

Introduction aux multiplicateurs de Lagrange : Recherche de minimum d'une fonction quadratique sous contraintes affines

Gilles LEBORGNE

4 février 2014

Table des matières

| | | |
|----------|--|-----------|
| 1 | Introduction | 2 |
| 1.1 | Plan tangent (hyperplan tangent) | 2 |
| 1.2 | Extremum sans contrainte | 4 |
| 1.3 | Extremum sous contrainte affine | 4 |
| 1.3.1 | La contrainte affine | 4 |
| 1.3.2 | L'espace affine P associé et l'espace vectoriel \vec{P} associé | 4 |
| 1.4 | La fonction $f _P$ à minimiser | 5 |
| 1.5 | Le multiplicateur de Lagrange | 5 |
| 1.6 | Le problème à résoudre | 6 |
| 1.7 | Le lagrangien | 7 |
| 1.8 | Forme matricielle usuelle | 7 |
| 2 | p contraintes : les p multiplicateurs de Lagrange | 8 |
| 2.1 | Caractérisation de l'intersection de p hyperplans dans \mathbb{R}^n | 8 |
| 2.2 | Le problème contraint | 9 |
| 2.3 | Les p multiplicateurs de Lagrange | 9 |
| 2.4 | Le problème contraint à résoudre | 10 |
| 2.5 | Le lagrangien | 10 |
| 2.6 | Matrice B | 11 |
| 2.7 | Forme matricielle usuelle | 11 |
| 2.8 | n valeurs propres positives et p valeurs propres négatives | 12 |
| 3 | Problèmes primal et dual | 12 |
| 3.1 | Problème primal | 12 |
| 3.2 | Problème dual et matrice d'Uzawa | 12 |
| 3.3 | Problème d'inf-sup pour le lagrangien (point selle du lagrangien) | 14 |
| 4 | Résolution | 14 |
| 4.1 | Pénalisation | 14 |
| 4.2 | Présentation de la méthode de résolution itérative d'Uzawa | 16 |
| 4.2.1 | Présentation | 16 |
| 4.2.2 | Première amélioration | 16 |
| 4.2.3 | Algorithme | 17 |
| 4.3 | Lagrangien augmenté | 17 |
| 4.4 | Lagrangien augmenté pénalisé : résolution itérative d'Uzawa | 18 |
| 4.5 | Lagrangien augmenté pénalisé : convergence du schéma d'Uzawa | 19 |
| 4.6 | Choix de ρ^n | 20 |
| 4.7 | Uzawa : c'est un algorithme de gradient | 22 |
| A | Annexe. Dimensions des noyaux et images | 22 |
| B | Annexe. Matrices similaires et congruentes | 23 |
| B.1 | Matrices similaires | 23 |
| B.2 | Matrices congruentes et loi d'inertie de Sylvester | 24 |

| | |
|--|-----------|
| C Annexe. Point selle | 25 |
| C.1 Définition | 25 |
| C.2 “sup inf \leq inf sup” | 26 |
| C.3 L’inf-sup | 27 |
| C.4 Cas d’un lagrangien | 27 |

1 Introduction

Soit \mathbb{R}^n l’espace affine usuel muni d’une origine. On note également \mathbb{R}^n l’espace vectoriel associé usuel, muni de son produit scalaire euclidien $(\cdot, \cdot)_{\mathbb{R}^n}$ et de sa base canonique.

Ainsi un point de \mathbb{R}^n espace affine sera représenté par le vecteur $\vec{x} \in \mathbb{R}^n$ (grâce à l’origine choisie dans \mathbb{R}^n).

1.1 Plan tangent (hyperplan tangent)

On note $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ un point dans \mathbb{R}^n . Pour $f : \mathbb{R}^n \rightarrow \mathbb{R}$, le graphe $G(f)$ de f est le sous-ensemble de \mathbb{R}^{n+1} défini par :

$$G(f) = \{(\vec{x}, z) \in \mathbb{R}^n \times \mathbb{R} \text{ t.q. } z = f(\vec{x})\} \subset \mathbb{R}^{n+1}. \quad (1.1)$$

Et l’hypersurface d’équation :

$$z = f(\vec{x}) \quad (1.2)$$

représente $G(f)$ dans \mathbb{R}^{n+1} , voir figure 1.1.

(Si $n = 1$ alors $G(f)$ est une courbe dans \mathbb{R}^2 , si $n = 2$ alors $G(f)$ est une surface \mathbb{R}^3 , et si $n \geq 3$ alors $G(f)$ est appelée une hypersurface, et le mot surface ou hypersurface est utilisé si n est un entier générique.)

Pour $f \in C^1(\mathbb{R}^n; \mathbb{R})$, on notera $\vec{\nabla} f(\vec{x})$ la matrice colonne = $\begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x}) \end{pmatrix} \in \mathbb{R}^n$, et $\vec{\nabla} f(\vec{x})^T$ la matrice ligne transposée. Ainsi, le développement limité de f au premier ordre au voisinage d’un point $\vec{x}_0 \in \mathbb{R}^n$ est donné par :

$$\begin{aligned} f(\vec{x}) &= f(\vec{x}_0) + (x_1 - x_{01}) \frac{\partial f}{\partial x_1}(\vec{x}_0) + \dots + (x_n - x_{0n}) \frac{\partial f}{\partial x_n}(\vec{x}_0) + o(\|\vec{x} - \vec{x}_0\|) \\ &= f(\vec{x}_0) + \vec{\nabla} f(\vec{x}_0)^T \cdot (\vec{x} - \vec{x}_0)_{\mathbb{R}^n} + o(\|\vec{x} - \vec{x}_0\|), \\ &= f(\vec{x}_0) + (\vec{\nabla} f(\vec{x}_0), \vec{x} - \vec{x}_0)_{\mathbb{R}^n} + o(\|\vec{x} - \vec{x}_0\|), \end{aligned} \quad (1.3)$$

où $(\cdot, \cdot)_{\mathbb{R}^n}$ est le produit scalaire euclidien de \mathbb{R}^n .

L’approximation de f au premier ordre au voisinage de \vec{x}_0 est la fonction affine $p_{\vec{x}_0} \in C^1(\mathbb{R}^n; \mathbb{R})$ définie par :

$$p_{\vec{x}_0}(\vec{x}) = f(\vec{x}_0) + (\vec{\nabla} f(\vec{x}_0), (\vec{x} - \vec{x}_0))_{\mathbb{R}^n}. \quad (1.4)$$

Le graphe de $p_{\vec{x}_0}$ est :

$$G(p_{\vec{x}_0}) = \{(\vec{x}, z) \in \mathbb{R}^n \times \mathbb{R} \text{ t.q. } z = f(\vec{x}_0) + (\vec{\nabla} f(\vec{x}_0), (\vec{x} - \vec{x}_0))_{\mathbb{R}^n}\} \subset \mathbb{R}^{n+1}. \quad (1.5)$$

Ainsi $G(p_{\vec{x}_0})$ est l’hyperplan dans \mathbb{R}^{n+1} d’équation :

$$z = f(\vec{x}_0) + (\vec{\nabla} f(\vec{x}_0), (\vec{x} - \vec{x}_0))_{\mathbb{R}^n}, \quad (1.6)$$

voir figure 1.1.

(Si $n = 1$ alors $G(p_{\vec{x}_0})$ est une droite dans \mathbb{R}^2 , si $n = 2$ alors $G(p_{\vec{x}_0})$ est un plan dans \mathbb{R}^3 , et si $n \geq 3$ alors $G(p_{\vec{x}_0})$ est appelée un hyperplan, et le mot plan ou hyperplan est utilisé si n est un entier générique.)

Et un vecteur normal à l'hyperplan $G(p_{\vec{x}_0})$ est par exemple :

$$\begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x}) \\ -1 \end{pmatrix} \stackrel{\text{noté}}{=} \begin{pmatrix} \vec{\nabla} f(\vec{x}_0) \\ -1 \end{pmatrix}, \quad (1.7)$$

puisque $\left(\begin{pmatrix} \vec{\nabla} f(\vec{x}_0) \\ -1 \end{pmatrix}, \begin{pmatrix} \vec{x} \\ z \end{pmatrix} \right)_{\mathbb{R}^n} = 0$, soit $(\vec{\nabla} f(\vec{x}_0), (\vec{x} - \vec{x}_0))_{\mathbb{R}^n} - z = 0$, soit $z = (\vec{\nabla} f(\vec{x}_0), (\vec{x} - \vec{x}_0))_{\mathbb{R}^n}$, équation de l'hyperplan vectoriel "parallèle" à l'hyperplan affine (1.6), voir figure 1.1.

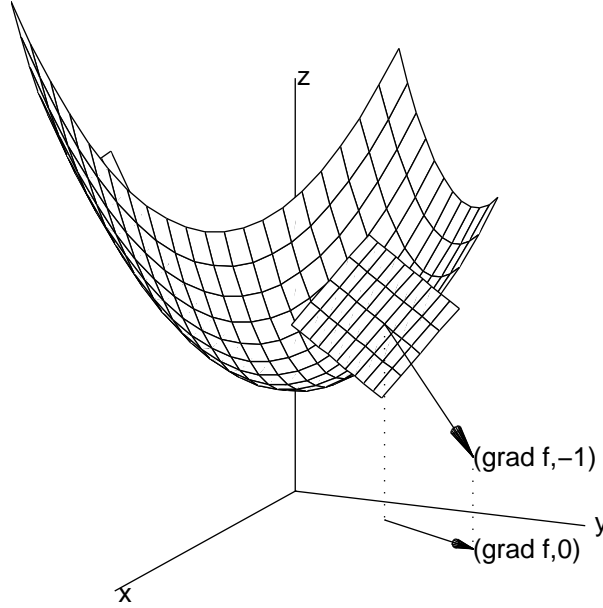


FIGURE 1.1 – Graphe d'une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ quadratique et le plan tangent en un point $\vec{x} = (x, y)$. Un vecteur normal sur le graphe de la fonction est le vecteur $(\vec{\nabla} f(\vec{x}), -1) \in \mathbb{R}^3$. Et le vecteur $\vec{\nabla} f(\vec{x}) \in \mathbb{R}^2$ donne, dans le plan (x, y) , la direction de la plus grande pente.

Remarque 1.1 Pour $\vec{x}_0 \in \mathbb{R}^n$, (1.3) donne en particulier, pour tout $\vec{v} \in \mathbb{R}^n$ et tout $h \in \mathbb{R}$ dans un voisinage de 0 :

$$f(\vec{x}_0 + h\vec{v}) = f(\vec{x}_0) + h(\vec{\nabla} f(\vec{x}_0), \vec{v})_{\mathbb{R}^n} + o(h), \quad (1.8)$$

et pour le plan tangent en \vec{x}_0 , pour tout $h \in \mathbb{R}$:

$$p_{\vec{x}_0}(\vec{x}_0 + h\vec{v}) = f(\vec{x}_0) + h(\vec{\nabla} f(\vec{x}_0), \vec{v})_{\mathbb{R}^n}. \quad (1.9)$$

Ainsi le taux de variation $\frac{p_{\vec{x}_0}(\vec{x}_0 + h\vec{v}) - p_{\vec{x}_0}(\vec{x}_0)}{h} = (\vec{\nabla} f(\vec{x}_0), \vec{v})_{\mathbb{R}^n}$ est maximal en valeur absolue quand on choisit $\vec{v} \parallel \vec{\nabla} f(\vec{x}_0)$ (cf. inégalité de Cauchy–Schwarz). En particulier $\vec{v} = \vec{\nabla} f(\vec{x}_0)$ donne la direction de "la plus grande pente" le long du plan tangent en \vec{x}_0 . Voir figure 1.1. ■

Exemple 1.2 Dans \mathbb{R}^2 . On note $\vec{x} = (x, y)$, $\vec{x}_m = (x_m, y_m)$. Soit α, β deux réels > 0 et soit $f(x, y) = 1 + \alpha(x - x_m)^2 + \beta(y - y_m)^2$. On a $\vec{\nabla} f(\vec{x}) = \begin{pmatrix} 2\alpha(x - x_m) \\ 2\beta(y - y_m) \end{pmatrix}$. Le graphe de f est le paraboloïde de \mathbb{R}^3 :

$$G(f) = \{(\vec{x}, z) \in \mathbb{R}^2 \times \mathbb{R} \text{ t.q. } z = 1 + \alpha(x - x_m)^2 + \beta(y - y_m)^2\},$$

voir figure 1.1. Le plan tangent en un point $\vec{x}_0 = (x_0, y_0) \in \mathbb{R}^2$ est :

$$z = f(\vec{x}_0) + (\vec{\nabla} f(\vec{x}_0), \vec{x} - \vec{x}_0)_{\mathbb{R}^2} = f(\vec{x}_0) + 2\alpha(x_0 - x_m)(x - x_0) + 2\beta(y_0 - y_m)(y - y_0).$$

C'est, dans \mathbb{R}^3 , le plan tangent au graphe $G(f)$ au point $(\vec{x}_0, f(\vec{x}_0))$, voir figure 1.1.

Et $\vec{\nabla}f(\vec{x}) = 0$ au point \vec{x}_m , avec $f(\vec{x}_m) = 1$: le graphe de f au point $(\vec{x}_m, 1) = (x_m, y_m, 1)$ admet pour plan tangent le plan $z = 1$ (horizontal) dont un vecteur normal est $\begin{pmatrix} -\vec{\nabla}f(\vec{x}_m) \\ 1 \end{pmatrix} = \begin{pmatrix} \vec{0} \\ 1 \end{pmatrix}$ (orthogonal au plan horizontal). \blacksquare

1.2 Extremum sans contrainte

Caractérisation d'un extremum : $f \in C^1(\mathbb{R}^n, \mathbb{R})$ admet un extremum en un point $\vec{x}_m \in \mathbb{R}^n$ ssi $\vec{\nabla}f(\vec{x}_m) = \vec{0}$ (la dérivée de f en ce point est nulle).

L'équation du plan tangent en \vec{x}_m est $p_{\vec{x}_m}(\vec{x}) = f(\vec{x}_m) + \vec{0} \cdot (\vec{x} - \vec{x}_m) = f(\vec{x}_m) = \text{constante}$, cf. (1.4), autrement dit le plan tangent à f en \vec{x}_m est "horizontal". En d'autres termes, le vecteur "vertical" $\begin{pmatrix} -\vec{\nabla}f(\vec{x}_m) \\ 1 \end{pmatrix} = \begin{pmatrix} \vec{0} \\ 1 \end{pmatrix}$ est un vecteur normal au graphe $G(f)$ de f au point $(\vec{x}_m, f(\vec{x}_m))$.

1.3 Extremum sous contrainte affine

1.3.1 La contrainte affine

Soit une fonction affine $g : \vec{x} \in \mathbb{R}^n \rightarrow z = g(\vec{x}) \in \mathbb{R}$, donc de la forme : il existe $\vec{b} \in \mathbb{R}^n$ et $c \in \mathbb{R}$ t.q. :

$$g(\vec{x}) = (\vec{b}, \vec{x})_{\mathbb{R}^n} - c \quad (= b_1x_1 + \dots + b_nx_n - c). \quad (1.10)$$

Comme $\frac{\partial g}{\partial x_i}(\vec{x}) = b_i$ pour tout i (immédiat), on a, pour tout $\vec{x} \in \mathbb{R}^n$:

$$\vec{\nabla}g(\vec{x}) = \vec{b}, \quad (1.11)$$

vecteur indépendant du point où on se trouve.

On ne s'intéressera qu'au cas $\vec{b} \neq \vec{0}$ et qu'aux points $\vec{x} \in \mathbb{R}^n$ tels que :

$$g(\vec{x}) = 0 \quad \text{soit} \quad (\vec{b}, \vec{x})_{\mathbb{R}^n} = c. \quad (1.12)$$

Ce sera notre contrainte, contrainte affine dans \mathbb{R}^n l'ensemble de définition de f .

Exemple 1.3 Soit $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ donné par $g(x_1, x_2) = b_1x_1 + b_2x_2 - c$ avec $\vec{b} = (b_1, b_2) \neq \vec{0}$.

Et donc l'ensemble des $\vec{x} \in \mathbb{R}^2$ t.q. $g(\vec{x}) = 0$ est la droite affine $b_1x_1 + b_2x_2 - c = 0$ de \mathbb{R}^2 . \blacksquare

Exemple 1.4 Soit $g : \mathbb{R}^3 \rightarrow \mathbb{R}$, $g(x_1, x_2, x_3) = b_1x_1 + b_2x_2 + b_3x_3 - c$ avec $\vec{b} = (b_1, b_2, b_3) \neq \vec{0}$.

