4.3 Contraintes de type ~~[illegible]~~ $Bx = c$ ou $Bx \geq c$

Nous voulons caractériser le minimum du problème

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

où $U = \{x \in \mathbb{R}^n \mid Bx = c\}$ ou $\{x \in \mathbb{R}^n \mid Bx \geq c\}$ ~~$g_1(x) = \ldots = g_m(x) = 0\}$~~

Ceci va conduire à la méthode des multiplicateurs de Lagrange.

Commençons par ~~un cas particulier~~ des contraintes de type égalité.

4.3.1 Contraintes (égalité) linéaires

$$U = \{x \in \mathbb{R}^n \mid Bx = c\}$$

où B = matrice de format $m \times n$

Débutons avec $c = 0$ $\Rightarrow$ $U = \underbrace{\operatorname{Ker} B}_{\text{sous-espace}} \subset \mathbb{R}^n$

A la première section, on a vu que le minimum $a \in U$ est caractérisé par la relation

$$\nabla f(a) \in U^* = U^\perp = (\operatorname{Ker} B)^\perp$$

Or $(\operatorname{Ker} B)^\perp = \operatorname{Im} B^t$

$\Rightarrow$ il existe $\lambda \in \mathbb{R}^m$ t.q.

$$\nabla f(a) = -B^t \lambda \quad \left(= B^t(-\lambda) \right)$$

$$\Rightarrow \begin{cases} \nabla f(a) + B^t \lambda = 0 \\ B\,a = 0 \end{cases}$$
et

Donc, le couple (a, λ) est solution du système.

$$\begin{cases} \nabla f(x) + B^t \lambda = 0 \\ B\,x = 0 \end{cases}$$

Ceci constitue un système non linéaire (en général) de $n+m$ équations à $n+m$ variables (x, λ).
Les variables λ portent le nom de multiplicateurs de Lagrange.

En fait, introduisons la fonction des variables (x, λ)

$$L(x, \lambda) = f(x) + (\lambda, Bx)$$

Il est facile de montrer que

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + (B^t \lambda)_i \; ; \text{ i.e. } \nabla_x L = \nabla f + B^t \lambda$$

et

$$\frac{\partial L}{\partial \lambda_i} = (B\,x)_i$$

Donc, les points critiques de L p/r aux variables x et λ

$$\nabla_{(x,\lambda)} L = 0$$

$$\Leftrightarrow \begin{cases} \nabla f(x) + B^t \lambda = 0 \\ B\,x = 0 \end{cases}$$

C'est la méthode des multiplicateurs de Lagrange pour ce problème.

Considérons le cas particulier où

$$f(x) = \frac{1}{2}(Ax, x) - (b, x) \quad \text{où } A > 0 \text{ sym. et } b \in \mathbb{R}^n$$

$$\Rightarrow Df(x) = Ax - b$$

Le système $DL = 0$ s'écrit:

$$\begin{cases} Ax + B^t\lambda = b \\ Bx = 0 \end{cases}$$

ou, sous forme matricielle,

$$\begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix} \quad (*)$$

En résumé, pour résoudre le problème de minimisation

$$\min_{Bx=0} \frac{1}{2}(Ax, x) - (b, x)$$

il suffit de résoudre le système linéaire ci-dessus (*)

Remarque: en général, le système (*) peut admettre une infinité de solutions en λ.

En effet (x, λ) et $(x, \lambda + m)$ où $m \in \operatorname{Ker} B^t$ sont 2 solutions de (*). Pour obtenir unicité, il faut exiger que

$$\operatorname{Ker} B^t = \{0\}$$

ou

$$\operatorname{Im} B = \mathbb{R}^m \quad (\text{par dualité})$$

ou

$$rg(B) = m$$

Exemple : $\quad \min\limits_{x_1+x_2+x_3=0} \dfrac{(x_1-1)^2+(x_2-2)^2+(x_3-3)^2}{2}$

La contrainte s'écrit :

$$B\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix} = (1\ \ 1\ \ 1)\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix}$$

$$B = (1\ 1\ 1)_{1\times 3} \qquad \text{et } rg(B)=1 \Rightarrow \operatorname{Ker} B^t = \{0\}$$

$$f(x_1,x_2,x_3) = \frac{1}{2}(Ax,x) - (b,x) \quad + \text{cte}$$

$$= \frac{1}{2}\left[x_1^2+x_2^2+x_3^2\right] - x_1 - 2x_2 - 3x_3 + \text{cte}$$

$$\Rightarrow A = \begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \text{ et } b = \begin{pmatrix} +1\\ +2\\ +3\end{pmatrix} = \begin{pmatrix} 1\\ 2\\ 3\end{pmatrix}$$

