v^*) < f(x^*)$$

Et la différence f (x*)- θ(u*, ν*) est appelé le saut de Dualité

Exemple:

$$\begin{cases} Min\,(x^2 + y^2) \\ s.c: \quad 2x + y \leq -4 \end{cases}$$

- Le lagrangien de f est défini par : L(x, y, λ)=$x^2 + y^2$+ λ($2x + y$+4).
- - le minimum en x de L(x, y, λ) est défini par:

$$\begin{cases} \dfrac{\partial L}{\partial x} = 2x + 2\lambda = 0 \Rightarrow x = -\lambda \\ \dfrac{\partial L}{\partial y} = 2y + \lambda = 0 \Rightarrow y = -\lambda/2 \end{cases}$$

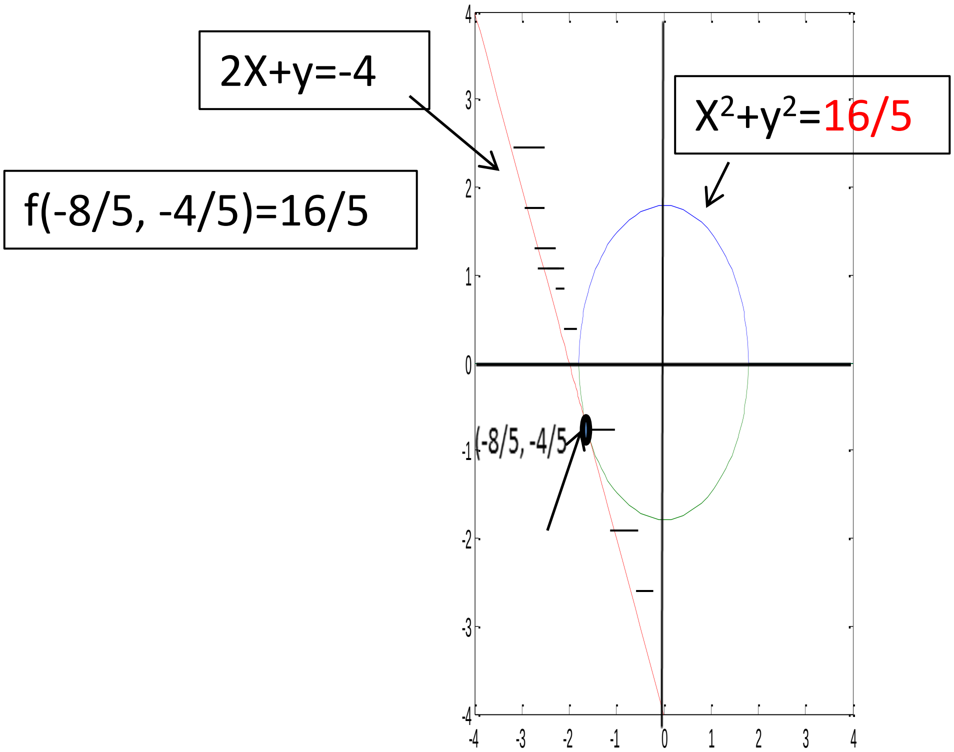
D'où la fonction duale:

$\theta(\lambda) = -\frac{5\lambda^2}{4} + 4\lambda$

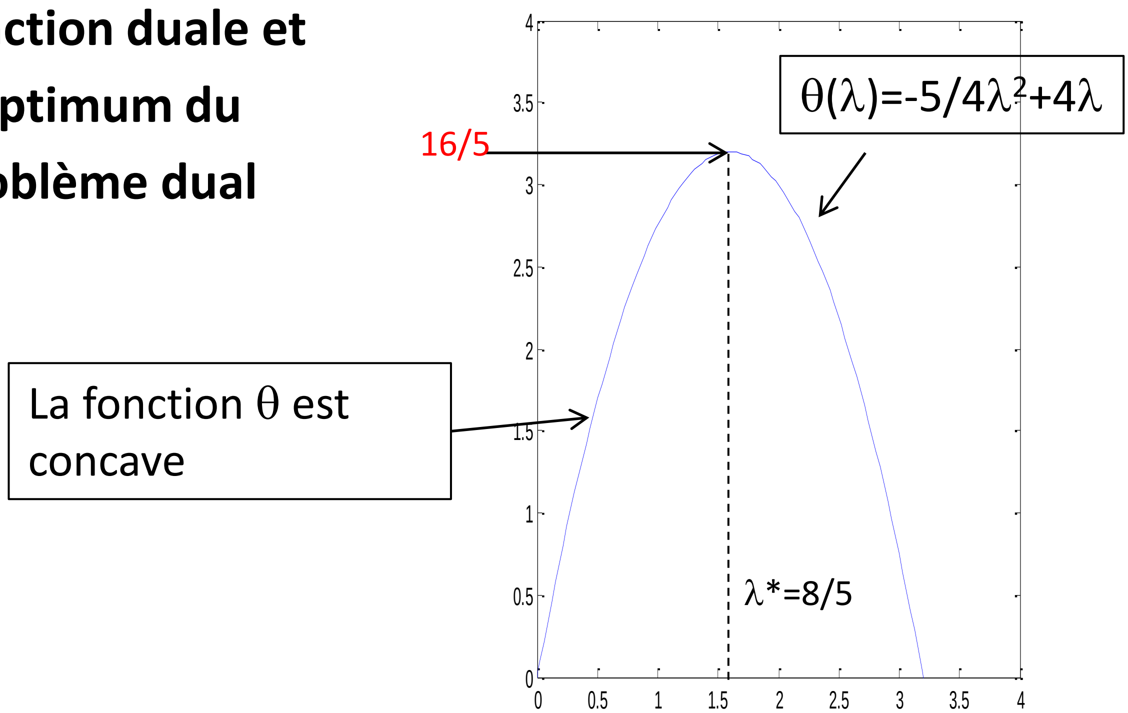
Le maximum de θ(λ) est obtenu lorsque:

$\frac{\partial\theta}{\partial\lambda} = -\frac{5\lambda}{2} + 4 = 0$ d'où $\lambda^* = 8/5$. On a alors: f(x*, y*)=θ(λ *)=16/5 et X*=-8/5 et y*=-4/5

Solution optimale du problème primal:



- **Fonction duale et l'optimum du problème dual**



Exemple 2:

soit (P) minimiser $f(x_1,x_2)=x_1^2+x_2^2$ sous les contraintes $\begin{cases} x_1+x_2\geq 1 \\ x_1\geq 0,\ x_2\geq 0 \end{cases}$

On met les contraintes sous la forme $\begin{cases} g_1(x_1,x_2)=-x_1-x_2+1\leq 0 \\ g_2(x_1,x_2)=-x_1\leq 0 \\ g_3(x_1,x_2)=-x_2\leq 0 \end{cases}$

Ici $X=R^n$.

La fonction de Lagrange est $L(x,u)=x_1^2+x_2^2+u_1(-x_1-x_2+1)+u_2(-x_1)+u_3(-x_2)$

La fonction duale est $\theta(u)=\inf\left\{L(x,u): x\in R^n\right\}$

Explicitons la fonction duale. Pour cela cherchons un point critique de $L(x,u)$ (u constant):

$$\begin{cases} 2x_1-u_1-u_2=0 \\ 2x_2-u_1-u_3=0 \end{cases} \Rightarrow \begin{cases} x_1=\dfrac{u_1+u_2}{2} \\ x_2=\dfrac{u_1+u_3}{2} \end{cases}$$

On en déduit: $\theta(u)=-\dfrac{u_1^2}{2}-\dfrac{u_2^2}{4}-\dfrac{u_3^2}{4}-\dfrac{u_1u_2}{2}-\dfrac{u_1u_3}{2}+u_1$

On remarque que $\theta(u)$ est bien concave (on vérifie que son hessien est défini négatif).

Le problème dual (D) est: maximiser $\theta(u)$ sous les contraintes $u_1\geq 0,\ u_2\geq 0,\ u_3\geq 0$

Résolvons le problème dual. Pour cela réécrivons (D) en un problème (D') de type minimiser:

minimiser $\frac{u_1^2}{2}+\frac{u_2^2}{4}+\frac{u_3^2}{4}+\frac{u_1u_2}{2}+\frac{u_1u_3}{2}-u_1$ sous les contraintes $-u_1 \le 0,\ -u_2 \le 0,\ -u_3 \le 0$

Ecrivons les conditions de Kuhn et Tucker:

$$\begin{cases} u_1+\frac{u_2}{2}+\frac{u_3}{2}-1-\lambda_1=0 \\ \frac{u_2}{2}+\frac{u_1}{2}-\lambda_2=0 \\ \frac{u_3}{2}+\frac{u_1}{2}-\lambda_3=0 \\ \\ \lambda_i \ge 0 \quad (i=1,2,3) \\ \lambda_i u_i=0 \quad (i=1,2,3) \end{cases}$$

Ce système admet une solution telle que (u_1,u_2,u_3) vérifie les contraintes de (D):

$\lambda_1=0,\ \lambda_2=\lambda_3=\frac{1}{2},\ u_1=1,\ u_2=u_3=0.$

Les hypothèses de qualification sont satisfaites puisque les conditions de l'indépendance linéaire sont satisfaites.

Le problème (D') vérifiant par ailleurs les conditions de convexité, les conditions de Kuhn-Tucker sont suffisantes et le point $(1,0,0)$ est minimum (global) donc maximum (global) pour (D).

$\inf\left\{L(x,(1,0,0)):\ x\in R^n\right\}$ est atteint par le point $(x_1,x_2)=\left(\frac{1}{2},\frac{1}{2}\right)$. Il est unique donc c'est la solution de (P) (les hypothèses de convexité et de superconsistence étant vérifiées on sait que la fonction de Lagrange admet un point selle).

Remarquons que $\theta(1,0,0)=\frac{1}{2}=f\left(\frac{1}{2},\frac{1}{2}\right)$ et que $x_1=x_2=\frac{1}{2},\ u_1=1,\ u_2=u_3=0$ vérifient les conditions de Kuhn-Tucker de (P).

Méthodes de résolution: Trouver une solution d'un problème d'optimisation sous contraintes fonctionnelles consiste à déterminer un point optimal x* et des multiplicateurs de Lagrange associés (λ*, μ*). Deux grandes familles de méthodes peuvent être définies pour la résolution de ce type de problèmes: les méthodes primales et les méthodes duales. Les méthodes primales ce concentrent sur la détermination du point x*, les multiplicateurs (λ, μ) ne servent souvent qu'à vérifier l'optimalité de x*. Les méthodes duales s'intéressent à la détermination des multiplicateurs (λ*, μ*) en travaillant sur le problème dual (D).

6. Résolution du problème dual

1. Algorithme de sous-gradient

1.1. Sous-gradient de la fonction duale

Notons: $w = \begin{pmatrix} u \\ v \end{pmatrix}, \beta(x) = \begin{pmatrix} g(x) \\ h(x) \end{pmatrix}, \; X(w) = \{x \in X : f(x) + w \bullet \beta(x) = \theta(w)\}.$

- $\beta(x) = \nabla\theta(w)$ est le sous gradient de la fonction duale $\theta(w)$

Théorème: si $\theta(w)$ est différentiable en w solution (optimale) du problème dual et si $x \in X(w) \neq \emptyset$ alors (x, w) est un point selle.