Et donc l'ensemble des $\vec{x} \in \mathbb{R}^3$ t.q. $g(\vec{x}) = 0$ est le plan affine $b_1x_1 + b_2x_2 + b_3x_3 - c = 0$ de \mathbb{R}^3 . \blacksquare

Exemple 1.5 Le cas $g : \mathbb{R} \rightarrow \mathbb{R}$ est sans intérêt : on aurait $g(x) = bx - c$ avec $b \neq 0$, et l'ensemble des $x \in \mathbb{R}$ t.q. $g(x) = 0$ est réduit à un point, le point $x = \frac{c}{b}$: le problème s'arrête ici. \blacksquare

1.3.2 L'espace affine P associé et l'espace vectoriel \vec{P} associé

Dans la suite, toujours avec $g(\vec{x}) = (\vec{b}, \vec{x})_{\mathbb{R}^n} - c$, cf. (1.10), on note P l'hyperplan affine dans \mathbb{R}^n défini par (toujours pour $\vec{b} \neq \vec{0}$) :

$$P = \{\vec{x} \in \mathbb{R}^n : g(\vec{x}) = 0\}, \quad (1.13)$$

soit :

$$P = \{\vec{x} \in \mathbb{R}^n : (\vec{b}, \vec{x})_{\mathbb{R}^n} = c\}. \quad (1.14)$$

(Pour $\mathbb{R}^n = \mathbb{R}^2$, P est une droite, pour $\mathbb{R}^n = \mathbb{R}^3$, P est un plan.) Et on note :

$$\vec{P} = \{\vec{v} \in \mathbb{R}^n : (\vec{b}, \vec{v})_{\mathbb{R}^n} = 0\} = \text{Vect}\{\vec{b}\}^\perp = \text{Vect}\{\vec{\nabla}g(\vec{x})\}^\perp, \quad (1.15)$$

l'espace vectoriel associé. Donc \vec{P} est un hyperplan vectoriel dans \mathbb{R}^n (sous-espace vectoriel de dimension $n-1$).

Ainsi, si $\vec{x}_P \in P$, i.e. si $(\vec{b}, \vec{x}_P)_{\mathbb{R}^n} = c$, alors on a :

$$P = \vec{x}_P + \vec{P}, \quad (1.16)$$

puisque $\vec{x} \in P$ (i.e. $(\vec{b}, \vec{x})_{\mathbb{R}^n} = c$) $\iff (\vec{b}, \vec{x} - \vec{x}_P)_{\mathbb{R}^n} = 0 \iff \vec{x} - \vec{x}_P \in \vec{P}$.

1.4 La fonction $f|_P$ à minimiser

Soit $f \in C^1(\mathbb{R}^n; \mathbb{R})$. Le problème sous contrainte considéré est : trouver un minimum de f dans P :

$$\text{trouver } \vec{x}_0 \in P \text{ t.q. : } f(\vec{x}_0) = \min_{\vec{x} \in P} f(\vec{x}). \quad (1.17)$$

Notons $f|_P$ la restriction de f à P :

$$f|_P : \begin{cases} P & \rightarrow \mathbb{R} \\ \vec{x} & \rightarrow f|_P(\vec{x}) = f(\vec{x}), \end{cases} \quad (1.18)$$

P étant l'hyperplan affine dans \mathbb{R}^n donné en (1.14). On supposera qu'il existe un point (dans P) où $f|_P$ atteint un extremum. Voir figure 1.2.

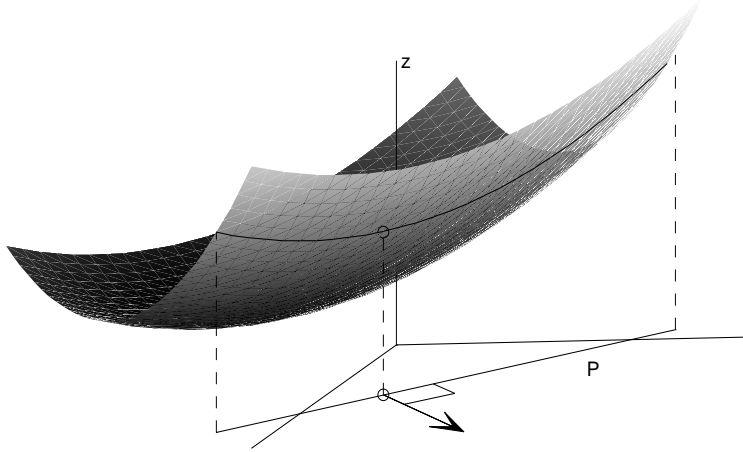


FIGURE 1.2 – $G(f)$ un parabolöide. Contrainte : $x + y - c = 0 = g(x, y)$. Espace de contrainte : $P = \{(x, y) : x + y - c = 0\} =$ la droite dans le plan (xOy) (plan horizontal). La fonction $f|_P$ est définie sur P . Le point (x_0, y_0) en lequel $f|_P$ atteint sur minimum (i.e. le point (x_0, y_0) en lequel f atteint sur minimum dans P) est représenté par le signe o dans le plan horizontal. Et le point du graphe $G(f|_P)$ de $f|_P$ où $f|_P$ est minimale, à savoir le point $(x_0, y_0, z_0 = f|_P(x_0, y_0)) \in \mathbb{R}^3$, est également représenté par le signe o .

Exemple 1.6 Soit $f(x, y) = 1 + \alpha(x - x_m)^2 + \beta(y - y_m)^2$, comme à l'exemple 1.2.

Soit $g(\vec{x}) = ax + by - c$ avec $(a, b) \neq (0, 0)$. On ne s'intéresse qu'aux points $(x, y) \in P$, i.e. t.q. $ax + by - c = 0$.

Dans ce cas simple, on peut calculer directement le point \vec{x}_0 : par exemple pour $b \neq 0$, pour $(x, y) \in P$ on a $y = \frac{-a}{b}x + \frac{c}{b}$, et donc $f(x, y) = f(x, \frac{-a}{b}x + \frac{c}{b}) =$ noté $h(x)$ fonction d'une seule variable x , avec ici $h(x) = 1 + \alpha(x - x_m)^2 + \beta(\frac{-a}{b}x + \frac{c}{b} - y_m)$. Le minimum est obtenu pour x_0 tel que $h'(x_0) = 0$. D'où x_0 (immédiat), d'où la valeur $y_0 = \frac{-a}{b}x_0 + \frac{c}{b}$ et la valeur $z_0 = f(x_0, y_0)$: le minimum de f le long de la droite $ax + by + c = 0$ au lieu au point $\vec{x}_0 = (x_0, y_0)$, et la valeur de f en ce point est z_0 . Voir figure 1.2.

Cette démarche d'explicitier y en fonction de x ne sera pas utilisée dans la suite : elle est simple dans \mathbb{R}^2 , mais peu agréable dans \mathbb{R}^n pour $n \geq 3$. On préférera la technique des multiplicateurs de Lagrange. \blacksquare

1.5 Le multiplicateur de Lagrange

Soit un point \vec{x}_0 où $f|_P$ atteint un extremum : à partir de ce point $\vec{x}_0 \in P$, pour rester dans P on ne peut se déplacer que dans les directions $\vec{v} \in \vec{P}$ tangentes à P , i.e. considérer les points $\vec{x}_0 + h\vec{v}$

pour $h \in \mathbb{R}$ et $\vec{v} \in \vec{P}$. L'extremum de $f|_P$ en \vec{x}_0 dans P est donc caractérisé par :

$$\forall \vec{v} \in \vec{P}, \quad \lim_{h \rightarrow 0} \frac{f|_P(\vec{x}_0 + h\vec{v}) - f|_P(\vec{x}_0)}{h} = 0 \quad (1.19)$$

(les dérivées directionnelles sont nulles pour tout $\vec{v} \in \vec{P}$), i.e. (cf. développement limité au premier ordre (1.3)) :

$$(\vec{\nabla} f(\vec{x}_0), \vec{v})_{\mathbb{R}^n} = 0, \quad \forall \vec{v} \in \vec{P}. \quad (1.20)$$

En d'autres termes $\vec{\nabla} f(\vec{x}_0)$ est orthogonal à \vec{P} (dans \mathbb{R}^n) :

$$\vec{\nabla} f(\vec{x}_0) \perp \vec{P}, \quad (1.21)$$

soit avec (1.15) :

$$\vec{\nabla} f(\vec{x}_0) \in \text{Vect}\{\vec{\nabla} g(\vec{x})\}. \quad (1.22)$$

Donc il existe $\lambda \in \mathbb{R}$, qui est appelé multiplicateur de Lagrange (de la contrainte $g = 0$), tel que :

$$\vec{\nabla} f(\vec{x}_0) + \lambda \vec{\nabla} g(\vec{x}_0) = \vec{0}. \quad (1.23)$$

Remarque 1.7 Le choix d'écrire $\vec{\nabla} f(\vec{x}_0) + \lambda \vec{\nabla} g(\vec{x}_0) = \vec{0}$, i.e. $\vec{\nabla} f(\vec{x}_0) = -\lambda \vec{\nabla} g(\vec{x}_0)$, i.e. le choix du signe devant λ est une convention de signe traditionnelle qui permettra d'avoir un lagrangien $\mathcal{L}(\vec{x}, \lambda) = f(\vec{x}) + \lambda g(\vec{x})$ où n'apparaîtra que des signes + comme dans (1.23). ■

On écrit souvent (1.22) sous la forme :

$$\vec{\nabla} f(\vec{x}_0) \parallel \vec{\nabla} g(\vec{x}_0), \quad (1.24)$$

i.e., $\vec{\nabla} f(\vec{x}_0)$ est parallèle à $\vec{\nabla} g(\vec{x}_0)$ en un extremum \vec{x}_0 de $f|_P$.

Remarque 1.8 P a été donné à l'aide d'une relation implicite " $g(\vec{x}) = 0$ " : cela permet d'avoir f et g définies sur le même domaine de définition, ici \mathbb{R}^n . Et f et $g : \mathbb{R}^n \rightarrow \mathbb{R}$ ayant le même domaine de définition, on peut les "comparer" à l'aide de (1.23).

C'est un des gros intérêts de considérer la surface P à l'aide de l'équation implicite $g(\vec{x}) = 0$, et non à l'aide d'une fonction explicite de type $x_n = h(x_1, \dots, x_{n-1})$: ici h et f n'ont pas le même domaine de définition, et on ne peut pas les comparer directement, voir exemple 1.6, à moins de passer par la fonction $g(x_1, \dots, x_n) = x_n - h(x_1, \dots, x_{n-1})$...

De plus, une équation explicite de type $x_n = h(x_1, \dots, x_{n-1})$ ne permet pas de considérer les hyperplans dans \mathbb{R}^n comme " $x_1 = 0$ ". ■

1.6 Le problème à résoudre

On vient d'introduire une nouvelle inconnue λ . Le nouveau problème à résoudre est donc : trouver $\vec{x} \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$ tels que :

$$\begin{cases} \vec{\nabla} f(\vec{x}) + \lambda \vec{\nabla} g(\vec{x}) = \vec{0}, \\ g(\vec{x}) = 0, \end{cases} \quad (1.25)$$

soit :

$$\begin{cases} \frac{\partial f}{\partial x_1}(\vec{x}) + \lambda \frac{\partial g}{\partial x_1}(\vec{x}) = 0, \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x}) + \lambda \frac{\partial g}{\partial x_n}(\vec{x}) = 0, \\ g(\vec{x}) = 0, \end{cases} \quad (1.26)$$

système de $n+1$ équations à $n+1$ inconnues (les x_i et λ).

Exemple 1.9 $f(x, y) = 1 + \alpha(x - x_m)^2 + \beta(y - y_m)^2$ et $g(x, y) = ax + by + c$, comme à l'exemple 1.6. On a $\vec{\nabla}f(x, y) = \begin{pmatrix} 2\alpha(x - x_m) \\ 2\beta(y - y_m) \end{pmatrix}$ et $\vec{\nabla}g(x, y) = \begin{pmatrix} a \\ b \end{pmatrix}$. Et $\vec{\nabla}f(\vec{x}) // \vec{\nabla}g(\vec{x})$ s'énonce : il existe un $\lambda \in \mathbb{R}$ t.q. $\vec{\nabla}f(\vec{x}) + \lambda\vec{\nabla}g(\vec{x}) = \vec{0}$, avec λ coefficient appelé le multiplicateur de Lagrange. Et le système (1.25) à résoudre s'énonce : trouver $(\vec{x}, \lambda) \in \mathbb{R}^2 \times \mathbb{R}$ t.q. :

$$\begin{cases} \vec{\nabla}f(\vec{x}) + \lambda\vec{\nabla}g(\vec{x}) = \vec{0}, \\ g(\vec{x}) = 0, \end{cases} \quad \text{i.e.} \quad \begin{cases} 2\alpha x + \lambda a = 2\alpha x_m, \\ 2\beta y + \lambda b = 2\beta y_m, \\ ax + by = c, \end{cases} \quad (1.27)$$

système de trois équations à trois inconnues x, y, λ . ▀

Remarque 1.10 La démarche d'introduction du multiplicateur de Lagrange semble compliquée dans l'exemple ci-dessus (dans \mathbb{R}^2). Mais dans \mathbb{R}^n avec n "grand", et avec plus d'une contrainte, cette démarche prendra tout son sens : cela simplifiera grandement la résolution. De plus, le multiplicateur de Lagrange (sa valeur) admet souvent une interprétation utile : c'est la valeur de la "pression" nécessaire au maintien de la contrainte. ▀

Exemple 1.11 Pour $\mathbb{R}^n = \mathbb{R}^2$, on peut résoudre le problème (1.17) par exemple à l'aide de la méthode de substitution : dans le cas où $\frac{\partial g}{\partial x_2}(x_1, x_2) \neq 0$, le lien $g(x_1, x_2) = 0$ est explicité sous la forme $x_2 = x_2(x_1)$, et on pose :

$$\psi(x_1) = f(x_1, x_2(x_1)).$$

Et l'extremum de f dans P est l'extremum x_1 de ψ , et on récupère ensuite $x_2 = x_2(x_1)$. Et $\psi'(x_1) = 0$ s'écrit :

$$\frac{\partial f}{\partial x_1}(x_1, x_2(x_1)) + \frac{\partial f}{\partial x_2}(x_1, x_2(x_1)) \frac{\partial x_2}{\partial x_1}(x_1) = 0, \quad (1.28)$$

équation dont la résolution donne x_1 . Et comme $g(x_1, x_2(x_1)) = 0$, on obtient par dérivation :

$$\frac{\partial g}{\partial x_1}(x_1, x_2(x_1)) + \frac{\partial g}{\partial x_2}(x_1, x_2(x_1)) \frac{\partial x_2}{\partial x_1}(x_1) = 0, \quad (1.29)$$

et en comparant (1.28) et (1.29) on retrouve en particulier $\vec{\nabla}f // \vec{\nabla}g$. ▀

1.7 Le lagrangien

Pour retrouver la caractérisation simple et essentielle d'un extremum d'une fonction \mathcal{L} , à savoir " $\vec{\nabla}\mathcal{L} = \vec{0}$ ", on pose :

$$\mathcal{L}(\vec{x}, \lambda) = f(\vec{x}) + \lambda g(\vec{x}). \quad (1.30)$$

\mathcal{L} est appelé le lagrangien (relatif à f sous la contrainte $g = 0$). Supposant $f, g \in C^1(\mathbb{R}^n; \mathbb{R})$, on a $\mathcal{L} \in C^1(\mathbb{R}^{n+1}; \mathbb{R})$ et :

$$\frac{\partial \mathcal{L}}{\partial \vec{x}}(\vec{x}, \lambda) = \vec{\nabla}f(\vec{x}) + \lambda\vec{\nabla}g(\vec{x}) \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \lambda}(\vec{x}, \lambda) = g(\vec{x}). \quad (1.31)$$

Ainsi :

$$\vec{\nabla}\mathcal{L} = \vec{0} \quad \iff \quad (1.25), \quad (1.32)$$

qui est un problème de recherche d'extremum sans contrainte : on a ainsi transformé notre problème contraint (pour f) en un problème sans contrainte (pour \mathcal{L}).

1.8 Forme matricielle usuelle

Soit $A = [a_{ij}]$ une matrice $n \times n$ symétrique définie positive, $\vec{\alpha} \in \mathbb{R}^n$ et $c_f \in \mathbb{R}$, et soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ la fonction quadratique à valeurs réelles :

$$f(\vec{x}) = \frac{1}{2}\vec{x}^T . A . \vec{x} + \vec{\alpha}^T . \vec{x} + c_f. \quad (1.33)$$

Comme on s'intéressera aux points \vec{x} où f est minimum, on pourra prendre $c_f = 0$. En effet, le point \vec{x} où le minimum sera atteint ne dépendra pas de cette constante c_f , puisque lors des

dérivations cette constante disparaît. Seule la valeur de f (la “hauteur” de f) en ce minimum dépend de c_f .