L'unique pt. minimisant $x \in \mathbb{R}^3$ est la solution du système linéaire

$$\begin{pmatrix} A & B^t\\ B & 0\end{pmatrix}\begin{pmatrix} x\\ \lambda\end{pmatrix} = \begin{pmatrix} b\\ 0\end{pmatrix}$$

c'est-à-dire

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 1\\ \hline 1 & 1 & 1 & 0\end{array}\right)\begin{pmatrix} x_1\\ x_2\\ x_3\\ \lambda\end{pmatrix} = \begin{pmatrix} 1\\ 2\\ 3\\ 0\end{pmatrix}$$

On trouve :
$$\begin{cases} x_1 = -1\\ x_2 = 0\\ x_3 = 1\\ \lambda = 2\end{cases}$$

Maintenant considérons le problème

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

où $U = \{ x \in \mathbb{R}^n \mid Bx = c \}$ = sous-espace affine de $\mathbb{R}^n$ avec $c \neq 0 \in \mathbb{R}^m$

Il est clair que ce problème a un sens si $c \in \operatorname{Im} B$ car sinon $U = \emptyset$.

La condition $c \in \operatorname{Im} B$ est toujours vérifiée si

$$rg(B) = m$$

car, dans ce cas, $\operatorname{Im} B = \mathbb{R}^m$.

A la section précédente, on a vu que la condition d'optimalité dans le cas d'un sous-espace affine

$$U = x_0 + \operatorname{Ker} B$$

où $x_0 \in \mathbb{R}^n$ vérifie $Bx_0 = c$ ($c \in \operatorname{Im} B$).

s'écrit

$$\nabla f(a) \in (\operatorname{Ker} B)^\perp$$

De nouveau $(\operatorname{Ker} B)^\perp = \operatorname{Im} B^t$

il existe $\lambda \in \mathbb{R}^m$ tel que

et $\begin{cases} \nabla f(a) + B^t \lambda = 0 \\ Ba = c \end{cases}$

Donc, le pt (a, λ) est solution du système

$$\begin{cases} \nabla f(x) + B^t \lambda = 0 \\ Bx = c \end{cases}$$

On peut interpréter ce système sous la forme des multiplicateurs de Lagrange.

$$L(x, \lambda) = f(x) + (\lambda, Bx - c)$$

$$\Rightarrow \quad D_x L = \nabla f + B^t \lambda = 0$$

$$D_\lambda L = Bx - c = 0$$

Exemple: $$\min_{x_1 + 2x_2 + 3x_3 = 4} \frac{x_1^2 + x_2^2 + x_3^2}{2}$$

$m = 1 \quad x_1 + 2x_2 + 3x_3 = 4 \quad \Leftrightarrow \quad Bx = c$

$$\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4$$

$f(x) = \frac{1}{2}(Ax, x) \quad \Rightarrow \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$\Rightarrow \nabla f = Ax$

Le pt. minimisant est l'unique solution du système

$$\begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 1 & 2 & 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 4 \end{pmatrix}$$

de solution:

$x_1 = 2/7$

$x_2 = 4/7$

$x_3 = 6/7$

$\lambda = -2/7$

Remarque: étant donné que

$$\max_{x \in U} f(x) = -\min_{x \in U} -f(x)$$

La condition d'optimalité pour le problème de maximisation correspond à :

$$\begin{cases} D(-f)(x) + B^t \lambda = 0 \\ Bx = c \end{cases}$$

ou

$$\begin{cases} Df(x) + B^t(-\lambda) = 0 \\ Bx = c \end{cases}$$

donc solution du <u>même</u> système

$$\begin{cases} Df(x) + B^t \lambda = 0 \\ Bx = c \end{cases}$$

Ceci signifie que la condition d'optimalité ne distingue pas les min et les max. Il faut les trier.