1.2. Maximisation de la fonction duale

On cherche ici à résoudre le problème dual c'est-à-dire maximiser la fonction duale sous la contrainte que certaines des variables doivent être positives ou nulles. La simplicité de ces contraintes fait que l'on peut adapter l'algorithme de plus forte pente utilisé dans le cas de l'optimisation sans contraintes. La différence est qu'on utilise un sous-gradient à la place du gradient et que l'on interdit les déplacements qui rendraient négatives des variables astreintes à être positives ou nulles.

Algorithme de sous-gradient (algorithme d'Uzawa)

On se donne une suite de nombres réels $\rho^{(k)} > 0$.

Soit $w^{(0)} = \begin{pmatrix} u^{(0)} \\ v^{(0)} \end{pmatrix}$ avec $u^{(0)} \geq 0$, k=0.

1) résoudre le problème $\min\limits_{x \in X} f(x) + w^{(k)} \bullet \beta(x)$: soit $x^{(k)}$ une solution.

$$w^{k+1} = w^{k} + \rho^{k} \beta\left(x^{k}\right)$$

Avec

$$u_i^{k+1} = \begin{cases} u_i^{k+1} & \text{si } u_i^{k+1} \geq 0 \\ 0 & \text{sinon} \end{cases}$$

si $w^{(k+1)} = w^{(k)}$ alors STOP

sinon k=k+1 et aller en 1)

Les choix possibles de $\rho^{(k)}$ sont:

- constant $\rho^{(k)} = \rho \ \forall k$

- la série divergente $\rho^{(k)} \to 0, \sum_k \rho^{(k)} = +\infty$

Théorème

Dans le cas de la minimisation d'une fonction quadratique $f(x) = \frac{1}{2} x \bullet Ax - b \bullet x$ (A définie positive) sous des contraintes d'inégalité affines $Dx - d \leq 0$ (D de rang m) et $X = R^n$, dans le cas du choix $\rho^{(k)} = \rho \ \forall k$, si $\rho \in \left]0, \frac{\lambda_{\min}(A)}{\|D\|^2}\right]$ ($\lambda_{\min}(A)$ est la plus petite valeur propre de A) alors la suite $\left(x^{(k)}, w^{(k)}\right)$ converge vers $\left(x^*, w^*\right)$ un point selle.

- Le pas de déplacement peut être choisi aussi comme suit:

$$\rho^k = \delta^k \frac{\left(z - \theta(w^k)\right)}{\left\|\beta(x^k)\right\|^2}$$

Où z est la valeur de l'objectif de la meilleur solution réalisable obtenu pour le primal (P), δ_k est un nombre choisi entre 0 et 2, puis δ_k est réduit par un facteur de 2 chaque fois que $\theta(w^k)$ n'est pas améliorée après un certain nombre d'itérations. $\theta(w^k)$ est la fonction duale.

- **Projection:** Soient $C \subseteq R^n$, $x \in R^n$, $x^* \in C$ tel que:

$\| x-x^* \| = \inf \{ \| x-y \| : y \in C\}$, **x*est appelé projection de x sur C**.

X*, s'il existe, est le point de C le plus proche de x.

Exemple: appliquons l'algorithme d'Uzawa au dual du problème minimiser $f(x_1,x_2)=x_1^2+x_2^2$

sous les contraintes $\begin{cases} x_1+x_2\geq 1 \\ x_1\geq 0,\ x_2\geq 0 \end{cases}$ en partant de $u^{(0)}=\begin{pmatrix}0\\1\\1\end{pmatrix}$ avec le pas constant $\rho=1$.

On pose: $g(x_1,x_2)=\begin{pmatrix} g_1(x_1,x_2)\\ g_2(x_1,x_2)\\ g_3(x_1,x_2)\end{pmatrix}=\begin{pmatrix}-x_1-x_2+1\\ -x_1\\ -x_2\end{pmatrix}$.

itération 1

On minimise $x_1^2+x_2^2-x_1-x_2$ $(x_1,x_2)\in R^2$. La fonction étant convexe et soumise à aucune contrainte, il suffit de chercher un point critique: $x_1=x_2=\frac{1}{2}$. Ce qui donne $\theta(0,1,1)=-\frac{1}{2}$

$$u^{(1)}=\begin{pmatrix}0\\1\\1\end{pmatrix}+\begin{pmatrix}0\\-\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}=\begin{pmatrix}0\\\frac{1}{2}\\\frac{1}{2}\end{pmatrix}$$

itération 2

On minimise $x_1^2+x_2^2-\frac{1}{2}x_1-\frac{1}{2}x_2$ $(x_1,x_2)\in R^2$. La solution est $x_1=x_2=\frac{1}{4}$. Ce qui donne $\theta\left(0,\frac{1}{2},\frac{1}{2}\right)=-\frac{1}{8}$

$$u^{(2)}=\begin{pmatrix}0\\\frac{1}{2}\\\frac{1}{2}\end{pmatrix}+\begin{pmatrix}\frac{1}{2}\\-\frac{1}{4}\\-\frac{1}{4}\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\\\frac{1}{4}\\\frac{1}{4}\end{pmatrix}$$

itération 3

On minimise $x_1^2 + x_2^2 - \frac{3}{4}x_1 - \frac{3}{4}x_2 + \frac{1}{2}\ \ (x_1, x_2) \in R^2$. La solution est $x_1 = x_2 = \frac{3}{8}$. Ce qui donne

$\theta\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right) = \frac{7}{32}$