Comme A est symétrique, on a :

$$\vec{\nabla} f(\vec{x}) = A.\vec{x} + \vec{\alpha}, \quad (1.34)$$

En effet :

$$\begin{aligned} f(\vec{x} + h\vec{v}) - f(\vec{x}) &= \frac{1}{2}(\vec{x} + h\vec{v})^T . A . \vec{x}(\vec{x} + h\vec{v}) + \vec{\alpha}^T . (\vec{x} + h\vec{v}) + \beta - \frac{1}{2}\vec{x}^T . A . \vec{x} - \vec{\alpha}^T . \vec{x} - \beta \\ &= \frac{1}{2}h\vec{v}^T . A . \vec{x} + \frac{1}{2}h\vec{x}^T . A . \vec{v} + \frac{1}{2}h^2\vec{v}^T . A . \vec{v} + \vec{\alpha}^T . h\vec{v} \\ &= h\vec{v}^T . A . \vec{x} + \vec{\alpha}^T . h\vec{v} + \frac{1}{2}h^2\vec{v}^T . A . \vec{v} \end{aligned}$$

car $A^T = A$. D'où $f(\vec{x} + h\vec{v}) - f(\vec{x}) = h(A.\vec{x} + \vec{\alpha}, \vec{v})_{\mathbb{R}^n} + o(h)$. D'où (1.34) (cf. (1.8)).

Et donc les équations (1.25) se mettent sous la forme, avec (1.11) et $B = \vec{b}^T$ (la contrainte $(\vec{b}, \vec{x})_{\mathbb{R}^n} = c = \vec{b}^T . \vec{x}$) :

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \vec{x} \\ \lambda \end{pmatrix} = \begin{pmatrix} -\vec{\alpha} \\ c \end{pmatrix}.$$

C'est le système sous forme matricielle générique à résoudre.

2 p contraintes : les p multiplicateurs de Lagrange

On se donne deux entiers n et p , avec $n \geq 2$ (dimension de l'espace) et $p \geq 1$ (nombre des contraintes), et on suppose $n > p$.

2.1 Caractérisation de l'intersection de p hyperplans dans \mathbb{R}^n

On se donne p vecteurs non nuls $\vec{b}_i \in \mathbb{R}^n$ et p réels c_i , et, pour $i = 1, \dots, p$, on note $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ la fonction :

$$g_i(\vec{x}) = (\vec{b}_i, \vec{x})_{\mathbb{R}^n} - c_i. \quad (2.1)$$

Donc les g_i sont affines (non constantes car $\vec{b}_i \neq \vec{0}$), avec $\vec{b}_i = \vec{\nabla} g_i(\vec{x})$ et $c_i = -g_i(\vec{0})$. On supposera de plus :

$$\text{les vecteurs } \vec{b}_1, \dots, \vec{b}_p \text{ sont indépendants dans } \mathbb{R}^n, \quad (2.2)$$

sinon les contraintes $g_i(\vec{x}) = 0$ seraient incompatibles ou redondantes.

Et on suppose $p < n$: on exclut le cas $p = n$, car dans ce cas il n'y a qu'un seul point \vec{x} tel que $g_i(\vec{x}) = 0$ pour tout i : le problème s'arrêterait ici.

Notons, pour $i = 1, \dots, p$:

$$P_i = \{\vec{x} \in \mathbb{R}^n : g_i(\vec{x}) = 0\} = \{\vec{x} \in \mathbb{R}^n : (\vec{x}, \vec{b}_i)_{\mathbb{R}^n} = c_i\}. \quad (2.3)$$

Les P_i sont des hyperplans affines de \mathbb{R}^n , et on note :

$$P = \bigcap_{i=1}^p P_i, \quad (2.4)$$

l'intersection des hyperplans affines.

Les hyperplans vectoriels associés aux P_i sont les :

$$\vec{P}_i = \{\vec{v} \in \mathbb{R}^n : (\vec{v}, \vec{b}_i)_{\mathbb{R}^n} = 0\} = (\text{Vect}\{\vec{b}_i\})^\perp, \quad (2.5)$$

hyperplans vectoriels dont des vecteurs normaux sont les $\vec{b}_i = \vec{\nabla} g_i(\vec{x})$. Et on note \vec{P} le sous-espace vectoriel associé à P :

$$\vec{P} = \bigcap_{i=1}^p \vec{P}_i = \bigcap_{i=1}^p (\text{Vect}\{\vec{b}_i\})^\perp. \quad (2.6)$$

\vec{P} est l'intersection des hyperplans vectoriels \vec{P}_i .

Proposition 2.1 Les vecteurs $\vec{b}_1, \dots, \vec{b}_p$ étant indépendants dans \mathbb{R}^n , on a les équivalences :

$$\begin{aligned} \vec{v} \in \vec{P} &\iff \forall i = 1, \dots, p, \quad \vec{v} \in \vec{P}_i, \\ &\iff \forall i = 1, \dots, p, \quad \vec{v} \perp \vec{b}_i, \\ &\iff \forall i = 1, \dots, p, \quad (\vec{v}, \vec{b}_i)_{\mathbb{R}^n} = 0, \\ &\iff \forall (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p : (\vec{v}, \lambda_1 \vec{b}_1 + \dots + \lambda_p \vec{b}_p)_{\mathbb{R}^n} = 0. \end{aligned} \quad (2.7)$$

Et donc \vec{P} donné en (2.6) est de dimension $n-p$, avec $(\vec{b}_1, \dots, \vec{b}_p)$ qui est une base de \vec{P}^\perp :

$$\vec{P}^\perp = \text{Vect}\{\vec{b}_1, \dots, \vec{b}_p\}. \quad (2.8)$$

Et si $\vec{x}_P \in P$ on a :

$$P = \vec{x}_P + \vec{P}. \quad (2.9)$$

Preuve. Les trois premières équivalences sont immédiates, et la quatrième l'est également grâce à la bilinéarité du produit scalaire.

Donc, avec la deuxième équivalence, $\vec{P} = \bigcap_{i=1}^p (\text{Vect}\{\vec{b}_i\})^\perp = (\text{Vect}\{\vec{b}_1, \dots, \vec{b}_p\})^\perp$. D'où \vec{P}^\perp est donné par (2.8), et est de dimension p car les \vec{b}_i sont indépendants, d'où \vec{P} est de dimension $n-p$.

Si $\vec{x}_P \in P$, on a $g(\vec{x}_P) = \vec{c}$, et si $\vec{v} \in \vec{P}$, alors $g_i(\vec{x}_P + \vec{v}) = (\vec{b}_i, \vec{x}_P + \vec{v})_{\mathbb{R}^n} = (\vec{b}_i, \vec{x}_P)_{\mathbb{R}^n} + (\vec{b}_i, \vec{v})_{\mathbb{R}^n} = \vec{c} + \vec{0} = \vec{c}$, ce pour tout i , et donc $\vec{x}_P + \vec{v} \in P$. Réciproquement, si $\vec{x} \in P$ et si $\vec{x}_P \in \vec{P}$ alors $\vec{v} = \vec{x} - \vec{x}_P$ vérifie $(\vec{v}, \vec{b}_i)_{\mathbb{R}^n} = (\vec{b}_i, \vec{x})_{\mathbb{R}^n} - (\vec{b}_i, \vec{x}_P)_{\mathbb{R}^n} = \vec{c} - \vec{c} = \vec{0}$, ce pour tout i , et donc $\vec{v} \in \vec{P}$. D'où (2.9). ■

2.2 Le problème contraint

On note $\vec{g} = \begin{pmatrix} g_1 \\ \vdots \\ g_p \end{pmatrix} : \vec{x} \in \mathbb{R}^n \rightarrow \vec{g}(\vec{x}) = \begin{pmatrix} g_1(\vec{x}) \\ \vdots \\ g_p(\vec{x}) \end{pmatrix} \in \mathbb{R}^p$ la fonction à valeurs vectorielles composées des g_i .

On impose la contrainte :

$$\vec{g}(\vec{x}) = \vec{0}. \quad (2.10)$$

On cherche un extremum de f dans l'espace des contraintes :

$$\left\{ \begin{array}{l} \text{Trouver } \vec{x}_0 \in \mathbb{R}^n \text{ t.q.} \\ f(\vec{x}_0) = \underset{\vec{x} \in \mathbb{R}^n : \vec{g}(\vec{x}) = \vec{0}}{\text{extremum}} f(\vec{x}) = \underset{\vec{x} \in P}{\text{extremum}} f(\vec{x}) \end{array} \right. \quad (2.11)$$

où on a utilisé (2.4). Soit encore, notant $f|_P : \vec{x} \in P \rightarrow f|_P(\vec{x}) = f(\vec{x})$ la restriction de f à P :

$$\left\{ \begin{array}{l} \text{Trouver } \vec{x}_0 \in P \text{ t.q.} \\ f|_P(\vec{x}_0) = \underset{\vec{x} \in P}{\text{extremum}} f|_P(\vec{x}). \end{array} \right. \quad (2.12)$$

2.3 Les p multiplicateurs de Lagrange

Corollaire 2.2 Soit $\vec{x}_0 \in \mathbb{R}^n$. La fonction $f|_P$ est extrémale en \vec{x}_0 ssi :

$$\forall \vec{v} \in \vec{P}, \quad \vec{\nabla} f|_P(\vec{x}_0) \cdot \vec{v} = 0, \quad (2.13)$$

i.e. ssi la fonction f vérifie :

$$\forall \vec{v} \in \vec{P}, \quad \vec{\nabla} f(\vec{x}_0) \cdot \vec{v} = 0, \quad (2.14)$$

i.e. ssi :

$$\vec{\nabla} f(\vec{x}_0) \in \vec{P}^\perp, \quad (2.15)$$

i.e. ssi, avec (2.8) :

$$\exists (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p, \quad \vec{\nabla} f(\vec{x}_0) + \sum_{i=1}^p \lambda_i \vec{b}_i = 0. \quad (2.16)$$

Preuve. Comme $f|_P$ est définie sur P , que $\vec{x}_0 \in P$, et donc que $P = \vec{x}_0 + \vec{P}$, cf. (2.9), on a $f|_P(\vec{x})$ est définie pour tout $\vec{x} = \vec{x}_0 + \vec{v}$, où $\vec{v} \in \vec{P}$. Et, pour tout $\vec{v} \in \vec{P}$ on a au voisinage de $h = 0$:

$$f|_P(\vec{x}_0 + h\vec{v}) - f|_P(\vec{x}_0) = h \vec{\nabla} f|_P(\vec{x}_0) \cdot \vec{v} + o(h).$$

Donc $f|_P$ est extrémale en \vec{x}_0 ssi (2.13). D'où (2.14), (2.15), puis (2.16) avec (2.8). \blacksquare

Définition 2.3 Les λ_i sont les multiplicateurs de Lagrange (associés à f sous les contraintes g_i).

Les λ_i sont, au signe près, les composantes de $\vec{\nabla} f(\vec{x}_0)$ dans la base $(\vec{b}_i)_{i=1, \dots, p}$ de \vec{P}^\perp .

2.4 Le problème contraint à résoudre

Le nouveau problème à résoudre est donc, avec (2.16) et (2.10) : trouver $(\vec{x}, \vec{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p$ t.q. :

$$\begin{cases} \vec{\nabla} f(\vec{x}) + \sum_{i=1}^p \lambda_i \vec{\nabla} g_i(\vec{x}) = \vec{0} & (n \text{ équations}), \\ \vec{g}(\vec{x}) = 0 & (p \text{ équations}), \end{cases} \quad (2.17)$$

système de $n + p$ équations à $n + p$ inconnues $(\vec{x}, \vec{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p$.

Exemple 2.4 Soit $\mathbb{R}^n = \mathbb{R}^3$ et $f(x, y, z) = x^2 + y^2 + z^2 + 1$. Donc $\vec{\nabla} f(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}$. Soit $g_1(x, y, z) = x + y + z - c_1$ et $g_2(x, y, z) = x - y + z - c_2$. Ici $\vec{\nabla} g_1(x, y, z) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \vec{b}_1$ et $\vec{\nabla} g_2(x, y, z) = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \vec{b}_2$, et les vecteurs \vec{b}_1 et \vec{b}_2 sont indépendants.

Et (2.17)₁ donne le système de 3 équations : $\begin{cases} 2x_0 + \lambda_1 + \lambda_2 = 0 \\ 2y_0 + \lambda_1 - \lambda_2 = 0 \\ 2z_0 + \lambda_1 + \lambda_2 = 0 \end{cases}$, Et (2.17)₂ (la contrainte

vectorielle) donne le système de 2 équations : $\begin{cases} x_0 + y_0 + z_0 = c_1 \\ x_0 - y_0 + z_0 = c_2 \end{cases}$. Donc 5 équations et 5 inconnues, ici avec une unique solution $(x_0, y_0, z_0, \lambda_1, \lambda_2) \in \mathbb{R}^3 \times \mathbb{R}^2$. Et le point (x_0, y_0, z_0) est le point en lequel $f|_P$ est minimum.

Calculs pour informations :

les plans $P_1 = \{g_1 = 0\}$ et $P_2 = \{g_2 = 0\}$ dans \mathbb{R}^3 ont pour intersection une droite dans \mathbb{R}^3 :

considérant z comme paramètre, on a $\begin{cases} x + y = c_1 - z, \\ x - y = c_2 - z, \end{cases}$ donc $\begin{cases} x = \frac{c_1 + c_2}{2} - z, \\ y = \frac{c_1 - c_2}{2}, \end{cases}$ équation d'une

droite passant par le point $\vec{x}_P = (\frac{c_1 + c_2}{2}, \frac{c_1 - c_2}{2}, 0)$, droite vectorielle associés $\begin{cases} x = -z, \\ y = 0, \end{cases}$ de vecteur

directeur $\vec{v} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ (ou encore $\vec{b}_1 \wedge \vec{b}_2 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$). (On a bien $\vec{x} = \vec{x}_P + \alpha \vec{v} = (\frac{c_1 + c_2}{2} - \alpha, \frac{c_1 - c_2}{2}, \alpha)$ qui vérifie $g_1(\vec{x}) = 0 = g_2(\vec{x})$ pour tout α). \blacksquare

2.5 Le lagrangien

Comme au paragraphe (1.7), pour retrouver la caractérisation simple et essentielle d'un extremum, à savoir en un extremum de \mathcal{L} on a " $\vec{\nabla} \mathcal{L} = \vec{0}$ ", on pose :

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) + (\vec{\lambda}, \vec{g}(\vec{x}))_{\mathbb{R}^p} = f(\vec{x}) + \sum_{i=1}^p \lambda_i g_i(\vec{x}) = f(\vec{x}) + \sum_{i=1}^p \lambda_i ((\vec{b}_i, \vec{x})_{\mathbb{R}^n} - c_i), \quad (2.18)$$

$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}$ est appelé le lagrangien. Supposant f et g C^1 on a :

$$\frac{\partial \mathcal{L}}{\partial \vec{x}}(\vec{x}, \lambda) = \vec{\nabla} f(\vec{x}) + \sum_{i=1}^p \lambda_i \vec{b}_i = \vec{\nabla} f(\vec{x}) + \sum_{i=1}^p \lambda_i \vec{\nabla} g_i(\vec{x}), \quad \text{et} \quad \frac{\partial \mathcal{L}}{\partial \vec{\lambda}}(\vec{x}, \vec{\lambda}) = \vec{g}(\vec{x}), \quad (2.19)$$

et $\vec{\nabla} \mathcal{L} = 0$ donne bien (2.17).