Exemple:
$$\max_{x_1+x_2+x_3=10} x_1 x_2 x_3$$

$\Rightarrow L(x,\lambda) = x_1 x_2 x_3 + \lambda(x_1+x_2+x_3-10)$

Il faut résoudre le système

$$\begin{cases} x_2 x_3 + \lambda = 0 \\ x_1 x_3 + \lambda = 0 \\ x_1 x_2 + \lambda = 0 \\ x_1+x_2+x_3 = 10 \end{cases} \Rightarrow \begin{array}{l} x_1 x_3 = x_2 x_3 \\ x_3 = 0 \text{ ou } x_1 = x_2 \end{array}$$

Si $x_3 = 0 \Rightarrow \lambda = 0 \Rightarrow x_1 x_2 = 0$

$\Rightarrow (10, 0, 0)$ ou $(0, 10, 0)$

Si $x_3 \neq 0 \Rightarrow x_1 = x_2$

$\Rightarrow x_1 x_2 = x_1 x_3 \Rightarrow x_1 = 0$ ou $x_1 = x_3$

si $x_1 = 0 \Rightarrow \lambda = 0 \Rightarrow x_2 = 0$

$\Rightarrow (0, 0, 10)$

si $x_1 \neq 0 \Rightarrow x_1 = x_3$ et $x_1 = x_2$

$\Rightarrow x_1 = x_2 = x_3 = 10/3$

Il est clair que le maximum est atteint au pt

$(10/3, 10/3, 10/3)$

4.3.2 Contraintes linéaires de type inégalité

On considère le problème de minimisation défini sur le cône convexe (fermé)

$$U = \{ x \in \mathbb{R}^n \mid Bx \geq 0 \}$$

où B est une matrice de format $m \times n$.

Suivant la théorie générale où U est convexe, la condition d'optimalité s'écrit

$$a \in U \qquad (\nabla f(a), x - a) \geq 0 \qquad \forall x \in U$$

Mais U est un cône convexe, on peut choisir
$$x = a + y \qquad \text{pour } y \in U$$

$$\Rightarrow \qquad (\nabla f(a), y) \geq 0 \qquad \forall y \in U$$

C'est-à-dire que $\quad \nabla f(a) \in U^*$

Mais, à la section 4.2, nous avons calculé U^*

$$U^* = \{ B^t \lambda \mid \lambda \geq 0 \} = B^t K^+$$

$\Rightarrow \exists \lambda \in \mathbb{R}^m$ avec $\lambda \geq 0$ tel que

$$\boxed{\nabla f(a) - B^t \lambda = 0}$$

Ceci constitue une première relation d'optimalité! Les λ_i portent aussi le nom de multiplicateurs de Lagrange.

Le fait que U soit un cône implique une autre relation.

$$\left(\nabla f(a), a\right) = 0$$

En effet, on prend respectivement $x = 2a$ et $x = 0 \in U$ dans la relation $(\nabla f(a), x-a) \geq 0$.

Mais $\nabla f(a) = B^t\lambda \Rightarrow (B^t\lambda, a) = 0$

$$\Rightarrow (\lambda, Ba) = 0$$

Aussi $\lambda \geq 0$ et $Ba \geq 0$ car $a \in U$

$$\Rightarrow \lambda_i (Ba)_i = 0 \quad \forall i = 1, \dots, m$$

En conclusion, le couple $(a, \lambda) \in U \times \mathbb{R}^m$ vérifie les conditions dites de Kuhn-Tucker (KKT)

$$\begin{cases} \nabla f(x) - B^t \lambda = 0 \\ (Bx)_i \geq 0 \\ \lambda_i \geq 0 \\ \lambda_i (Bx)_i = 0 \quad \forall i = 1, \dots, m \end{cases}$$

Exemples:

$$\min_{\substack{x_1 \geq 0 \\ x_2 \geq 0}} \underbrace{\frac{1}{2}\left[2x_1^2 + 2x_1x_2 + x_2^2\right] - x_2}_{f(x_1, x_2)}$$

$$Bx = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\nabla f(x) = (2x_1 + x_2, x_1 + x_2 - 1)$$

Les conditions KKT s'écrivent

$$\begin{cases} 2x_1 + x_2 = \lambda_1 \geq 0 \\ x_1 + x_2 - 1 = \lambda_2 \geq 0 \\ \lambda_1 x_1 = 0 \\ \lambda_2 x_2 = 0 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