$u^{(2)} + g\left(\frac{3}{8}, \frac{3}{8}\right) = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix} + \begin{pmatrix} \frac{1}{4} \\ -\frac{3}{8} \\ -\frac{3}{8} \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ -\frac{1}{8} \\ -\frac{1}{8} \end{pmatrix}$. En projetant sur $K = \left(R^+\right)^3$ on obtient: $u^{(3)} = \begin{pmatrix} \frac{3}{4} \\ 0 \\ 0 \end{pmatrix}$.

itération 4

On minimise $x_1^2 + x_2^2 - \frac{3}{4}x_1 - \frac{3}{4}x_2 + \frac{3}{4}\ \ (x_1, x_2) \in R^2$. La solution est $x_1 = x_2 = \frac{3}{8}$. Ce qui donne

$\theta\left(\frac{3}{4}, 0, 0\right) = \frac{15}{32}$

$u^{(3)} + g\left(\frac{3}{8}, \frac{3}{8}\right) = \begin{pmatrix} \frac{3}{4} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{4} \\ -\frac{3}{8} \\ -\frac{3}{8} \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{3}{8} \\ -\frac{3}{8} \end{pmatrix}$. En projetant on obtient: $u^{(4)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

itération 5

On minimise $x_1^2+x_2^2-x_1-x_2+1 \quad (x_1,x_2)\in R^2$. La solution est $x_1=x_2=\frac{1}{2}$. Ce qui donne $\theta(1,0,0)=\frac{1}{2}$

$u^{(4)}+g\left(\frac{1}{2},\frac{1}{2}\right)=\begin{pmatrix}1\\0\\0\end{pmatrix}+\begin{pmatrix}0\\-\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}=\begin{pmatrix}1\\-\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}$. En projetant on obtient: $u^{(5)}=\begin{pmatrix}1\\0\\0\end{pmatrix}$.

$u^{(5)}=u^{(4)}$ donc on arrête.

Vérifions que $\left(x^*,u^*\right)=\left(\frac{1}{2},\frac{1}{2},1,0,0\right)$ est un point selle:

- les relations de complémentarités 3) sont vérifiées: $u^*\bullet g\left(x^*\right)=\begin{pmatrix}1\\0\\0\end{pmatrix}\bullet\begin{pmatrix}0\\-\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}=0$

- la condition 2) (x^* réalisable) est vérifiée: $g(x^*)=\begin{pmatrix}0\\-\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}\leq 0$

- la condition 1) (x^* minimise la fonction de Lagrange $L(x,u^*)$ en u^*) est vérifiée par définition de x^*.

1.2 Méthode des plans sécants

Le problème dual (D) peut s'écrire:

maximiser z

sous les contraintes $\begin{cases} z \leq f(x) + u \bullet g(x) + v \bullet h(x) \text{ pour tout } x \in X \\ u \geq 0 \end{cases}$

En effet, pour u, v fixés, si z est maximum alors $z = \inf\left\{ f(x) + u \bullet g(x) + v \bullet h(x) : x \in X \right\} = \theta(u,v)$. Donc le problème ci-dessus revient à maximiser la fonction duale $\theta(u,v)$ sous les conditions $u \geq 0$ soit au problème dual.

Ce problème est un problème linéaire (en les variables z, u, v) qui comporte autant de contraintes (en dehors des contraintes $u \geq 0$) qu'il y a d'éléments dans X (en particulier cardinal de X peut être infini).

C'est pourquoi à la place de (D), on résout une relaxation de (D) où seulement quelques contraintes sont prises en compte. Si la solution obtenue vérifie toutes les contraintes c'est fini sinon on rajoute une contrainte non vérifiée et on réitère.

algorithme:

soit $X^{(0)}$ contenant quelques éléments de X, $k \leftarrow 0$.

(1) on résout le problème maître:

maximiser z

sous les contraintes $\begin{cases} z \leq f(x) + u \bullet g(x) + v \bullet h(x) & \text{pour tout } x \in X^{(k)} \\ u \geq 0 \end{cases}$

Soit $z^{(k)}, u^{(k)}, v^{(k)}$ une solution. Pour tester si cette solution vérifie toutes les contraintes de (D), on résout le sous-problème: $z^* = \min\limits_{x \in X} \left\{ f(x) + u^{(k)} \bullet g(x) + v^{(k)} \bullet h(x) \right\}$

Si $z^{(k)} \leq z^*$

alors STOP (on a résolu (D))

sinon

soit $x^{(k)}$ une solution du sous-problème

$X^{(k+1)} \leftarrow X^{(k)} + \left\{ x^{(k)} \right\}$ (on rajoute une contrainte au problème maître)

$k \leftarrow k+1$ et aller en (1)

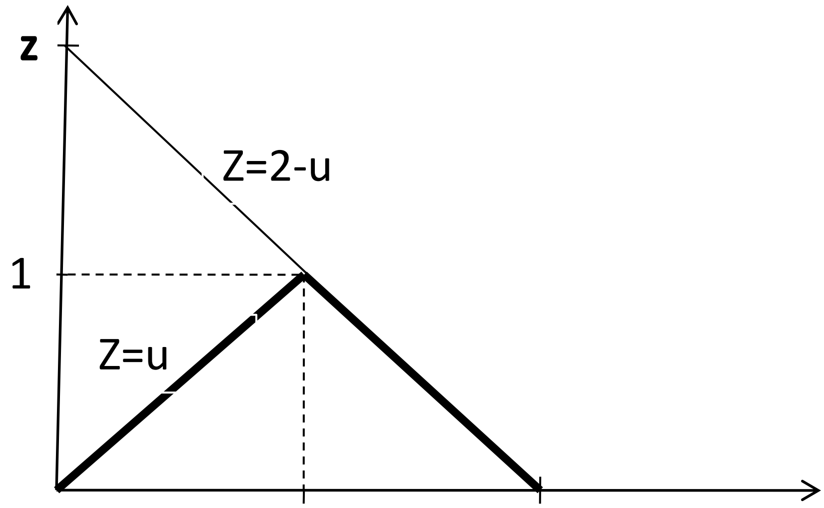
Exemple: appliquons l'algorithme des plans sécants au dual du problème minimiser

$$f(x_1, x_2) = x_1^2 + x_2^2 \text{ sous les contraintes } \begin{cases} g_1(x_1, x_2) = -x_1 - x_2 + 1 \leq 0 \\ (x_1, x_2) \in X = \{(x_1, x_2) : x_1 \geq 0,\ x_2 \geq 0\} \end{cases}$$