2.6 Matrice B

On note B la matrice et \vec{c} le vecteur donnés par :

$$B = \begin{pmatrix} [\vec{b}_1^T] \\ \vdots \\ [\vec{b}_p^T] \end{pmatrix} \quad \text{et} \quad \vec{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_p \end{pmatrix}. \quad (2.20)$$

($B = [dg(\vec{x})]$ est la matrice jacobienne de g en \vec{x} , i.e. la matrice de la différentielle $dg(\vec{x})$ de g en \vec{x} exprimée dans la base canonique.) Ainsi :

$$\vec{g}(\vec{x}) = B \cdot \vec{x} - \vec{c} \quad \left(= \begin{pmatrix} g_1(\vec{x}) = \vec{b}_1^T \cdot \vec{x} + c_1 \\ \vdots \\ g_p(\vec{x}) = \vec{b}_p^T \cdot \vec{x} + c_p \end{pmatrix} \right). \quad (2.21)$$

Ainsi :

$$\vec{x} \in P \quad \iff \quad B \cdot \vec{x} = \vec{c}. \quad (2.22)$$

Et :

$$\vec{x} \in \vec{P} \quad \iff \quad B \cdot \vec{x} = \vec{0}. \quad (2.23)$$

Exemple 2.5 Pour $\mathbb{R}^n = \mathbb{R}^3$ et $\mathbb{R}^p = \mathbb{R}^2$ on prend $B_j^i = \frac{\partial g_i}{\partial x^j}(\vec{x})$:

$$\begin{cases} g_1(x, y, z) = B_1^1 x + B_2^1 y + B_3^1 z - c_1 = \vec{b}_1^T \cdot \vec{x} - c_1, \\ g_2(x, y, z) = B_1^2 x + B_2^2 y + B_3^2 z - c_2 = \vec{b}_2^T \cdot \vec{x} - c_2, \end{cases}$$

où donc $\vec{b}_1^T = (B_1^1, B_2^1, B_3^1)$, $\vec{b}_2^T = (B_1^2, B_2^2, B_3^2)$, et \vec{b}_1 et \vec{b}_2 supposés indépendants. Ici on a :

$$\vec{g}(\vec{x}) = B \cdot \vec{x} - \vec{c} \quad \text{où} \quad B = [B_j^i] = \begin{pmatrix} B_1^1 & B_2^1 & B_3^1 \\ B_1^2 & B_2^2 & B_3^2 \end{pmatrix} = \begin{pmatrix} [\vec{b}_1^T] \\ [\vec{b}_2^T] \end{pmatrix}.$$

Et $P = \{\vec{x} : \vec{g}(\vec{x}) = \vec{0}\}$ est l'intersection des deux plans affines non parallèles $g_1(\vec{x}) = 0$ et $g_2(\vec{x}) = 0$ et est donc une droite affine passant par un point \vec{x}_c tel que $B \cdot \vec{x}_c = -\vec{c} \in \mathbb{R}^2$, droite de vecteur directeur $\vec{b}_1 \wedge \vec{b}_2$ (vecteur à la fois perpendiculaire à \vec{b}_1 et à \vec{b}_2). Et B est la matrice jacobienne $[d\vec{g}(\vec{x})] = [\frac{\partial g_i}{\partial x_j}(\vec{x})] = B$ de \vec{g} en \vec{x} , indépendante de \vec{x} puisque \vec{g} est affine. \blacksquare

2.7 Forme matricielle usuelle

On reprend la forme quadratique f donnée en (1.33) :

$$f(\vec{x}) = \frac{1}{2} \vec{x}^T \cdot A \cdot \vec{x} + \vec{\alpha}^T \cdot \vec{x} + c_f. \quad (2.24)$$

A étant une matrice symétrique, on a $\vec{\nabla} f(\vec{x}) = A \cdot \vec{x} + \vec{\alpha}$.

Avec la contrainte $\vec{g}(\vec{x}) = \vec{0}$ qui s'écrit $B \cdot \vec{x} = \vec{c}$, et avec $\sum_{i=1}^p \lambda_i \vec{\nabla} g_i(\vec{x}) = \sum_{i=1}^p \lambda_i \vec{b}_i = B^T \cdot \vec{\lambda}$ (combinaison linéaires des colonnes \vec{b}_i de B^T), le problème matriciel (2.17) s'écrit : trouver $(\vec{x}, \vec{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p$ t.q. :

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \cdot \begin{pmatrix} \vec{x} \\ \vec{\lambda} \end{pmatrix} = \begin{pmatrix} -\vec{\alpha} \\ \vec{c} \end{pmatrix}. \quad (2.25)$$

C'est l'équation "prototype".

2.8 n valeurs propres positives et p valeurs propres négatives

Proposition 2.6 Avec A matrice $n \times n$ inversible symétrique et B matrice $p \times n$, on a :

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ B.A^{-1} & I \end{pmatrix} \cdot \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \cdot \begin{pmatrix} I & A^{-1}.B^T \\ 0 & I \end{pmatrix}, \quad (2.26)$$

où :

$$S = -B.A^{-1}.B^T. \quad (2.27)$$

Et si $\text{rang}(B) = p < n$ et A symétrique définie positive, alors S est une matrice définie négative. Dans ce cas $\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}$ est une matrice inversible qui a n valeurs propres strictement positives et p valeurs propres strictement négatives.

Preuve. Soit $N = \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}$, $P = \begin{pmatrix} I & 0 \\ B.A^{-1} & I \end{pmatrix}$ et $M = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix}$. On a $P^T = \begin{pmatrix} I & A^{-1}.B^T \\ 0 & I \end{pmatrix}$, car A symétrique, d'où $M.P^T = \begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix}$, d'où $P.M.P^T = \begin{pmatrix} A & B^T \\ B & B.A^{-1}.B^T + S \end{pmatrix}$. D'où :

$$N = P.M.P^T \quad \text{quand} \quad S = -B.A^{-1}.B^T,$$

soit (2.26). Et P est inversible (d'inverse $\begin{pmatrix} I & 0 \\ B.A^{-1} & I \end{pmatrix}^{-1} = \begin{pmatrix} I & 0 \\ -B.A^{-1} & I \end{pmatrix}$). Donc M et N sont congruentes, donc ont même nombre de valeurs propres nulles, positives et négatives (voir annexe B théorème d'inertie de Sylvester).

Cas A matrice $n \times n$ symétrique définie positive, donc A a n valeurs propres strictement positives. Donc A^{-1} a n valeurs propres strictement positives (ce sont les inverses de celles de A).

Si de plus B est de rang p , alors $-S = B.A^{-1}.B^T$ est une matrice $p \times p$ symétrique (trivial) qui est de plus définie positive : en effet :

$$(B.A^{-1}.B^T.\vec{\lambda}, \vec{\lambda})_{\mathbb{R}^p} = (A^{-1}.B^T.\vec{\lambda}, B^T.\vec{\lambda})_{\mathbb{R}^n} \geq \gamma \|B^T.\vec{\lambda}\|_{\mathbb{R}^n}^2 > 0,$$

où γ est la plus petite valeur propre de A^{-1} , le > 0 car B matrice $p \times n$ étant de rang p on a $\text{Ker} B^T = \{0\}$ (voir annexe).

Donc $S = -B.A^{-1}.B^T$ est symétrique définie négative, donc a p valeurs propres strictement négatives. Donc M a n valeurs propres strictement positives et p valeurs propres strictement négatives. Donc N aussi par congruence (théorème d'inertie de Sylvester). \blacksquare

3 Problèmes primal et dual

3.1 Problème primal

C'est le problème avec contrainte (2.11) :

$$\min_{\vec{x}: \vec{g}(\vec{x}) = \vec{0}} f(\vec{x}). \quad (3.1)$$

3.2 Problème dual et matrice d'Uzawa

Une méthode de résolution possible de (3.1) à l'aide du lagrangien $\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) + (\vec{\lambda}, \vec{g}(\vec{x}))_{\mathbb{R}^p}$ est séquentielle : d'abord :

$$\begin{cases} \text{à } \vec{\lambda} \text{ fixé, trouver } \vec{x} = \vec{x}(\vec{\lambda}) \text{ t.q. :} \\ \frac{\partial \mathcal{L}}{\partial \vec{x}}(\vec{x}, \vec{\lambda}) = 0. \end{cases} \quad (3.2)$$

C'est l'équation (2.17)₁. Donc à $\vec{\lambda}$ fixé, trouver $\vec{x} = \vec{x}(\vec{\lambda})$ t.q. :

$$\vec{\nabla} f(\vec{x}) + \sum_{i=1}^p \lambda_i \vec{\nabla} g_i(\vec{x}) = \vec{0}. \quad (3.3)$$

On détermine ainsi une solution $\vec{x} = \vec{x}(\vec{\lambda})$, et on définit :

$$G(\vec{\lambda}) \stackrel{\text{déf}}{=} \mathcal{L}(\vec{x}(\vec{\lambda}), \vec{\lambda}), \quad (3.4)$$

appelée fonction duale (la fonction duale est une fonction de $\vec{\lambda}$ alors que la fonction primale est une fonction de \vec{x}).

On suppose f quadratique définie positive :

$$f(\vec{x}) = \frac{1}{2} \vec{x}^T . A . \vec{x} + \vec{\alpha}^T . \vec{x} + c_f, \quad (3.5)$$

avec A matrice $n \times n$ symétrique définie positive, $\vec{\alpha} \in \mathbb{R}^n$ et $c_f \in \mathbb{R}$. On suppose \vec{g} affine :

$$\vec{g}(\vec{x}) = B . \vec{x} - \vec{c}, \quad (3.6)$$

avec B matrice $p \times n$ de rang p et $\vec{c} \in \mathbb{R}^p$. Le lagrangien est donc :

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = \frac{1}{2} \vec{x}^T . A . \vec{x} + \vec{\alpha}^T . \vec{x} + c_f + \vec{\lambda}^T . (B . \vec{x} - \vec{c}). \quad (3.7)$$

Proposition 3.1 et définition de la matrice d'Uzawa. Avec \mathcal{L} définie en (3.7), le vecteur $\vec{x}(\vec{\lambda})$ réalisant le minimum $\inf_{\vec{x} \in \mathbb{R}^n} \mathcal{L}(\vec{x}, \vec{\lambda}) = \mathcal{L}(\vec{x}(\vec{\lambda}), \vec{\lambda})$, cf. (3.2) et (3.3), vaut :

$$\vec{x}(\vec{\lambda}) = -A^{-1} . (B^T . \vec{\lambda} + \vec{\alpha}). \quad (3.8)$$

Et la fonction duale G , définie en (3.4), est la fonction quadratique symétrique définie négative donnée par, pour $\vec{\lambda} \in \mathbb{R}^p$:

$$G(\vec{\lambda}) = -\frac{1}{2} \vec{\lambda}^T . U . \vec{\lambda} - \vec{\lambda}^T . \vec{\beta} - \gamma \quad (= \mathcal{L}(\vec{x}(\vec{\lambda}), \vec{\lambda})), \quad (3.9)$$

où :

$$U \stackrel{\text{déf}}{=} B . A^{-1} . B^T \quad \text{et} \quad \vec{\beta} = \vec{c} + B . A^{-1} . \vec{\alpha} \quad \text{et} \quad \gamma = \frac{1}{2} \vec{\alpha}^T . A^{-1} . \vec{\alpha}. \quad (3.10)$$

Et la matrice $p \times p$:

$$U = B . A^{-1} . B^T \quad (\text{matrice d'Uzawa}). \quad (3.11)$$

est symétrique définie positive, dite matrice d'Uzawa. D'où G est quadratique strictement concave et admet un maximum en :

$$\vec{\lambda}_0 = U^{-1} . \vec{\beta}, \quad (3.12)$$

avec :

$$G(\vec{\lambda}_0) = \frac{1}{2} \vec{\beta}^T . U^{-1} . \vec{\beta} + \gamma. \quad (3.13)$$

(En particulier $G(\vec{\lambda}_0) \geq \gamma$.)

On pose $\vec{x}_0 = \vec{x}(\vec{\lambda}_0)$ calculée avec (3.8), et \vec{x}_0 est la solution cherchée du problème primal (3.1), solution qui vérifie $G(\vec{\lambda}_0) = f(\vec{x}_0)$.

Preuve. Soit $\vec{\lambda}$ fixé, et $\mathcal{L}_{\vec{\lambda}}(\vec{x}) = \mathcal{L}(\vec{x}, \vec{\lambda})$. Ici $\mathcal{L}_{\vec{\lambda}}$ est quadratique strictement convexe continue (en \vec{x}) car f l'est et g est affine. Donc $\mathcal{L}_{\vec{\lambda}}$ a un unique extremum qui est un minimum.

Comme A est symétrique, l'égalité $\frac{\partial \mathcal{L}}{\partial \vec{x}}(\vec{x}, \vec{\lambda}) = \vec{0}$ donne :

$$A . \vec{x} + \vec{\alpha} + B^T . \vec{\lambda} = \vec{0}, \quad (3.14)$$

cf. (1.34), et donc le point $\vec{x} = \vec{x}(\vec{\lambda})$ réalisant le minimum est donné par :

$$\vec{x}(\vec{\lambda}) = -A^{-1} . (B^T . \vec{\lambda} + \vec{\alpha}) = -(A^{-1} . B^T . \vec{\lambda} + A^{-1} . \vec{\alpha}), \quad (3.15)$$

soit (3.8). Comme A est symétrique inversible, on a A^{-1} symétrique, et donc :

$$\begin{aligned}
G(\vec{\lambda}) &= \mathcal{L}(\vec{x}(\vec{\lambda}), \vec{\lambda}) \\
&= \frac{1}{2}(\vec{\lambda}^T . B . A^{-1} + \vec{\alpha}^T . A^{-1}) . A . (A^{-1} . B^T . \vec{\lambda} + A^{-1} . \vec{\alpha}) - \vec{\alpha}^T . (A^{-1} . B^T . \vec{\lambda} + A^{-1} . \vec{\alpha}) \\
&\quad + \vec{\lambda}^T . (-B . A^{-1} . B^T . \vec{\lambda} + B . A^{-1} . \vec{\alpha} - \vec{c}) \\
&= \frac{1}{2}\vec{\lambda}^T . B . A^{-1} . B^T . \vec{\lambda} + \vec{\lambda}^T . B . A^{-1} . \vec{\alpha} + \frac{1}{2}\vec{\alpha}^T . A^{-1} . \vec{\alpha} - \vec{\alpha}^T . A^{-1} . B^T . \vec{\lambda} - \vec{\alpha}^T . A^{-1} . \vec{\alpha} \\
&\quad - \vec{\lambda}^T . B . A^{-1} . B^T . \vec{\lambda} - \vec{\lambda}^T . B . A^{-1} . \vec{\alpha} - \vec{\lambda}^T . \vec{c} \\
&= -\frac{1}{2}\vec{\lambda}^T . (B . A^{-1} . B^T) . \vec{\lambda} + (-\vec{c}^T - \vec{\alpha}^T . A^{-1} . B^T) . \vec{\lambda} - \frac{1}{2}\vec{\alpha}^T . A^{-1} . \vec{\alpha},
\end{aligned}$$

i.e. (3.9).

Et la matrice $U = B . A^{-1} . B^T$ de taille $p \times p$ est symétrique définie positive : en effet, la symétrie est immédiate (car A^{-1} est symétrique), et pour tout $\vec{\mu} \in \mathbb{R}^p - \{0\}$ on a $B^T \vec{\mu} \neq \vec{0}$ (car B est de rang p donne B^T injective), et donc $(B^T \vec{\mu})^T A^{-1} (B^T \vec{\mu}) \neq 0$ car A^{-1} est symétrique définie positive (puisque A l'est).

Et donc G donnée par (3.9) est bien quadratique symétrique définie négative.

D'où G atteint son maximum pour $\vec{\lambda}_0$ donné par $G'(\vec{\lambda}_0) = 0$, i.e. pour $\vec{\lambda}_0$ donné par $-U . \vec{\lambda}_0 + \vec{\beta} = 0$, soit (3.12).

D'où $G(\vec{\lambda}_0) = -\frac{1}{2}(\vec{\beta}^T . U^{-1}) . U . (U^{-1} . \vec{\beta}) + (\vec{\beta}^T . U^{-1}) . \vec{\beta} + \gamma = \frac{1}{2}\vec{\beta}^T . U^{-1} . \vec{\beta} + \gamma$, soit (3.13).

Puis $\vec{x}_0 = \vec{x}(\vec{\lambda}_0)$ est solution du problème primal, voir paragraphe suivant. ▀

3.3 Problème d'inf-sup pour le lagrangien (point selle du lagrangien)

Proposition 3.2 et point selle. On considère le cas de la proposition 3.1. L'extremum $(\vec{x}_0, \vec{\lambda}_0)$ de \mathcal{L} est la solution du problème inf-sup :

$$\left\{ \begin{array}{l} \text{trouver } (\vec{x}_0, \vec{\lambda}_0) \in \mathbb{R}^n \times \mathbb{R}^p \text{ t.q. :} \\ \mathcal{L}(\vec{x}_0, \vec{\lambda}_0) = \sup_{\vec{\lambda} \in \mathbb{R}^p} \inf_{\vec{x} \in \mathbb{R}^n} \mathcal{L}(\vec{x}, \vec{\lambda}) = \inf_{\vec{x} \in \mathbb{R}^n} \sup_{\vec{\lambda} \in \mathbb{R}^p} \mathcal{L}(\vec{x}, \vec{\lambda}), \end{array} \right. \quad (3.16)$$

le point solution étant appelé le point selle de \mathcal{L} (minimum en \vec{x} et maximum en λ), donné par (2.25) :

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \cdot \begin{pmatrix} \vec{x} \\ \vec{\lambda} \end{pmatrix} = \begin{pmatrix} -\vec{\alpha} \\ \vec{c} \end{pmatrix}. \quad (3.17)$$

Preuve. Avec (3.9), on a :

$$G(\vec{\lambda}) = \inf_{\vec{x} \in \mathbb{R}^n} \mathcal{L}(\vec{x}, \vec{\lambda}) = \mathcal{L}(\vec{x}(\vec{\lambda}), \vec{\lambda}).$$

Et avec (3.12) et (3.13), G atteint son maximum en $\vec{\lambda}_0$:

$$\mathcal{L}(\vec{x}(\lambda_0), \vec{x}_0) = G(\lambda_0) = \sup_{\lambda \in \mathbb{R}^p} \mathcal{L}(\vec{x}(\vec{\lambda}), \vec{\lambda}) = \sup_{\lambda \in \mathbb{R}^p} \inf_{\vec{x} \in \mathbb{R}^n} \mathcal{L}(\vec{x}, \vec{\lambda}),$$

et on note $\vec{x}_0 = \vec{x}(\lambda_0)$. D'où l'existence et l'unicité de la solution.