Si $\lambda_1 = 0 \Rightarrow 2x_1 + x_2 = 0 \Rightarrow x_2 = -2x_1$

impossible car $x_1, x_2 \geq 0$

sauf en $(0,0)$

$\Rightarrow \lambda_2 = -1$ impossible

$\Rightarrow \lambda_1 \neq 0 \Rightarrow x_1 = 0 \Rightarrow$ $x_2 = \lambda_1$

$x_2 - 1 = \lambda_2 \Rightarrow x_2 = 1 + \lambda_2 \geq 1$

≥ 0

$x_2 = \lambda_1 > 0$ (car $\lambda_1 \neq 0$) aussi

$\Downarrow$

$\lambda_2 = 0 \Rightarrow x_2 - 1 = \lambda_2 = 0 \Rightarrow x_2 = 1$

D'où $(0,1)$ est l'unique pt. min.

Les multiplicateurs λ_1 et λ_2 sont :

$$\lambda_1 = x_2 = 1 \geq 0$$

$$\lambda_2 = x_2 - 1 = 0$$

Traitons maintenant le cas plus général

$$U = \{x \in \mathbb{R}^n \,/\, Bx \geq c\} \subset \mathbb{R}^n$$
convexe fermé

où B est une matrice de format $m \times n$ et $c \in \mathbb{R}^m$.

On fera l'hypothèse qu'il existe un point $x_0 \in \mathbb{R}^n$ de sorte que

$$Bx_0 = c.$$

On peut écrire U sous la forme

$$U = \{x \in \mathbb{R}^n \,/\, B(x - x_0) \geq 0\}$$

Ceci montre que U est le translaté d'un cône de sommet 0 $(\{By \geq 0\})$

Etudions la condition d'optimalité du problème

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

qui est

$$a \in U, \quad (\nabla f(a), x - a) \geq 0 \quad \forall x \in U$$

Prenons x de la forme : $x = a + y$ où $By \geq 0$

$\Rightarrow x \in U$ car $B(a+y) = Ba + By \geq Ba \geq c$

$$\Rightarrow (\nabla f(a), y) \geq 0 \quad \forall y \text{ t.q. } By \geq 0$$

$$\Rightarrow \nabla f(a) \in \{y \,/\, By \geq 0\}^*$$

On a déjà calculé ce cône dual

$\Rightarrow \quad \nabla f(a) \in B^t K^+$

$\Rightarrow \exists\, \lambda \in \mathbb{R}^m$ et $\lambda \geq 0$ de sorte que

$$\boxed{\nabla f(a) - B^t \lambda = 0}$$

D'autre part, on a aussi une autre condition

$$(\nabla f(a),\ a - x_0) = 0$$

En effet, on prend $x = x_0 \in U$

$\Rightarrow (\nabla f(a),\ x_0 - a) \geq 0$

et

$x = x_0 + 2(a - x_0) \in U$

car $Bx_0 + 2(Ba - Bx_0)$

$= c + 2(Ba - c) = 2Ba - c$

$\geq 2c - c = c$

$\Rightarrow (\nabla f(a),\ x - a) = (\nabla f(a),\ a - x_0) \geq 0$

d'où l'égalité.

Maintenant, on utilise le fait que $B^t \lambda = \nabla f(a)$

$\Rightarrow \quad (\nabla f(a),\ a - x_0) = 0$

$\Leftrightarrow \quad (B^t \lambda,\ a - x_0) = 0$

$\Leftrightarrow \quad (\lambda,\ B(a - x_0)) = 0$

$$\Leftrightarrow \quad (\lambda, Ba - c) = 0$$

En conclusion, le couple (a, λ) doit satisfaire le système de Kuhn-Tucker

$$\begin{cases} \nabla f(x) - B^t \lambda = 0 \\ \lambda_i \geq 0 \\ (Bx)_i \geq c_i & \forall i = 1, \dots m \\ \lambda_i \, (Bx - c)_i = 0 \end{cases}$$

Remarque: la dernière relation est une relation de complémentarité.

Si, pour un indice i donné, on a que

$$(Bx)_i > c_i$$

(on dit que la contrainte i est <u>inactive</u>)

on doit avoir $\lambda_i = 0$ (le multiplicateur est nul)

Par contre, si le multiplicateur $\lambda_i > 0$, la contrainte i doit être <u>active</u>

$$(Bx)_i = c_i.$$