Partons avec $X^{(0)} = \{(0,0),(1,1)\}$.

itération k=0:

$$\text{Le problème maître est: maximiser } z \text{ s.c.} \begin{cases} z \leq u_1 \\ z \leq 2 - u_1 \\ u_1 \geq 0 \end{cases}$$



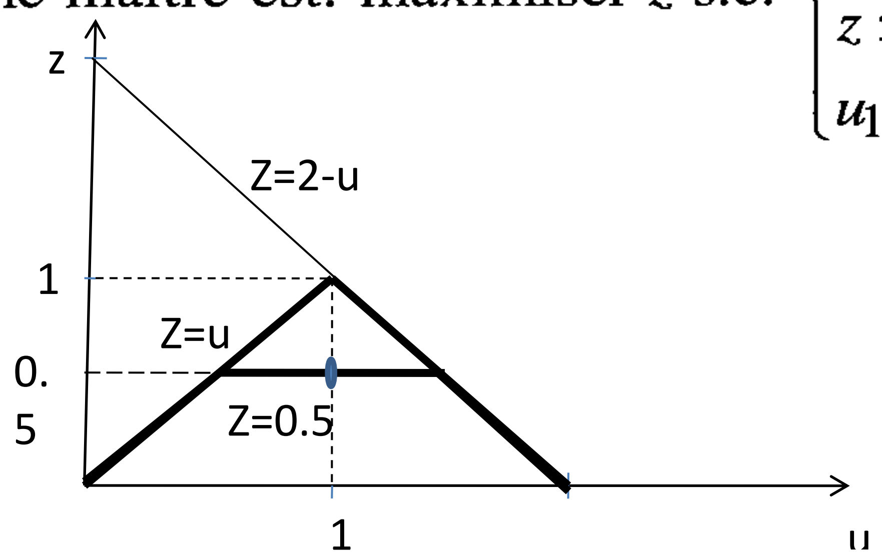
La solution est $z^{(0)}=1, u_1^{(0)}=1$.

Le sous-problème est: $z^*=\min_{x\in X}\left\{x_1^2+x_2^2-x_1-x_2+1\right\}=\frac{1}{2}$

La solution est $x_1^{(0)}=x_2^{(0)}=\frac{1}{2}$. $z^{(0)}>z^*$ donc on rajoute au problème maître, la contrainte engendrée par $x_1^{(0)}=x_2^{(0)}=\frac{1}{2}$ i.e. $X^{(1)}=\left\{(0,0),(1,1),\left(\frac{1}{2},\frac{1}{2}\right)\right\}$.

itération k=1:

Le problème maître est: maximiser z s.c. $\begin{cases} z \leq u_1 \\ z \leq 2-u_1 \\ z \leq \frac{1}{2} \\ u_1 \geq 0 \end{cases}$



$z^{(1)}=\frac{1}{2}, u_1^{(1)}=1$ est une solution.

Méthodes de pénalités et de barrière:

Le principe de ces méthodes est de transformer le problème initial en un problème sans contraintes dont la résolution est a priori plus facile surtout dans le cas de contraintes non linéaires.

méthode de pénalités :

sont des méthodes qui consistent à atteindre la solution du problème par l'extérieur de l'ensemble des solutions réalisables. Dès que l'on atteint une solution réalisable, celle-ci se trouve être une solution optimale. Cette méthode consiste à résoudre une suite de problèmes sans contraintes.

Principe de la méthode:

On considère le problème (P): minimiser $f(x) \quad x \in \mathbf{X} \subset R^n$

où $\mathbf{X} = \{x\colon g_i(x) \leq 0 \; i = 1,\ldots,m \;\; h_i(x) = 0 \; i = 1,\ldots,p\}$

A partir de (P) on construit le problème sans contraintes (P_μ): minimiser $f(x)+\mu P(x) \quad x \in R^n$

où $\mu \in R^+$ et $P\colon R^n \to R$ vérifie $\begin{cases} 1) \; P \text{ continue} \\ 2) \; P(x) \geq 0 \;\; \forall x \in R^n \\ 3) \; P(x) = 0 \Leftrightarrow x \in S \end{cases}$

$P(x)$ est une pénalité infligée à tout x qui n'est pas solution réalisable de (P).

On peut prendre comme pénalité:

$$P(x) = \sum_{i=1}^{m} \left(g_i^+(x)\right)^2 + \sum_{i=1}^{p} \left(h_i(x)\right)^2 \text{ où } g_i^+(x) = \max\{0, g_i(x)\}$$

Intuitivement pour μ "grand", toute solution (optimale) de (P_μ) a une pénalité nulle c'est-à-dire est réalisable pour (P). Mais on ne sait pas a priori quelle valeur assigner à μ. C'est pourquoi, pour chercher cette valeur, on se donne une suite (μ_k) (positif) strictement croissante et qui tend vers l'infini (par exemple k).

La démarche des méthodes de pénalité est alors la suivante:

- pour chaque k on résout le problème sans contraintes (P_k): $f(x) + \mu_k P(x) \quad x \in R^n$
- on obtient (x_k) la suite des solutions des (P_k).

On se donne une suite (μ_k) telle que $\mu_{k+1} > \mu_k > 0$ et qui tend vers l'infini.