Et dans ce cas, le sup de l'inf est égal à l'inf du sup, voir annexe C. ▀

4 Résolution

4.1 Pénalisation

But : résoudre (3.17) en découplant le système. On commence par "pénaliser".

Soit $\varepsilon \geq 0$ et on transforme (3.17) en :

$$\begin{pmatrix} A & B^T \\ B & -\varepsilon I \end{pmatrix} \cdot \begin{pmatrix} \vec{x}^\varepsilon \\ \vec{\lambda}^\varepsilon \end{pmatrix} = \begin{pmatrix} -\vec{\alpha} \\ \vec{c} \end{pmatrix}. \quad (4.1)$$

Proposition 4.1 *Si A est une matrice $n \times n$ symétrique définie positive, si B une matrice $p \times n$ de rang p , et si $\varepsilon > 0$, alors (4.1) a une unique solution qu'on notera $(\vec{x}^\varepsilon, \vec{\lambda}^\varepsilon)$.*

Et la solution $(\vec{x}^\varepsilon, \vec{\lambda}^\varepsilon)$ (solution de (4.1)) approche $(\vec{x}^0, \vec{\lambda}^0)$ (solution de (3.17)) :

$$\begin{cases} \|\vec{x}^0 - \vec{x}^\varepsilon\|_{\mathbb{R}^n} = O(\sqrt{\varepsilon}), \\ \|\vec{\lambda} - \vec{\lambda}^\varepsilon\|_{\mathbb{R}^p} = O(\sqrt{\varepsilon}). \end{cases} \quad (4.2)$$

Et le problème pénalisé (4.1) donne une solution à $\sqrt{\varepsilon}$ -près de (3.17).

Preuve. On réécrit (4.1) sous la forme :

$$\begin{pmatrix} A & B^T \\ -B & \varepsilon I \end{pmatrix} \cdot \begin{pmatrix} \vec{x}^\varepsilon \\ -\vec{\lambda}^\varepsilon \end{pmatrix} = \begin{pmatrix} -\vec{\alpha} \\ -\vec{c} \end{pmatrix}.$$

Et la matrice $\begin{pmatrix} A & B^T \\ -B & \varepsilon I \end{pmatrix}$ est définie positive (non symétrique) : en effet, pour tout $(\vec{y}, \vec{\mu}) \in \mathbb{R}^{n+p}$:

$$\begin{aligned} \left(\begin{pmatrix} A & B^T \\ -B & \varepsilon I \end{pmatrix} \cdot \begin{pmatrix} \vec{y} \\ \vec{\mu} \end{pmatrix}, \begin{pmatrix} \vec{y} \\ \vec{\mu} \end{pmatrix} \right)_{\mathbb{R}^{n+p}} &= \left(\begin{pmatrix} A.\vec{y} + B^T.\vec{\mu} \\ -B.\vec{y} + \varepsilon\vec{\mu} \end{pmatrix}, \begin{pmatrix} \vec{y} \\ \vec{\mu} \end{pmatrix} \right)_{\mathbb{R}^{n+p}} = (A.\vec{y}, \vec{y})_{\mathbb{R}^n} + \varepsilon(\vec{\mu}, \vec{\mu})_{\mathbb{R}^p} \\ &\geq \beta_{\min}(\|\vec{y}\|_{\mathbb{R}^n}^2 + \|\vec{\mu}\|_{\mathbb{R}^p}^2) = \beta_{\min}\|(\vec{y}, \vec{\mu})\|_{\mathbb{R}^{n+p}}^2, \end{aligned}$$

où $\beta_{\min} = \min(\alpha_{\min}, \varepsilon)$, où α_{\min} est la plus petite valeur propre de A . En particulier cette matrice carrée est inversible, d'où l'existence et l'unicité de la solution.

Calcul de l'erreur. On a $A.(\vec{x} - \vec{x}^\varepsilon) + B^T.(\vec{\lambda} - \vec{\lambda}^\varepsilon) = 0$, donc $B^T.(\vec{\lambda} - \vec{\lambda}^\varepsilon) = -A.(\vec{x} - \vec{x}^\varepsilon)$. En particulier $-A.(\vec{x} - \vec{x}^\varepsilon) \in \text{Im}B^T$. Et B^T est injective car B de rang max. Donc $B^T : \mathbb{R}^p \rightarrow \text{Im}B^T$ est bijective, d'où $B^{-T} : \text{Im}B^T \rightarrow \mathbb{R}^p$ est bien définie, d'où $(\vec{\lambda} - \vec{\lambda}^\varepsilon) = -B^{-T}.A.(\vec{x} - \vec{x}^\varepsilon)$ et :

$$\|\vec{\lambda} - \vec{\lambda}^\varepsilon\|_{\mathbb{R}^p} \leq \|B^{-T}\| \|A\| \|\vec{x} - \vec{x}^\varepsilon\|_{\mathbb{R}^n}.$$

Donc si on a (4.2)₁, on a (4.2)₂.

Montrons (4.2)₁. Les systèmes (3.17) et (4.1) équivalent à : pour tous $\vec{y} \in \mathbb{R}^n$ et $\vec{\mu} \in \mathbb{R}^p$:

$$\begin{cases} (A.\vec{x}, \vec{y})_{\mathbb{R}^n} + (B^T.\vec{\lambda}, \vec{y})_{\mathbb{R}^n} = -(\vec{\alpha}, \vec{y})_{\mathbb{R}^n}, \\ (B.\vec{x}, \vec{\mu})_{\mathbb{R}^p} = (\vec{c}, \vec{\mu})_{\mathbb{R}^p}, \end{cases} \quad \text{et} \quad \begin{cases} (A.\vec{x}^\varepsilon, \vec{y})_{\mathbb{R}^n} + (B^T.\vec{\lambda}^\varepsilon, \vec{y})_{\mathbb{R}^n} = -(\vec{\alpha}, \vec{y})_{\mathbb{R}^n}, \\ (B.\vec{x}^\varepsilon, \vec{\mu})_{\mathbb{R}^p} - \varepsilon(\vec{\lambda}^\varepsilon, \vec{\mu})_{\mathbb{R}^p} = (\vec{c}, \vec{\mu})_{\mathbb{R}^p}, \end{cases}$$

où $(\vec{x}, \vec{\lambda})$ et $(\vec{x}^\varepsilon, \vec{\lambda}^\varepsilon)$ sont les solutions de (3.17) et (4.1) D'où, pour tous $\vec{y} \in \mathbb{R}^n$ et $\vec{\mu} \in \mathbb{R}^p$:

$$\begin{cases} (A.(\vec{x} - \vec{x}^\varepsilon), \vec{y})_{\mathbb{R}^n} + (B^T.(\vec{\lambda} - \vec{\lambda}^\varepsilon), \vec{y})_{\mathbb{R}^n} = 0, \\ (B.(\vec{x} - \vec{x}^\varepsilon), \vec{\mu})_{\mathbb{R}^p} - \varepsilon(\vec{\lambda} - \vec{\lambda}^\varepsilon, \vec{\mu})_{\mathbb{R}^p} = -\varepsilon(\vec{\lambda}, \vec{\mu})_{\mathbb{R}^p}. \end{cases}$$

En particulier on prend $\vec{y} = \vec{x} - \vec{x}^\varepsilon$ et $\vec{\mu} = \vec{\lambda} - \vec{\lambda}^\varepsilon$, et on obtient en soustrayant les deux équations :

$$(A.(\vec{x} - \vec{x}^\varepsilon), \vec{x} - \vec{x}^\varepsilon)_{\mathbb{R}^n} + \varepsilon(\vec{\lambda} - \vec{\lambda}^\varepsilon, \vec{\lambda} - \vec{\lambda}^\varepsilon)_{\mathbb{R}^p} = \varepsilon(\vec{\lambda}, \vec{\lambda} - \vec{\lambda}^\varepsilon)_{\mathbb{R}^p}.$$

Par Cauchy-Schwarz, on déduit que, notant α_{\min} la plus petite valeur propre de A :

$$\alpha_{\min}\|\vec{x} - \vec{x}^\varepsilon\|_{\mathbb{R}^n}^2 + \varepsilon\|\vec{\lambda} - \vec{\lambda}^\varepsilon\|_{\mathbb{R}^p}^2 \leq \varepsilon\|\vec{\lambda}\|_{\mathbb{R}^p}\|\vec{\lambda} - \vec{\lambda}^\varepsilon\|_{\mathbb{R}^p},$$

et avec l'inégalité générique " $ab \leq \frac{a^2}{2} + \frac{b^2}{2}$ ", on obtient :

$$\alpha_{\min}\|\vec{x} - \vec{x}^\varepsilon\|_{\mathbb{R}^n}^2 + \frac{\varepsilon}{2}\|\vec{\lambda} - \vec{\lambda}^\varepsilon\|_{\mathbb{R}^p}^2 \leq \frac{\varepsilon}{2}\|\vec{\lambda}\|_{\mathbb{R}^p}^2.$$

En particulier $\alpha_{\min}\|\vec{x} - \vec{x}^\varepsilon\|_{\mathbb{R}^n}^2 \leq \frac{\varepsilon}{2}\|\vec{\lambda}\|_{\mathbb{R}^p}^2$. D'où $\|\vec{x} - \vec{x}^\varepsilon\|_{\mathbb{R}^n} = \frac{\|\vec{\lambda}\|_{\mathbb{R}^p}}{\alpha_{\min}^2} \varepsilon = O(\varepsilon)$, i.e. (4.2)₁. \blacksquare

Remarque 4.2 Malheureusement, la matrice dans (4.1) est mal conditionnée en générale quand ε est "petit". On transformera (4.1) pour obtenir la méthode du lagrangien augmenté, voir paragraphe 4.3. \blacksquare

4.2 Présentation de la méthode de résolution itérative d'Uzawa

4.2.1 Présentation

On veut résoudre le système (3.17) en le découplant. Ce système s'écrit : trouver $(\vec{x}, \vec{\lambda})$ t.q. :

$$\begin{cases} A.\vec{x} + B^T.\vec{\lambda} = -\vec{\alpha}, \\ B\vec{x} = \vec{c}, \end{cases} \quad (4.3)$$

Le problème est que dans (4.3)₂ il manque de l'information sur $\vec{\lambda}$. On remplace ce problème par le problème pénalisé (4.1) :

$$\begin{cases} A.\vec{x}^\varepsilon + B^T.\vec{\lambda}^\varepsilon = -\vec{\alpha}, \\ B\vec{x}^\varepsilon - \varepsilon\vec{\lambda} = \vec{c}. \end{cases} \quad (4.4)$$

Pour découpler ce problème et le résoudre séquentiellement, on essaie d'appliquer la méthode du point fixe : on suppose connu une approximation (\vec{x}^n, λ^n) , et on cherche une meilleure approximation $(\vec{x}^{n+1}, \vec{\lambda}^{n+1})$.

En particulier, connaissant $\vec{\lambda}^n$, on va calculer \vec{x}^{n+1} à l'aide de (4.4)₁ sous la forme :

$$A.\vec{x}^{n+1} = -B^T.\vec{\lambda}^n - \vec{\alpha}. \quad (4.5)$$

Puis connaissant \vec{x}^{n+1} on utilise (4.4)₂ pour calculer $\vec{\lambda}^{n+1}$:

$$\vec{\lambda}^{n+1} = \frac{B.\vec{x}^{n+1} - \vec{c}}{\varepsilon}. \quad (4.6)$$

À partir de (\vec{x}^n, λ^n) , on a obtenu $(\vec{x}^{n+1}, \vec{\lambda}^{n+1})$. Et on itère le procédé.

Malheureusement, l'erreur sur $\vec{\lambda}^{n+1}$ est mal contrôlée à cause de la division par ε .

4.2.2 Première amélioration

Plutôt que de considérer (4.6) avec la division par ε , on préfère calculer la différence $\Delta\vec{\lambda} = \vec{\lambda}^{n+1} - \vec{\lambda}^n$ en fonction de $\Delta\vec{x} = \vec{x}^{n+1} - \vec{x}^n$ (où \vec{x}^{n+1} est donné par (4.5)). Remarque : si on avait :

$$B.\vec{x}^{n+1} - \varepsilon\vec{\lambda}^{n+1} = \vec{c} \quad \text{et} \quad B.\vec{x}^n - \varepsilon\vec{\lambda}^n = \vec{c}, \quad (4.7)$$

alors par différence on obtiendrait :

$$\vec{\lambda}^{n+1} - \vec{\lambda}^n = \frac{B.(\vec{x}^{n+1} - \vec{x}^n)}{\varepsilon}. \quad (4.8)$$

Et on "espère" que $\Delta\vec{\lambda}$ est de l'ordre de ε ...

Puis on remarque que pour initialiser le calcul, on a besoin uniquement de $\vec{\lambda}^0$, et pas de \vec{x}^0 , ce pour calculer \vec{x}^1 à l'aide de (4.5), puis pour calculer $\vec{\lambda}^1$ à l'aide de (4.8).

On peut donc prendre un \vec{x}^0 quelconque (qu'on a pas besoin de donner), en particulier un \vec{x}^0 tel que :

$$B.\vec{x}^0 = \vec{c}, \quad (4.9)$$

qui correspond à une solution exacte, et alors pour la première itération avec donc $n = 0$ et $n+1 = 1$, (4.8) donne :

$$\vec{\lambda}^1 - \vec{\lambda}^0 = \frac{B.\vec{x}^1 - \vec{c}}{\varepsilon}.$$

Et on conserve ce schéma pour les itérations suivantes :

$$\vec{\lambda}^{n+1} - \vec{\lambda}^n = \frac{B.\vec{x}^{n+1} - \vec{c}}{\varepsilon}. \quad (4.10)$$

4.2.3 Algorithme

D'où l'algorithme :

1. Initialisation : on se donne un $\vec{\lambda}^0$ et un $\rho^0 = \frac{1}{\varepsilon^0} \gg 1$.
2. Boucle pour $n = 0, \dots$
 - * Trouver \vec{x}^{n+1} solution de (4.5), i.e. :

$$A.\vec{x}^{n+1} = -B^T.\vec{\lambda}^n - \vec{\alpha}. \quad (4.11)$$

(Résolution par exemple à l'aide de la méthode du gradient conjugué.)

- * Calcul de $\vec{\lambda}^{n+1}$ à l'aide de (4.10), i.e. :

$$\vec{\lambda}^{n+1} = \vec{\lambda}^n + \rho^n(B.\vec{x}^{n+1} - \vec{c}), \quad (4.12)$$

puis on se donne un $\rho^{n+1} \gg 1$ (choix dicté par la convergence requise, voir § suivant).

- * Arrêt si $\|\vec{\lambda}^{n+1} - \vec{\lambda}^n\|$ est suffisamment petit, sinon on boucle.

(Il reste à choisir les $\rho^n = \frac{1}{\varepsilon^n}$, voir plus loin théorème 4.6.)

4.3 Lagrangien augmenté

Pour améliorer le conditionnement de la matrice $\begin{pmatrix} A & B^T \\ B & -\varepsilon I \end{pmatrix}$ dans (4.1), et pour pouvoir choisir un $\rho^n = \frac{1}{\varepsilon^n}$ "optimal" à chaque étape, on utilise la méthode du lagrangien augmenté.

On rappelle que le lagrangien est donné par :

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) + (\vec{\lambda}, \vec{g}(\vec{x}))_{\mathbb{R}^p}. \quad (4.13)$$

On introduit le Lagrangien augmenté : pour $r \gg 1$ (qu'on prendra "assez grand"),

$$\mathcal{L}_r(\vec{x}, \vec{\lambda}) = \mathcal{L}(\vec{x}, \vec{\lambda}) + \frac{r}{2} \|\vec{g}(\vec{x})\|_{\mathbb{R}^p}^2. \quad (4.14)$$

Comme on doit satisfaire la contrainte $g(\vec{x}) = 0$, on a pas changé le problème mathématique.

Remarque 4.3 Par contre on a changé le conditionnement de la matrice à inverser, ce qui change le problème numérique : d'un problème non inversible numériquement (matrice mal conditionnée), on va passer à un problème inversible numériquement (matrice suffisamment bien conditionnée). Ici un choix de r assez grand permettra de "compenser" ε petit (voir théorème 4.6). ■

Donc, ayant $f(\vec{x}) = \frac{1}{2}(A.\vec{x}, \vec{x})_{\mathbb{R}^n} + (\vec{\alpha}, \vec{x})_{\mathbb{R}^n} + c_f$ et $g(x) = B.\vec{x} - \vec{c}$:

$$\begin{aligned} \mathcal{L}_r(\vec{x}, \vec{\lambda}) &= \frac{1}{2}(A.\vec{x}, \vec{x})_{\mathbb{R}^n} + (\vec{\alpha}, \vec{x})_{\mathbb{R}^n} + (\vec{\lambda}, B.\vec{x} - \vec{c})_{\mathbb{R}^p} + \frac{r}{2}(B.\vec{x} - \vec{c}, B.\vec{x} - \vec{c})_{\mathbb{R}^p} \\ &= \frac{1}{2}(A.\vec{x} + rB^T.B.\vec{x}, \vec{x})_{\mathbb{R}^n} + (\vec{\alpha} + B^T.\vec{\lambda} - rB^T.\vec{c}, \vec{x})_{\mathbb{R}^n} - (\vec{\lambda}, \vec{c})_{\mathbb{R}^p} + \frac{r}{2}\|\vec{c}\|^2. \end{aligned} \quad (4.15)$$

(voir exercice 4.5.) On pose alors :

$$A_r = A + rB^T.B, \quad (4.16)$$

matrice symétrique définie positive dès que A l'est, et (4.15) donne :

$$\begin{cases} \frac{\partial \mathcal{L}_r}{\partial \vec{x}}(\vec{x}, \vec{\lambda}) = A_r.\vec{x} + B^T.\vec{\lambda} + \vec{\alpha} - rB^T.\vec{c}, \\ \frac{\partial \mathcal{L}_r}{\partial \vec{\lambda}}(\vec{x}, \vec{\lambda}) = B.\vec{x} - \vec{c}. \end{cases}$$

Un extremum de \mathcal{L}_r satisfait à $\frac{\partial \mathcal{L}_r}{\partial \vec{x}}(\vec{x}, \vec{\lambda}) = \vec{0}$ et à $\frac{\partial \mathcal{L}_r}{\partial \vec{\lambda}}(\vec{x}, \vec{\lambda}) = 0$, et la recherche d'un extremum de \mathcal{L}_r s'écrit donc :

$$\begin{cases} A_r.\vec{x} + B^T.\vec{\lambda} = -\vec{\alpha} + rB^T.\vec{c}, \\ B.\vec{x} = \vec{c}, \end{cases} \quad (4.17)$$

à comparer avec (4.3).

Sous forme matricielle, le problème s'écrit : trouver $(\vec{x}, \vec{\lambda})$ t.q. :

$$\begin{pmatrix} A_r & B^T \\ B & 0 \end{pmatrix} \cdot \begin{pmatrix} \vec{x} \\ \vec{\lambda} \end{pmatrix} = \begin{pmatrix} -\vec{\alpha} + rB^T.\vec{c} \\ \vec{c} \end{pmatrix}. \quad (4.18)$$

Proposition 4.4 Si $(\vec{x}_r, \vec{\lambda}_r)$ est solution de (4.18) (extremum du lagrangien augmenté (4.14)), alors si $(\vec{x}, \vec{\lambda})$ est solution de (3.17) (extremum du lagrangien (4.13)), on a $(\vec{x}_r, \vec{\lambda}_r) = (\vec{x}, \vec{\lambda})$. (Voir remarque 4.3.)

Preuve. La solution de (4.17) s'écrit $A.\vec{x}_r + rB^T.B.\vec{x}_r + B^T.\vec{\lambda}_r = -\vec{\alpha} + rB^T.\vec{c}$ avec $B.\vec{x}_r = \vec{c}$, et donc $A.\vec{x}_r + B^T.\vec{\lambda}_r = -\vec{\alpha}$ avec $B.\vec{x}_r = \vec{c}$. Donc $(\vec{x}_r, \vec{\lambda}_r)$ est solution de (4.13).

La solution de (3.17) s'écrit $A.\vec{x} + B^T.\vec{\lambda} = -\vec{\alpha}$ avec $B.\vec{x} = \vec{c}$, donc avec $B^T.B.\vec{x} = B^T.\vec{c}$, et donc $A.\vec{x}_r + rB^T.B.\vec{x}_r + B^T.\vec{\lambda}_r = -\vec{\alpha} + rB^T.\vec{c}$ avec $B.\vec{x} = \vec{c}$. Donc $(\vec{x}, \vec{\lambda})$ est solution de (4.18). ■

Exercice 4.5 Montrer que la fonction :

$$\varphi : \vec{x} \in \mathbb{R}^n \rightarrow \varphi(\vec{x}) = \frac{1}{2} \|B.\vec{x} - \vec{c}\|_{\mathbb{R}^p}^2 \in \mathbb{R}$$

est différentiable en tout $\vec{x} \in \mathbb{R}^n$ et que :

$$\vec{\nabla}\varphi(\vec{x}) = B^T.B.\vec{x} - B^T.\vec{c}.$$

Réponse. On a $\varphi(\vec{x}) = \frac{1}{2}(B.\vec{x} - \vec{c}, B.\vec{x} - \vec{c})_{\mathbb{R}^p} = \frac{1}{2}(B.\vec{x}, B.\vec{x})_{\mathbb{R}^p} - (B.\vec{x}, \vec{c})_{\mathbb{R}^p} + \frac{1}{2}\|\vec{c}\|_{\mathbb{R}^p}^2$, fonction quadratique donc C^∞ . Posons $\varphi_1(\vec{x}) = (B.\vec{x}, \vec{c})_{\mathbb{R}^p} = (B^T.\vec{c}, \vec{x})_{\mathbb{R}^n}$. On a $\varphi_1(\vec{x} + h\vec{v}) - \varphi_1(\vec{x}) = h\varphi_1(\vec{v}) = h(B.\vec{v}, \vec{c})_{\mathbb{R}^p} = h(\vec{\nabla}\varphi_1(\vec{x}), \vec{v})_{\mathbb{R}^p} + o(h)$, d'où $\vec{\nabla}\varphi_1(\vec{x}) = B^T.\vec{c}$ (la fonction φ_1 est linéaire).

Posons $\varphi_2(\vec{x}) = \frac{1}{2}(B.\vec{x}, B.\vec{x})_{\mathbb{R}^p}$. La fonction φ_2 est quadratique, et

$$\varphi_2(\vec{x} + h\vec{v}) - \varphi_2(\vec{x}) = \frac{1}{2}h(B.\vec{v}, B.\vec{x})_{\mathbb{R}^p} + \frac{1}{2}h(B.\vec{x}, B.\vec{v})_{\mathbb{R}^p} + \frac{1}{2}h^2(B.\vec{v}, B.\vec{v})_{\mathbb{R}^p} = h(B^T.B.\vec{x}, \vec{v})_{\mathbb{R}^n} + o(h).$$

D'où $\vec{\nabla}\varphi_2(\vec{x}) = B^T.B.\vec{x}$. ■

4.4 Lagrangien augmenté pénalisé : résolution itérative d'Uzawa

Pour appliquer la méthode d'Uzawa à ce nouveau problème, on pénalise (4.17) : trouver $(\vec{x}^\varepsilon, \vec{\lambda}^\varepsilon)$ t.q. :

$$\begin{cases} A_r.\vec{x}^\varepsilon + B^T.\vec{\lambda}^\varepsilon = -\vec{\alpha} + rB^T.\vec{c}, \\ B.\vec{x}^\varepsilon - \varepsilon\vec{\lambda}^\varepsilon = \vec{c}, \end{cases} \quad (4.19)$$

où A_r est donnée par (4.16), i.e. :

$$\begin{pmatrix} A_r & B^T \\ B & -\varepsilon I \end{pmatrix} \cdot \begin{pmatrix} \vec{x}^\varepsilon \\ \vec{\lambda}^\varepsilon \end{pmatrix} = \begin{pmatrix} -\vec{\alpha} + rB^T.\vec{c} \\ \vec{c} \end{pmatrix}, \quad (4.20)$$

à comparer avec (4.1).

On applique le schéma donné par (4.11) et (4.12) où on utilise le lagrangien augmenté, i.e., si on connaît $\vec{\lambda}^n$ alors on calcule \vec{x}^{n+1} à l'aide de :

$$A_r.\vec{x}^{n+1} + B^T.\vec{\lambda}^n = -\vec{\alpha} + rB^T.\vec{c}, \quad (4.21)$$

puis connaissant \vec{x}^{n+1} on calcule $\vec{\lambda}^{n+1}$ à l'aide de :

$$\vec{\lambda}^{n+1} = \vec{\lambda}^n + \rho^n(B.\vec{x}^{n+1} - \vec{c}) \quad (\text{où } \rho^n = \frac{1}{\varepsilon^n}). \quad (4.22)$$

Algorithme :

1. Initialisation : on se donne un $\vec{\lambda}^0$ et un $\rho^0 = \frac{1}{\varepsilon^0} \gg 1$ (petit).
2. Boucle :
 - * Pour $\vec{\lambda}^n$ donné, trouver \vec{x}^{n+1} solution de (4.21). Solution calculée par exemple à l'aide de la méthode du gradient conjugué.
 - * Pour $\rho^n = \frac{1}{\varepsilon^n}$ donné, calcul de $\vec{\lambda}^{n+1}$ à l'aide de (4.22). Puis on se donne un ρ^{n+1} .
 - * Arrêt si $\|\vec{\lambda}^{n+1} - \vec{\lambda}^n\|$ est suffisamment petit, sinon retour à l'étape 2.

4.5 Lagrangien augmenté pénalisé : convergence du schéma d'Uzawa

Théorème 4.6 *On suppose que la matrice A de taille $n \times n$ est symétrique définie positive et la matrice B de taille $p \times n$ est de rang p . Et soit deux réels $r > 0$ et $\tilde{\rho} > 0$.*

Si la suite $(\rho^n)_{n \in \mathbb{N}}$ de l'algorithme du paragraphe précédent 4.4 est strictement positive et telle que, pour tout n :

$$\tilde{\rho} < \rho^n < 2r, \quad (4.23)$$

alors les suites $(\bar{x}^n)_{n \in \mathbb{N}^}$ et $(\bar{\lambda}^n)_{n \in \mathbb{N}}$ solutions du schéma (4.21)–(4.22) convergent vers la solution $(\bar{x}, \bar{\lambda})$ de (4.3).*

Preuve. La solution de (4.17) vérifie $B\bar{x} - \bar{c} = \vec{0}$, et donc (4.17) équivaut à :

$$\begin{cases} A_r \cdot \bar{x} + B^T \cdot \bar{\lambda} = -\bar{\alpha} + rB^T \cdot \bar{c}, \\ \bar{\lambda} = \bar{\lambda} + \rho^n (B \cdot \bar{x} - \bar{c}). \end{cases}$$

Et la solution du schéma d'Uzawa (4.21)–(4.22) vérifie (avec $\bar{\lambda}^n$ connu) :

$$\begin{cases} A_r \cdot \bar{x}^{n+1} + B^T \cdot \bar{\lambda}^n = -\bar{\alpha} + rB^T \cdot \bar{c}, \\ \bar{\lambda}^{n+1} = \bar{\lambda}^n + \rho^n (B \cdot \bar{x}^{n+1} - \bar{c}). \end{cases} \quad (4.24)$$

On pose $\Delta \bar{x}^n = \bar{x}^n - \bar{x}$ et $\Delta \bar{\lambda}^n = \bar{\lambda}^n - \bar{\lambda}$ les erreurs au pas n . On obtient (par soustraction) :

$$\begin{cases} A_r \cdot \Delta \bar{x}^{n+1} + B^T \cdot \Delta \bar{\lambda}^n = \vec{0}, \\ \Delta \bar{\lambda}^{n+1} = \Delta \bar{\lambda}^n + \rho^n (B \cdot \Delta \bar{x}^{n+1}). \end{cases} \quad (4.25)$$

(4.25)₂ donne :

$$\|\Delta \bar{\lambda}^{n+1}\|^2 = \|\Delta \bar{\lambda}^n\|^2 + 2\rho^n (\Delta \bar{\lambda}^n, B \cdot \Delta \bar{x}^{n+1}) + (\rho^n)^2 \|B \cdot \Delta \bar{x}^{n+1}\|^2. \quad (4.26)$$

Et $(\Delta \bar{\lambda}^n, B \cdot \Delta \bar{x}^{n+1})$ est donné à l'aide de (4.25)₁, sachant $A_r = A + rB^T \cdot B$:

$$(A \cdot \Delta \bar{x}^{n+1}, \Delta \bar{x}^{n+1}) + r(B^T \cdot B \cdot \Delta \bar{x}^{n+1}, \Delta \bar{x}^{n+1}) + (B^T \cdot \Delta \bar{\lambda}^n, \Delta \bar{x}^{n+1}) = 0,$$

soit :

$$(\Delta \bar{\lambda}^n, B \cdot \Delta \bar{x}^{n+1}) = -(A \cdot \Delta \bar{x}^{n+1}, \Delta \bar{x}^{n+1}) - r(B \cdot \Delta \bar{x}^{n+1}, B \cdot \Delta \bar{x}^{n+1}),$$

ce qui reporté dans (4.26) donne :

$$\|\Delta \bar{\lambda}^{n+1}\|^2 - \|\Delta \bar{\lambda}^n\|^2 = -2\rho^n (A \cdot \Delta \bar{x}^{n+1}, \Delta \bar{x}^{n+1}) - \rho^n (2r - \rho^n) \|B \cdot \Delta \bar{x}^{n+1}\|^2. \quad (4.27)$$

Si on prend $2r - \rho^n > 0$, alors la suite $(\|\Delta \bar{\lambda}^n\|)_{n \in \mathbb{N}}$ est décroissante, et comme cette suite est positive, elle est minorée par 0 et est donc convergente. Et la suite $(\|\Delta \bar{\lambda}^{n+1}\|)_{n \in \mathbb{N}}$ est donc convergente vers la même limite, et donc $\|\Delta \bar{\lambda}^n\|^2 - \|\Delta \bar{\lambda}^{n+1}\|^2 \rightarrow 0$.

Donc le membre de droite de (4.27) tend vers 0, et les deux termes étant négatifs (car A est définie positive), on a $\Delta \bar{x}^{n+1} \rightarrow \vec{0}$ (car A est définie positive et $(A \cdot \Delta \bar{x}^{n+1}, \Delta \bar{x}^{n+1}) \geq 0$). Et de (4.25)₁ on déduit que $\Delta \bar{\lambda}^n \rightarrow 0$ (en multipliant au préalable par B^T puisque $B^T \cdot B$ matrice $p \times p$ est inversible car B de rang maximum p). D'où la convergence du schéma. \blacksquare

Remarque 4.7 Le schéma converge également dans le cas $r = 0$: dans l'équation (4.27) on veut $\|\Delta \bar{\lambda}^n\|^2 - \|\Delta \bar{\lambda}^{n+1}\|^2 > 0$, i.e., avec $\rho^n > 0$, $2 + (2r - \rho^n) \frac{\|B \cdot \vec{v}\|_{\mathbb{R}^p}^2}{(A \cdot \vec{v}, \vec{v})_{\mathbb{R}^n}} > 0$ pour tout \vec{v} . On pose :

$$\beta^2 = \max_{\vec{v} \in \mathbb{R}^n} \frac{\|B \cdot \vec{v}\|_{\mathbb{R}^p}^2}{(A \cdot \vec{v}, \vec{v})_{\mathbb{R}^n}} \quad (= \max_{\vec{v} \in \mathbb{R}^n} \frac{\|B \cdot \vec{v}\|_{\mathbb{R}^p}^2}{\|\vec{v}\|_A^2}), \quad (4.28)$$

i.e. β^2 est la plus grande valeur propre de la matrice $C = A^{-1} \cdot B^T \cdot B$ (voir exercice suivant 4.8). On a donc convergence dès que :

$$\tilde{\rho} < \rho^n < 2(r + \frac{1}{\beta^2}). \quad (4.29)$$

Comme β^2 peut être très grand, prendre $r = 0$ imposerait de prendre ρ^n très petit, et la convergence de l'algorithme est alors très lente. La matrice $C = A^{-1} \cdot B^T \cdot B$ sera très utilisée dans la suite, bien que cette matrice $n \times n$ ne soit pas inversible (quand $n > p$ puisque B n'est alors pas injective). \blacksquare

Exercice 4.8 Montrer que pour A symétrique définie positive, le réel $\gamma = \max_{\vec{v} \in \mathbb{R}^n} \frac{\|B \cdot \vec{v}\|_{\mathbb{R}^p}^2}{(A \cdot \vec{v}, \vec{v})_{\mathbb{R}^n}}$ est la plus grande valeur propre de la matrice $C = A^{-1}B^T \cdot B$. Montrer que toutes les valeurs propres de $A^{-1}B^T \cdot B$ sont ≥ 0 .

N.B. : 0 est toujours valeur propre de C quand $p < n$, car B n'est pas injective dans ce cas.