On note:

- $q(x,\mu_k) = f(x) + \mu_k P(x)$
- x_k la solution de (P_k): minimiser $q(x,\mu_k) = f(x) + \mu_k P(x) \;\; x \in R^n$

Hypothèse: $\forall k$ (P_k) admet une solution x_k.

Lemme1

$q(x_k,\mu_k) \leq q(x_{k+1},\mu_{k+1})$

$P(x_k) \geq P(x_{k+1})$

$f(x_k) \leq f(x_{k+1})$

Lemme2

Soit x^* une solution de (P). $f(x^*) \geq q(x_k,\mu_k) \geq f(x_k) \;\; \forall k$

Théorème

Si f continue alors toute sous-suite convergente $\left(x_{k_p}\right)$ converge vers une solution de (P).

Le théorème précédent nous affirme la convergence de cette méthode.

Approximation des multiplicateurs de Lagrange:

Soit $S = \{x\colon g_i(x) \leq 0 \; i = 1,\ldots,m \;\; h_i(x) = 0 \; i = 1,\ldots,p\}$

On considère la pénalité:

$$P(x) = \sum_{i=1}^{m}\left(g_i^+(x)\right)^2 + \sum_{i=1}^{p}\left(h_i(x)\right)^2 \;\text{ où }\; g_i^+(x) = \max\{0, g_i(x)\}$$

Lemme: Si g_i est de classe C^1 alors $\left(g_i^+(x)\right)^2$ est de classe C^1 et son gradient est donné par

$\nabla {g_i^+}^2(x) = 2g_i^+(x)\nabla g_i(x)$.

Théorème

Si x^* solution de (P) vérifie le critère de qualification de l'indépendance linéaire

Si f et g, h sont C^1

Si $x_k \xrightarrow{k} x^*$

alors $2\mu_k g_i^+(x_k) \xrightarrow{k} \lambda_i$ et $2\mu_k h_i(x_k) \xrightarrow{k} \alpha_i$ les multiplicateurs de Lagrange associés respectivement aux g_i $(i=1,\ldots,m)$, h_i $(i=1,\ldots,p)$.

Exemple: soit (P) minimiser $f(x_1,x_2)=x_1^2+x_2^2$ sous les contraintes $\begin{cases} x_1+x_2\geq 1 \\ x_1\geq 0,\ x_2\geq 0 \end{cases}$

On met les contraintes sous la forme $\begin{cases} g_1(x_1,x_2)=-x_1-x_2+1\leq 0 \\ g_2(x_1,x_2)=-x_1\leq 0 \\ g_3(x_1,x_2)=-x_2\leq 0 \end{cases}$

Résolvons ce problème à l'aide des pénalités de Courant-Beltrami.

On pose: $\begin{cases} g_1^+(x_1,x_2)=\max\{0,-x_1-x_2+1\} \\ g_2^+(x_1,x_2)=\max\{0,-x_1\} \\ g_3^+(x_1,x_2)=\max\{0,-x_2\} \end{cases}$

La pénalité est: $P(x_1,x_2)=\left(g_1^+(x_1,x_2)\right)^2+\left(g_2^+(x_1,x_2)\right)^2+\left(g_3^+(x_1,x_2)\right)^2$

En prenant $\mu_k=k$, on doit résoudre la suite de problèmes (P_k): minimiser $f(x_1,x_2)+kP(x_1,x_2)$, $(x_1,x_2)\in R^2$.

$f(x_1,x_2)+kP(x_1,x_2)$ est convexe, il suffit de déterminer ses points critiques c'est-à-dire résoudre l'équation:

$$\nabla f(x_1,x_2)+k\nabla P(x_1,x_2)=\nabla f(x_1,x_2)+k\sum_{i=1}^{3}2g_i^+(x_1,x_2)\nabla g_i(x_1,x_2)=0$$

1er cas: $-x_1-x_2+1>0,\ -x_1>0,\ -x_2>0$

$$\begin{cases} 2x_1+k\left(-2(-x_1-x_2+1)-2(-x_1)\right)=0 \\ 2x_2+k\left(-2(-x_1-x_2+1)-2(-x_2)\right)=0 \end{cases}\Rightarrow\begin{cases} x_1+k(2x_1+x_2-1)=0 \\ x_2+k(x_1+2x_2-1)=0 \end{cases}\Rightarrow\begin{cases} (1+2k)x_1+kx_2=k \\ (1+2k)x_2+kx_1=k \end{cases}$$

$\Rightarrow x_1=x_2=\dfrac{k}{(1+3k)}$ >0 donc impossible.

2è cas: $-x_1 - x_2 + 1 > 0,\ -x_1 \le 0,\ -x_2 > 0$

$$\begin{cases} 2x_1 + k\left(-2\left(-x_1 - x_2 + 1\right)\right) = 0 \\ 2x_2 + k\left(-2\left(-x_1 - x_2 + 1\right) - 2\left(-x_2\right)\right) = 0 \end{cases} \Rightarrow \begin{cases} x_1 + k\left(x_1 + x_2 - 1\right) = 0 \\ x_2 + k\left(x_1 + 2x_2 - 1\right) = 0 \end{cases}$$

$$\Rightarrow \begin{cases} (1+k)x_1 + kx_2 = k \\ (1+2k)x_2 + kx_1 = k \end{cases} \Rightarrow \begin{cases} x_1 = \dfrac{k(k+1)}{(k+1)^2 + k} \\ x_2 = \dfrac{k}{(k+1)^2 + k} > 0 \end{cases}$$ donc impossible.

3è cas: $-x_1 - x_2 + 1 > 0,\ -x_1 > 0,\ -x_2 \le 0$

Ce cas est symétrique au cas précédent (permuter x_1 et x_2) donc impossible.