Réponse. 1ère méthode. (Voir polycopié "valeurs propres généralisées".) Une valeur propre de la matrice $A^{-1}B^T \cdot B$ est un réel μ tel qu'il existe un vecteur $\vec{v} \neq 0$ vérifiant $A^{-1}B^T \cdot B \cdot \vec{v} = \mu \vec{v}$, donc tel que $B^T \cdot B \cdot \vec{v} = \mu A \cdot \vec{v}$. C'est un problème de valeurs propres généralisées (A étant une matrice symétrique définie positive) ; la matrice $B^T \cdot B$ étant symétrique, elle est A -diagonalisable (voir 3ème méthode). On note γ la plus grande des valeurs propres, et γ vérifie $\gamma = \max_{\vec{v} \in \mathbb{R}^n} \frac{(B^T \cdot B \cdot \vec{v}, \vec{v})_{\mathbb{R}^p}}{(A \cdot \vec{v}, \vec{v})_{\mathbb{R}^n}}$ (quotient de Rayleigh). Et les valeurs propres μ sont positives car, avec \vec{v} vecteur propre associé à μ , on a $0 \leq \|B \cdot \vec{v}\|_{\mathbb{R}^p}^2 = (B^T \cdot B \cdot \vec{v}, \vec{v})_{\mathbb{R}^p} = \mu (A \cdot \vec{v}, \vec{v})_{\mathbb{R}^n}$, avec $(A \cdot \vec{v}, \vec{v})_{\mathbb{R}^n} > 0$ puisque A est définie positive.

2ème méthode. Si on ne connaît pas ce résultat sur les valeurs propres généralisées, on se ramène à un problème de valeurs propres classique à l'aide de la décomposition de Cholevski $A = L \cdot L^T$, on posant de plus $L^T \cdot \vec{v} = \vec{w}$: le problème $A^{-1}B^T \cdot B \cdot \vec{v} = \mu \vec{v}$ s'écrit alors $L^{-1} \cdot B^T \cdot B \cdot L^{-T} \cdot \vec{w} = \mu \vec{w}$, soit $M \cdot \vec{w} = \mu \vec{w}$, où $M = L^{-1} \cdot B^T \cdot B \cdot L^{-T}$, avec M trivialement symétrique, donc diagonalisable. De plus M est positive car $(M \cdot \vec{w}, \vec{w})_{\mathbb{R}^n} = \|B \cdot L^{-T} \cdot \vec{w}\|_{\mathbb{R}^p}^2 \geq 0$. On note γ la plus grande valeur propre de M . Donc $\gamma = \sup_{\vec{w}} \frac{(M \cdot \vec{w}, \vec{w})_{\mathbb{R}^n}}{\|\vec{w}\|_{\mathbb{R}^n}^2} = \sup_{\vec{w}} \frac{(B \cdot L^{-T} \cdot \vec{w}, B \cdot L^{-T} \cdot \vec{w})_{\mathbb{R}^p}}{\|\vec{w}\|_{\mathbb{R}^n}^2} = \sup_{\vec{v}} \frac{(B \cdot \vec{v}, \vec{v})_{\mathbb{R}^p}}{(L^T \cdot \vec{v}, L^T \cdot \vec{v})_{\mathbb{R}^n}} = \sup_{\vec{v}} \frac{(B \cdot \vec{v}, \vec{v})_{\mathbb{R}^p}}{(A \cdot \vec{v}, \vec{v})_{\mathbb{R}^n}}$ comme annoncé.

3ème méthode. Variante de la 1ère méthode (problème de valeurs propres généralisées). Notons, pour $\vec{v}, \vec{w} \in \mathbb{R}^n$:

$$(\vec{v}, \vec{w})_A = (A \cdot \vec{v}, \vec{w})_{\mathbb{R}^n}.$$

$(\cdot, \cdot)_A$ est un produit scalaire car A est une matrice symétrique définie positive. Et pour $\vec{v} \in \mathbb{R}^n$ on a :

$$\|B \cdot \vec{v}\|_{\mathbb{R}^p}^2 = (B \cdot \vec{v}, B \cdot \vec{v})_{\mathbb{R}^p} = (B^T \cdot B \cdot \vec{v}, \vec{v})_{\mathbb{R}^n} = (A \cdot A^{-1} \cdot B^T \cdot B \cdot \vec{v}, \vec{v})_{\mathbb{R}^n} = (A^{-1} \cdot B^T \cdot B \cdot \vec{v}, \vec{v})_A = (C \cdot \vec{v}, \vec{v})_A.$$

Donc $\frac{\|B \cdot \vec{v}\|_{\mathbb{R}^p}^2}{(A \cdot \vec{v}, \vec{v})_{\mathbb{R}^n}} = \frac{(C \cdot \vec{v}, \vec{v})_A}{\|\vec{v}\|_A^2}$. De plus la matrice C est $(\cdot, \cdot)_A$ -adjointe (i.e. est "symétrique relativement au produit scalaire A "), i.e. $(C \cdot \vec{v}, \vec{w})_A = (\vec{v}, C \cdot \vec{w})_A$ pour tout $\vec{v}, \vec{w} \in \mathbb{R}^n$. En effet :

$$(C \cdot \vec{v}, \vec{w})_A = (B^T \cdot B \cdot \vec{v}, \vec{w})_{\mathbb{R}^n} = (\vec{v}, B^T \cdot B \cdot \vec{w})_{\mathbb{R}^n} = (\vec{v}, C \cdot \vec{w})_A.$$

Donc C est diagonalisable relativement au produit scalaire $(\cdot, \cdot)_A$, et la plus grande valeur propre de C est donnée par $\gamma = \max_{\vec{v} \in \mathbb{R}^n} \frac{(C \cdot \vec{v}, \vec{v})_A}{\|\vec{v}\|_A^2} = \max_{\vec{v} \in \mathbb{R}^n} \frac{\|B \cdot \vec{v}\|_{\mathbb{R}^p}^2}{\|\vec{v}\|_A^2}$ comme indiqué. ■

Remarque 4.9 β^2 est aussi la plus grande valeur propre de la matrice symétrique $A^{-\frac{1}{2}} \cdot B^T \cdot B \cdot A^{-\frac{1}{2}}$. En effet, A étant symétrique définie positive, on a, posant $\vec{w} = \sqrt{A} \cdot \vec{v}$:

$$\beta^2 = \max_{\vec{w} \in \mathbb{R}^n} \frac{(B \cdot A^{-\frac{1}{2}} \cdot \vec{w}, B \cdot A^{-\frac{1}{2}} \cdot \vec{w})_{\mathbb{R}^p}}{(\vec{w}, \vec{w})_{\mathbb{R}^n}} = \max_{\vec{w} \in \mathbb{R}^n} \frac{(A^{-\frac{1}{2}} \cdot B^T \cdot B \cdot A^{-\frac{1}{2}} \cdot \vec{w}, \vec{w})_{\mathbb{R}^n}}{(\vec{w}, \vec{w})_{\mathbb{R}^n}}.$$

Et β^2 est le max du quotient de rayleigh $\frac{(M \cdot \vec{v}, \vec{v})_{\mathbb{R}^n}}{\|\vec{v}\|_{\mathbb{R}^n}^2}$ pour la matrice symétrique $M = A^{-\frac{1}{2}} \cdot B^T \cdot B \cdot A^{-\frac{1}{2}}$, i.e. β^2 est la plus grande valeur propre de cette matrice M .

On a donc $A^{-\frac{1}{2}} \cdot B^T \cdot B \cdot A^{-\frac{1}{2}} \cdot \vec{w} = \mu \vec{w}$ pour \vec{w} vecteur propre associé à μ . En multipliant par $A^{-\frac{1}{2}}$ et posant $\vec{v} = A^{-\frac{1}{2}} \cdot \vec{w}$ on a $A^{-1} \cdot B^T \cdot B \cdot \vec{v} = \mu \vec{v}$, donc $A^{-1} \cdot B^T \cdot B$ est diagonalisable et a mêmes valeurs propres que M , d'où β^2 est aussi la plus grande valeur propre de $A^{-1} \cdot B^T \cdot B$. ■

Remarque 4.10 Avec (4.24) aux pas n et $n+1$:

$$\begin{cases} A_r \cdot (\vec{x}^{n+1} - \vec{x}^n) + B^T \cdot (\vec{\lambda}^n - \vec{\lambda}^{n-1}) = 0 \\ \vec{\lambda}^n - \vec{\lambda}^{n-1} = \rho^{n-1} (B \cdot \vec{x}^n - \vec{c}). \end{cases}$$

Quitte à changer d'origine, on prend $\vec{c} = \vec{0}$. D'où :

$$\vec{x}^{n+1} = \vec{x}^n - \rho^{n-1} C_r \cdot \vec{x}^n \quad \text{où} \quad C_r = A_r^{-1} \cdot B^T \cdot B.$$

■

4.6 Choix de ρ^n

Proposition 4.11 1- Si on prend $r = 0$ (lagrangien non augmenté) et $\rho^n = \rho^0$ pour tout n (schéma à pas ρ^n constant), le choix optimal pour $\rho^n = \rho^0$ est donné par :

$$\rho^0 = \frac{2}{\alpha_m + \alpha_M} \quad (\text{cas } r = 0), \quad (4.30)$$

où α_m désigne la valeur propre minimale non nulle et α_M la valeur propre maximale de la matrice $C = A^{-1} \cdot B^T \cdot B$. Le taux de convergence de la méthode est alors $\tau = \frac{1 - \frac{\alpha_m}{\alpha_M}}{1 + \frac{\alpha_m}{\alpha_M}}$.

2- Pour $r \neq 0$ et :

$$\rho^n = \rho^0 = r \quad (\text{cas } r > 0), \quad (4.31)$$

(choix classique), le taux de convergence de la méthode est $\tau = \frac{1}{1+r\alpha_m}$. En particulier, si on choisit $r = \frac{1}{\alpha_m}$, alors $\tau \leq \frac{1}{2}$.

Preuve. Quitte à changer d'origine, on prend $\vec{c} = \vec{0}$.

Dans l'algorithme d'Uzawa (4.21)-(4.22) on a :

$$\vec{x}^{n+1} = -A_r^{-1}.B^T.\vec{\lambda}^n - A_r^{-1}\vec{\alpha},$$

d'où $\Delta\vec{x}^{n+1} = -A_r^{-1}.B^T.\Delta\vec{\lambda}^n$, d'où, avec (4.25)₂ :

$$\Delta\vec{\lambda}^{n+1} = \Delta\vec{\lambda}^n - \rho^n B.A_r^{-1}.B^T.\Delta\vec{\lambda}^n.$$

On pose :

$$\vec{\mu}^n = A^{-1}.B^T.\Delta\vec{\lambda}^n. \quad (4.32)$$

L'égalité précédente s'écrit :

$$\begin{aligned} \vec{\mu}^{n+1} &= \vec{\mu}^n - \rho^n A^{-1}.B^T.B.A_r^{-1}.A.\vec{\mu}^n \\ &= (I - \rho^n C.A_r^{-1}.A).\vec{\mu}^n \quad \text{où} \quad C = A^{-1}.B^T.B. \end{aligned}$$

Et $A_r^{-1} = (A + rB^T.B)^{-1} = (A(I + rA^{-1}.B^T.B))^{-1} = (I + rC)^{-1}.A^{-1}$, et donc :

$$\vec{\mu}^{n+1} = (I - \rho^n C.(I + rC)^{-1}).\vec{\mu}^n. \quad (4.33)$$

Notons ici $\mathbb{R}^N = \mathbb{R}^n$, où donc N la dimension de l'espace, pour ne pas confondre avec l'itération n . Soit une base $(\vec{w}_i)_{i=1,\dots,N}$ de vecteurs propres de la matrice $C = A^{-1}.B^T.B$ associées aux valeurs propres α_i (voir remarque 4.7) : pour tout $i = 1, \dots, N$.

$$C.\vec{w}_i = \alpha_i \vec{w}_i \quad (4.34)$$

On décompose $\vec{\mu}^n$ sur la base $(\vec{w}_i)_{i=1,\dots,N}$:

$$\vec{\mu} = \sum_{i=1}^N \mu_i \vec{w}_i, \quad C.\vec{\mu} = \sum_{i=1}^N \alpha_i \mu_i \vec{w}_i, \quad (I + rC).\vec{\mu} = \sum_{i=1}^N (1 + r\alpha_i) \mu_i \vec{w}_i.$$

Composante par composante sur la base (\vec{w}_i) , (4.33) donne :

$$\begin{cases} \mu_i^{n+1} = (1 - \frac{\rho^n \alpha_i}{1 + r\alpha_i}) \mu_i^n = (\frac{1 + (r - \rho^n) \alpha_i}{1 + r\alpha_i}) \mu_i^n, & \text{si } \alpha_i \neq 0, \\ \mu_i^{n+1} = 0, & \text{si } \alpha_i = 0. \end{cases} \quad (4.35)$$

On veut la décroissance la plus rapide pour la suite $(\vec{\mu}^n)$, i.e on veut $\max_i |(\frac{1 + (r - \rho^n) \lambda_i}{1 + r\lambda_i})|$ le plus petit possible.

Pour les $\alpha_i = 0$, on a déjà converger ($\mu_i^{n+1} = 0$, donc $\Delta\lambda^{n+1} = 0$, car B^T est injective car B de rang p). On ne s'intéresse donc qu'aux composantes μ_i^{n+1} pour i t.q $\alpha_i \neq 0$.

1- En particulier, si $r=0$ et $\rho^n = \rho^0$ pour tout n , alors (4.35) donne $\mu_i^{n+1} = (1 - \rho^0 \alpha_i) \mu_i^n$ pour tout i et tout n , et $\max_i |1 - \rho^0 \alpha_i|$ est minimal pour $\rho^0 = \frac{2}{\alpha_m + \alpha_M}$. (Dessiner les graphes des fonctions $\rho \rightarrow |1 - \rho \lambda_i|$.) Et pour $\rho^0 = \frac{2}{\alpha_m + \alpha_M}$, on obtient :

$$|\mu_i^{n+1}| \leq \frac{1 - \frac{\alpha_m}{\alpha_M}}{1 + \frac{\alpha_m}{\alpha_M}} |\mu_i^n|;$$

Et $\tau = \frac{1 - \frac{\alpha_m}{\alpha_M}}{1 + \frac{\alpha_m}{\alpha_M}}$ est le taux de convergence de la méthode.

2- Si $r \neq 0$ et $\rho^n = r$, (4.35) donne $\mu_i^{n+1} = (\frac{1}{1+r\alpha_i}) \mu_i^n$. En prenant les valeurs absolues, on obtient $\tau = \frac{1}{1+r\alpha_m}$. \blacksquare

Remarque 4.12 Quand $\rho^n = r > 0$, le taux de convergence montre qu'il est souhaitable de prendre r "le plus grand possible". On ne peut cependant pas prendre r trop grand, le conditionnement de la matrice $A_r = A + rB^T.B$ se détériorant avec r : on peut montrer que $\text{cond}(A_r) = r \frac{\|B\|^2}{\lambda_{\min}(A)}$, où $\|B\| = \sup \frac{\|B.\vec{x}\|_{\mathbb{R}^n}}{\|\vec{x}\|_{\mathbb{R}^n}}$ est la norme de B et $\lambda_{\min}(A)$ est la plus petite valeur propre de A . \blacksquare

4.7 Uzawa : c'est un algorithme de gradient

Quitte à changer d'origine, on prend $\vec{c} = \vec{0}$. Par élimination de \vec{x}^{n+1} dans (4.21)-(4.22), on obtient :

$$\frac{\vec{\lambda}^{n+1} - \vec{\lambda}^n}{\rho^n} = B.A_r^{-1}.B^T.\vec{\lambda}^n - B.A_r^{-1}.\vec{\alpha}. \quad (4.36)$$

On note :

$$U_r = B.A_r^{-1}.B^T$$

la matrice d'Uzawa relative à A_r . On considère la fonctionnelle "duale" (à une constante près) :

$$G_r(\vec{\lambda}) = -\frac{1}{2}\vec{\lambda}^T.U_r.\vec{\lambda} + \vec{\lambda}^T.(B.A_r^{-1}.\vec{\alpha}),$$

à comparer avec (3.9)

Et on cherche à maximiser G_r . On a, la matrice $B.A_r^{-1}.B^T$ étant symétrique :

$$-\vec{\nabla}G_r(\vec{\lambda}) = B.A_r^{-1}.B^T.\vec{\lambda} - B.A_r^{-1}.\vec{\alpha},$$

qui est bien la direction de descente (ici direction de montée pour le maximum) choisie pour trouver $\vec{\lambda}^{n+1}$ à partir de $\vec{\lambda}^n$ dans (4.36).

Donc, la méthode d'Uzawa est un algorithme de gradient appliqué à la fonctionnelle duale.

A Annexe. Dimensions des noyaux et images

(Voir aussi le polycopié "algèbre linéaire".)

Une matrice $U = [u_{ij}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ de taille $m \times n$ est dite triangulaire supérieure ssi $u_{ij} = 0$ pour tout i, j t.q. $i > j$:

$$U = \begin{pmatrix} * & * & \dots & * & \dots & * \\ 0 & * & \dots & * & \dots & * \\ \vdots & \ddots & \ddots & & & \\ 0 & \dots & 0 & * & \dots & * \end{pmatrix}$$

où les '*' représentent des éléments éventuellement non nuls.