4è cas: $-x_1 - x_2 + 1 > 0,\ -x_1 \le 0,\ -x_2 \le 0$

$$\begin{cases} 2x_1 + k\left(-2\left(-x_1 - x_2 + 1\right)\right) = 0 \\ 2x_2 + k\left(-2\left(-x_1 - x_2 + 1\right)\right) = 0 \end{cases} \Rightarrow \begin{cases} x_1 + k\left(x_1 + x_2 - 1\right) = 0 \\ x_2 + k\left(x_1 + x_2 - 1\right) = 0 \end{cases} \Rightarrow \begin{cases} (1+k)x_1 + kx_2 = k \\ (1+k)x_2 + kx_1 = k \end{cases}$$

$$\Rightarrow x_1 = x_2 = \frac{k}{1+2k}$$

5è cas: $-x_1 - x_2 + 1 \leq 0,\ -x_1 > 0,\ -x_2 > 0$

$$\begin{cases} 2x_1 + k\left(-2(-x_1)\right) = 0 \\ 2x_2 + k\left(-2(-x_2)\right) = 0 \end{cases} \Rightarrow \begin{cases} (1+k)x_1 = 0 \\ (1+k)x_2 = 0 \end{cases} \Rightarrow x_1 = x_2 = 0 \text{ donc impossible.}$$

6è cas: $-x_1 - x_2 + 1 \leq 0,\ -x_1 \leq 0,\ -x_2 > 0$

$$\begin{cases} 2x_1 = 0 \\ 2x_2 + k\left(-2(-x_2)\right) = 0 \end{cases} \Rightarrow \begin{cases} x_1 = 0 \\ (1+k)x_2 = 0 \end{cases} \Rightarrow x_1 = x_2 = 0 \text{ donc impossible.}$$

7è cas: $-x_1 - x_2 + 1 \leq 0,\ -x_1 > 0,\ -x_2 \leq 0$

Ce cas est symétrique au cas précédent (permuter x_1 et x_2) donc impossible.

8è cas: $-x_1 - x_2 + 1 \leq 0,\ -x_1 \leq 0,\ -x_2 \leq 0$

$$\begin{cases} 2x_1 = 0 \\ 2x_2 = 0 \end{cases} \Rightarrow x_1 = x_2 = 0 \text{ donc impossible.}$$

La suite des $x_k = \begin{pmatrix} \frac{k}{1+2k} \\ \frac{k}{1+2k} \end{pmatrix}$ est convergente. D'après théorème, elle converge vers la solution de

$$(P):\ x_k = \begin{pmatrix} \frac{k}{1+2k} \\ \frac{k}{1+2k} \end{pmatrix} \xrightarrow{k} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

La solution de (P) vérifie le critère de qualification de l'indépendance linéaire, les fonctions sont C^1 donc les limites suivantes nous donnent les multiplicateurs de Lagrange de (P):

$$\begin{cases} 2kg_1^+(x_k) = 2k\left(-\frac{2k}{1+2k}+1\right) = \frac{2k}{1+2k} \xrightarrow{k} 1 \\ 2kg_2^+(x_k) = 0 \\ 2kg_3^+(x_k) = 0 \end{cases}$$

Méthode de Barrière:

La méthode de barrière est une méthode primale dans le sens où les déplacements se font à l'intérieur de l'ensemble des solutions réalisables, une fonction barrière empêchant d'en sortir. On résout une suite de problèmes qui par la présence de la barrière sont en pratique sans contraintes.

On considère le problème (P): minimiser $f(x) \quad x \in S \subset R^n$

tel que l'intérieur de S est non vide et pour tout x^* solution de (P) tout voisinage $V(x^*)$ de x^* rencontre l'intérieur de S. S est dit robuste. Cela signifie que l'on peut atteindre x^* par l'intérieur de S.

B une fonction barrière est une fonction définie sur l'intérieur de S et

$$\text{t.q.} \begin{cases} 1) \ B(x) \text{ continue} \\ 2) \ B(x) \geq 0 \\ 3) \ B(x) \to +\infty \text{ ssi } x \text{ s' approche de la frontière de } S \end{cases}$$

Exemple: $S=\{x: g_i(x)\le 0\ i=1,\ldots,m\}$ où les g_i sont continues et t.q. $\overset{o}{S}=\{x: g_i(x)<0\ i=1,\ldots,m\}$ ($\overset{o}{S}$ est l'intérieur de S). Sous ces hypothèses la frontière de S est l'ensemble $\{x\in S: \exists i \text{ t.q. } g_i(x)=0\}$.

Alors on peut prendre comme barrière:

$$B(x)=-\sum_{i=1}^{m}\frac{1}{g_i(x)},\ B(x)=-\sum_{i=1}^{m}Log\left(-g_i(x)\right) \text{ avec } -1\le g_i(x)$$

On considère le problème: minimiser $f(x)+\frac{1}{c}B(x)\ \ x\in\overset{o}{S}$ où $c>0$ ($\overset{o}{S}$ est l'intérieur de S)

Ce problème peut être résolu par des méthodes d'optimisation sans contraintes, les propriétés 1 et 3 de la barrière B empêchant de sortir de S.

Cependant pour que ce problème soit équivalent à (P), il faut c grand pour que $\frac{1}{c}B(x)$ soit négligeable et ainsi minimiser $f(x)$ sans écarter les éventuelles solutions de (P) situées sur la frontière de S. La valeur de c n'étant pas connue a priori, on se donne une suite (c_k) positive et strictement croissante qui tend vers l'infini (par exemple k) et on résout la suite de problèmes (P_k) minimiser $f(x)+\frac{1}{c_k}B(x)\ \ x\in\overset{o}{S}$. On obtient (x_k) la suite des solutions.

Exemple: soit (P) minimiser $f(x_1,x_2)=x_1^2+x_2^2$ sous la contrainte $x_1+x_2\geq 1$

On met la contrainte sous la forme $g(x_1,x_2)=-x_1-x_2+1\leq 0$

Utilisons la barrière $B(x_1,x_2)=-Log(x_1+x_2-1)$

On résout la suite de problèmes (P_k):

minimiser $x_1^2+x_2^2+\frac{1}{k}\left(-Log(x_1+x_2-1)\right)$ sous la contrainte $x_1+x_2>1$

Le domaine des solutions réalisables étant ouvert, le minimum est atteint nécessairement en un point critique. De plus la fonction étant convexe (comme somme de fonctions convexes), tout point critique est optimum global.

On a donc à résoudre: (1) $\begin{cases} 2x_1+\frac{1}{k}\left(\frac{-1}{x_1+x_2-1}\right)=0 \\ 2x_2+\frac{1}{k}\left(\frac{-1}{x_1+x_2-1}\right)=0 \end{cases}$

On en déduit le système $\begin{cases} x_1=x_2 \\ 2x_1+\frac{1}{k}\left(\frac{-1}{2x_1-1}\right)=0 \end{cases}$

La deuxième équation est du second degré et finalement on trouve 2 solutions au système:

$x_1 = x_2 = \frac{1 \pm \sqrt{1+\frac{4}{k}}}{4}$. Seule la solution $x_1 = x_2 = \frac{1+\sqrt{1+\frac{4}{k}}}{4}$ vérifie la contrainte $x_1 + x_2 > 1$.

$\frac{1+\sqrt{1+\frac{4}{k}}}{4} \xrightarrow{k} \frac{1}{2}$ donc la suite des solutions de (P_k) converge vers la solution de (P).

Observant le système (1), on peut considérer $\lambda_k = \frac{1}{k}\left(\frac{1}{x_1+x_2-1}\right)$ comme une approximation du multiplicateur de Lagrange de (P). Remplacant x_1, x_2 on obtient $\lambda_k = \frac{1}{k}\frac{1}{-\frac{1}{2}+\frac{\sqrt{1+\frac{4}{k}}}{2}} \xrightarrow{k} 1$

Exercices

<u>Exercice 1 :</u> Soit le problème d'optimisation suivant :

$$\text{Min } f(x,y)= (x+y)$$

$$\text{s. à :}$$

$$g(x,y)=x^2+(y-1)^2-2=0$$

1. Résoudre graphiquement ce problème.
2. Soit le point $x^*= (-1,0)$.
   - Vérifier que ce point est régulier et donner l'ensemble $T(S, x^*)$.
   - Indiquer sur le graphe le gradient de f et le gradient de h.
   - Donner l'équation de la tangente au point x^*.
3. Résoudre ce problème en utilisant la méthode du sous gradient avec $\lambda^{(0)}=1/4$ et un pas : $\rho=1/4$

Exercice 2 : Soit le problème d'optimisation suivant :

$$\text{Min } (16x_1+10x_2+4x_4)$$

$$\text{s. c. : } 8x_1+2x_2+x_3+4x_4 \geq 10$$

$$x_1+x_2 \geq 1$$

$$x_3+x_4 \geq 1$$

$$x_i \in \{0,1\}$$

1. Résoudre ce problème en utilisant la méthode des plans coupant avec $X^0=\{ (1,1,1,1), (1,0, 1,0) \}$
2. En utilisant la méthode des sous gradient avec
   I. un pas $\rho_k= 1/2^{(k+3)}$ (k=0,1,.. .) et $\lambda^{(0)}=0$
   II. Un pas de l'itération k égal :

$$\rho^k = \frac{\delta_k (z-\theta(\lambda^k)}{\left\|\beta(x^k)\right\|^2}$$

Prendre $\delta_k=2/2^k$ (k=0, 1,…) et $\lambda^{(0)}=0$

Exercice 3 : on considère le problème : $\text{Min}(x_1+x_2)$

s.c. $g(x)=x_1^2-x_2-2\leq 0$

Résoudre ce problème par la méthode de Barrière en prenant $\mu_k=k$ et comme fonction de Barrière B(x) :

- $B(x)=-1/g(x)$
- $B(x)=-\log(-g(x))$

Exercice 4:

Soit le problème suivant :

(P) $$\min_{(x_1,x_2)\in R^2} f(x_1,x_2)=3x_1^2+2x_2^2 \text{ sous la contrainte } x_1+x_2\geq 1$$

1) Conditions de Kuhn-Tucker.

a) Résoudre les conditions de Kuhn-Tucker.

b) Sont-elles des conditions suffisantes d'optimalité dans le cas de (P) et pourquoi ?

c) En déduire x^* la solution de (P) et calculer $f(x^*)$.

2) Dualité.

On note $\theta(\lambda)=\min_{(x_1,x_2)\in R^2} 3x_1^2+2x_2^2+\lambda(-x_1-x_2+1)$ la fonction duale.

a) Donner l'expression de $\theta(\lambda)$ en fonction de λ uniquement. Vérifier que θ est concave.

b) Enoncer (D) le problème dual de (P) et le résoudre.

c) Soit λ^* la solution optimale de (D). Calculer $\theta(\lambda^*)$. Existe-t-il un saut de dualité entre (P) et (D) ?

3) Pénalité de Beltrami.

On pose $g(x)=1-x_1-x_2$. La contrainte de (P) s'écrit donc $g(x)\leq 0$.

a) Résoudre (P) par la méthode de pénalité de Beltrami avec $\mu_k=k$.

b) Retrouver le multiplicateur de Lagrange des conditions de Kuhn-Tucker.