Une matrice $U = [u_{ij}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ de taille $m \times n$ triangulaire supérieure est dit de type échelon ssi elle est de la forme :

$$U = \begin{pmatrix} * & * & \dots & * & \dots & * \\ \boxed{0} & * & * & * & \dots & * \\ 0 & \boxed{0} & 0 & * & \dots & * \\ 0 & 0 & \dots & \boxed{0} & * & * \end{pmatrix}$$

i.e. : si j_i est l'indice tel que u_{i,j_i} est le premier élément non nul de la ligne i , alors pour la ligne $i+1$ on a $u_{i+1,j} = 0$ pour tout $j \leq j_i+1$. Par exemple, la matrice $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}$ est une matrice échelon, et toute matrice triangulaire supérieure est une matrice échelon.

Soit B la matrice $m \times n$ représentant l'application linéaire $\tilde{B} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ dans des bases données. On appelle rang de l'application linéaire le réel $r = \text{rang} B = \stackrel{\text{def}}{\dim(\text{Im}\tilde{B})}$. Et on appelle rang de la matrice B ce même réel r .

Proposition A.1 Soit B une matrice $m \times n$. Il existe une matrice triangulaire inférieure L de taille $m \times m$ avec des 1 sur la diagonale, il existe une matrice P inversible de tailles $m \times m$ (matrice de permutation des lignes de B), et il existe une matrice triangulaire supérieure échelon U de taille $m \times n$ telles que :

$$PB = LU. \quad (\text{A.1})$$

avec B et U de même rang : $\text{rang}(B) = \text{rang}(U)$.

Preuve. On applique la technique usuelle de la décomposition LU de Gauss pour les matrices carrées.

Et comme L et P sont inversibles, B et U ont même rang. ▀

Exemple A.2 Exemple pris dans Strang : Linear Algebra and its Applications. Soit $A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix}$. Le premier pivot est $1 \neq 0$, et on “soustrait” la ligne 1 aux autres lignes :

soit $E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, on obtient $E_1.A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 6 & 2 \end{pmatrix}$. On “soustrait” la ligne 2 à la

ligne 3 : soit $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$, on obtient $E_2.E_1.A = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \text{noté } U$. Ici $P = I$ et

$$L = E_1^{-1}.E_2^{-2} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}. \quad \blacksquare$$

Proposition A.3 Soit B une matrice $m \times n$.

$$(i) \quad \text{rang}(B) = \text{rang}(B^T) \quad (= r), \quad (\text{A.2})$$

$$(ii) \quad \text{Ker}(B) = (\text{Im}B^T)^\perp, \quad (\text{A.3})$$

$$(iii) \quad \mathbb{R}^n = \text{Ker}B \oplus \text{Im}B^T, \quad (\text{A.4})$$

$$(iv) \quad \dim(\text{Ker}B) + \dim(\text{Im}B) = n \quad (\text{i.e.} \quad \dim(\text{Ker}B) = n - r). \quad (\text{A.5})$$

Preuve. Avec (A.1), il s’agit de montrer que $\text{rang}(U) = \text{rang}(U^T)$: par construction de U , les r premières lignes de U sont indépendantes (et les $m-r$ dernières lignes sont nulles). Et donc les r premières colonnes de U^T sont indépendantes (et les $m-r$ dernières colonnes sont nulles). D’où (A.2).

Puis $(\text{Im}B^T)^\perp = \{\vec{x} \in \mathbb{R}^n : \forall \vec{\lambda} \in \mathbb{R}^m, (\vec{x}, B^T.\vec{\lambda})_m = 0\} = \{\vec{x} \in \mathbb{R}^n : \forall \vec{\lambda} \in \mathbb{R}^m, (B.\vec{x}, \vec{\lambda})_m = 0\} = \{\vec{x} \in \mathbb{R}^n : B.\vec{x} = 0\} = \text{Ker}B$. I.e. (A.3).

D’où $(\text{Ker}B)^\perp = \text{Im}B^T$, avec $\mathbb{R}^n = \text{Ker}B \oplus (\text{Ker}B)^\perp$, d’où $\mathbb{R}^n = \text{Ker}B \oplus \text{Im}B^T$, i.e. (A.4).

Et donc $\dim(\text{Ker}B) + \dim(\text{Im}B^T) = n$ et donc (A.5) à l’aide de (A.2). \blacksquare

Remarque A.4 On rappelle que $\text{Im}B = \text{Vect}\{\text{col}_i(B); i = 1, \dots, n\}$ espace vectoriel engendré par les vecteurs colonnes de B . En effet, $B.\vec{x}$ est le vecteur $= \sum_i x_i \text{col}_i(B)$. Donc $\text{Im}B^T = \text{Vect}\{\text{ligne}_i(B^T); i = 1, \dots, n\}$.

Et $\vec{x} \in \text{Ker}B$ ssi $B.\vec{x} = 0$, i.e. ssi $\text{ligne}_i(B).\vec{x} = 0$ pour tout i , i.e. ssi \vec{x} est orthogonal à tous les vecteurs ligne constituant B , i.e. orthogonal à tous les vecteurs colonne constituant B^T , i.e. orthogonal à $\text{Im}B^T$: c’est (A.3). \blacksquare

B Annexe. Matrices similaires et congruentes

(Voir aussi le polycopié “algèbre linéaire”).

B.1 Matrices similaires

Définition B.1 Deux matrices carrées A et B de \mathbb{R}^{n^2} sont similaires ssi il existe une matrice inversible $P \in \mathbb{R}^{n^2}$ telle que :

$$B = P^{-1}.A.P \quad (\text{B.1})$$

Définition B.2 La matrice carrée A est dite diagonalisable ssi A est similaire à une matrice diagonale $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, et on notera alors $D = P^{-1}.A.P$.

En particulier, si A est diagonalisable alors $A.P = P.D$ indique que les colonnes de P sont les vecteurs propres associés aux valeurs propres qui sont les éléments diagonaux de D . En effet, notant \vec{p}_j la j -ème colonne de P , l’égalité $A.P = P.D$ est l’identité $A.\vec{p}_j = \lambda_j \vec{p}_j$ pour tout j .

Calcul des valeurs propres : on calcule $\det(A - \lambda I) = P(\lambda) =$ le polynôme caractéristique de la matrice A , polynôme de degré n , et on cherche les racines λ de P , i.e. t.q. $P(\lambda) = 0$ (on cherche les λ t.q. la matrice $A - \lambda I$ n’est pas inversible) : on obtient ainsi toutes les valeurs propres λ_i .

Proposition B.3 Deux matrices similaires ont mêmes valeurs propres.

Preuve. $\det(B - \lambda I) = \det(P^{-1}.A.P - \lambda I) = \det(P^{-1}(A - \lambda I)P) = \det(P^{-1})\det(A - \lambda I)\det(P) = \det(A - \lambda I)$. \blacksquare

B.2 Matrices congruentes et loi d'inertie de Sylvester

Définition B.4 Si A et B sont deux matrices $n \times n$, elles sont dites congruentes ssi il existe une matrice Z inversible telle que :

$$B = Z^T . A . Z. \quad (\text{B.2})$$

Dans ce cas, comme $B^T = Z^T . A^T . Z$, on a l'équivalence A symétrique $\Leftrightarrow B$ symétrique.

Remarque B.5 Les matrices congruentes sont utilisées pour les fonctions quadratiques ou les formes bilinéaires symétriques, alors que les matrices similaires sont utilisées pour les endomorphismes. \blacksquare

Remarque B.6 Utilisation courante de la congruence : si on considère le changement de variables $\vec{x} \rightarrow \vec{y} = Z^{-1} . \vec{x}$, alors la forme quadratique $\Phi : \vec{x} \rightarrow \Phi(\vec{x}) = (A . \vec{x}, \vec{x})_{\mathbb{R}^n}$ de matrice A devient la forme quadratique $\Psi : \vec{y} \rightarrow (A . Z . \vec{y}, Z . \vec{y})_{\mathbb{R}^n} = (Z^T . A . Z . \vec{y}, \vec{y})_{\mathbb{R}^n} = (B . \vec{y}, \vec{y})_{\mathbb{R}^n}$ de matrice B congruente à A . \blacksquare

Proposition B.7 (*Loi d'inertie de Sylvester.*) Deux matrices symétriques congruentes ont même signature, i.e. ont même nombre de valeurs propres strictement positives, strictement négatives et nulles.

Preuve. Soit A et B deux matrices congruentes. Il existe donc Z inversible t.q. $B = Z^T . A . Z$. On considère la décomposition QR de Z , soit $Z = QR$, avec $Q^T = Q^{-1}$ (matrice orthogonale) et R matrice triangulaire supérieure (Right triangular matrix) de diagonale strictement positive (procédé d'orthogonalisation de Gram-Schmidt à partir des colonnes de Z qui forment une base car Z est inversible).

Pour $t \in [0, 1]$, la matrice $(1-t)R + tI$ (barycentre de R et I) est inversible car est triangulaire supérieure à diagonale strictement positive (immédiat).

On pose $Z(t) = Q . ((1-t)R + tI)$ pour $t \in [0, 1]$, qui est donc inversible pour $t \in [0, 1]$ (produit de deux matrices inversibles). On a $Z(0) = Q . R = Z$ et $Z(1) = Q$.

Notons $\mathcal{B}(t) = Z^T(t) . A . Z(t)$. Donc $\mathcal{B}(0) = Z^T . A . Z = B$, et $\mathcal{B}(1) = Q^T . A . Q = Q^{-1} . A . Q$ est similaire à A , donc a mêmes valeurs propres que A .

Z est une fonction affine, donc continue (pour la norme matricielle).

Donc \mathcal{B} est une fonction quadratique, donc continue (pour la norme matricielle).

Donc, pour λ fixé, $t \rightarrow \det(\mathcal{B}(t) - \lambda I)$ est une fonction continue.

Comme $\mathcal{B}(t)$ est symétrique pour tout t , $\mathcal{B}(t)$ est diagonalisable.

Si A est inversible, alors, pour $t \in [0, 1]$, $\mathcal{B}(t)$ est inversible et n'a donc aucune valeur propre nulle. Donc, pour $t \in [0, 1]$, par continuité de \mathcal{B} , $\mathcal{B}(t)$ a même signature que $\mathcal{B}(0) = B$ (sinon il existerait t t.q. 0 est valeur propre de $\mathcal{B}(t)$, et alors $\mathcal{B}(t)$ n'est pas inversible). En particulier $\mathcal{B}(1)$ a même signature que B . Et comme $\mathcal{B}(1)$ est similaire à A , $\mathcal{B}(1)$ à la même signature que A , cf. proposition B.3. Donc B à même signature que A .

Si A n'est pas inversible, on considère $A + \varepsilon I$ qui a pour valeurs propres les $\lambda_i + \varepsilon$ où les λ_i sont les valeurs propres de A . Et $A + \varepsilon I$ est inversible pour $\varepsilon > 0$ suffisamment petit, et on applique le raisonnement précédent, puis on fait tendre ε vers 0. \blacksquare

Remarque B.8 Autre preuve (toujours dans le cas A symétrique et donc B symétrique). On pose $\lambda_1 \leq \dots \leq \lambda_n$ les valeurs propres de A et $\mu_1 \leq \dots \leq \mu_n$ celles de B . Supposons $\lambda_k > 0$.

Le quotient de Rayleigh $R_A(\vec{v}) = \frac{(A . \vec{v}, \vec{v})_{\mathbb{R}^n}}{(\vec{v}, \vec{v})_{\mathbb{R}^n}}$ de A donne $\lambda_k = \sup_{V : \dim V = k} \min_{\vec{v} \in V} R_A(\vec{v})$. Et avec Z inversible, on a $\frac{(A . \vec{v}, \vec{v})_{\mathbb{R}^n}}{(\vec{v}, \vec{v})_{\mathbb{R}^n}} = \frac{(A . Z . \vec{w}, Z . \vec{w})_{\mathbb{R}^n}}{(Z . \vec{w}, Z . \vec{w})_{\mathbb{R}^n}} = \frac{(Z^T . A . Z . \vec{w}, \vec{w})_{\mathbb{R}^n}}{(\vec{w}, \vec{w})_{\mathbb{R}^n}} = \frac{(B . \vec{w}, \vec{w})_{\mathbb{R}^n}}{(Z . \vec{w}, Z . \vec{w})_{\mathbb{R}^n}} \leq \frac{(B . \vec{w}, \vec{w})_{\mathbb{R}^n}}{(\vec{w}, \vec{w})_{\mathbb{R}^n}} \frac{1}{\alpha}$, où $\alpha > 0$ est la plus petite valeur propre de $Z^T . Z$ (comme $Z^T . Z$ est une matrice symétrique définie positive, elle est diagonalisable de plus petite valeur propre $\alpha > 0$). D'où $\lambda_k \leq \mu_k \frac{1}{\alpha}$, et μ_k est ≥ 0 .

D'où B a au moins autant de valeurs propres positives que A . En inversant les rôles de A et B , on déduit que A a autant de valeurs propres positives que B . Puis en considérant $-A$ et $-B$, on déduit que A a autant de valeurs propres négatives que B . D'où le résultat. \blacksquare

Corollaire B.9 Si A est une matrice symétrique, si $A = L . C . L^T$ est sa décomposition de Cholesky, avec L matrice triangulaire inférieure avec des 1 sur la diagonale et C matrice diagonale contenant les pivots, alors il y a autant de pivots strictement positifs, strictement négatifs et nuls qu'il y a de valeurs propres strictement positives, strictement négatives et nulles.

Preuve. A étant symétrique, on peut choisir une décomposition diagonale $A = P.D.P^{-1}$ où $P^{-1} = P^T$. Donc A est congruente à D . Et $A = L.C.L^T$ est congruente à C . D'où C est congruente à D car $C = (L^{-1}.P).D.(P^{-1}.L^{-T})$ avec $(L^{-1}.P)^T = P^T.L^{-T} = P^{-1}.L^{-T}$. Et on applique loi d'inertie de Sylvester. \blacksquare

Exercice B.10 Montrer que si A et B sont deux matrices symétriques congruentes, $B = Z^T.A.Z$ avec Z inversible, alors leurs matrices diagonales des valeurs propres sont congruentes. I.e. si

$$\begin{cases} D = P^{-1}.A.P, & D = \text{diag}(\lambda_1, \dots, \lambda_n), & P^{-1} = P^T, \\ E = Q^{-1}.B.Q, & E = \text{diag}(\mu_1, \dots, \mu_n), & Q^{-1} = Q^T, \end{cases} \quad (\text{B.3})$$

alors :

$$B = Z^T.A.Z \iff E = Y^T.D.Y, \quad (\text{B.4})$$

où $Y = P^T.Z.Q$ est inversible.

Réponse. On a $(Q.E.Q^{-1}) = Z^T.(P.D.P^{-1}).Z$, d'où $E = (Q^{-1}.Z^T.P).D.(P^{-1}.Z.Q)$ et $(Q^{-1}.Z^T.P)^T = P^T.Z.(Q^{-1})^T = P^{-1}.Z.Q$. \blacksquare

Exercice B.11 Démonstration simple de : si D et E sont deux matrices diagonales congruentes et si D est une matrice définie positive, alors E est définie positive.

Réponse. On a $E = Y^T.D.Y$ avec Y inversible. Donc $(E.\vec{v}, \vec{v})_{\mathbb{R}^n} = (D.Y.\vec{v}, Y.\vec{v})_{\mathbb{R}^n}$. Ordonnons les valeurs propres de E : $\mu_1 \leq \dots \leq \mu_n$, et quitte à réordonner la base canonique de \mathbb{R}^n , on a $E = \text{diag}(\mu_1, \dots, \mu_n)$. D'où :

$$\mu_1 = (E.\vec{e}_1, \vec{e}_1)_{\mathbb{R}^n} = (D.Y.\vec{e}_1, Y.\vec{e}_1)_{\mathbb{R}^n} = (D.\vec{y}_1, \vec{y}_1) = \sum_{i=1}^n \lambda_i y_{i1}^2,$$

où \vec{y}_1 dénote la première colonne de la matrice Y , et où $Y = [y_{ij}]$. On en déduit que si tous les λ_i sont > 0 alors $\mu_1 > 0$ et donc tous les μ_i sont > 0 . \blacksquare

Exercice B.12 Soit $y_{11} \in \mathbb{R}^*$, Y_n matrice $n \times n$ inversible, $\vec{c}_2 \in \mathbb{R}^n$ et $Y = \begin{pmatrix} y_{11} & 0 \\ \vec{c}_2^T & Y_n \end{pmatrix}$. Montrer par récurrence : si E et D sont diagonales t.q. $E = Y^T.D.Y$, alors E et D ont la même signature.

Réponse. Notant $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & D_n \end{pmatrix}$ avec D_n matrice diagonale :

$$Y^T.D.Y = \begin{pmatrix} y_{11} & \vec{c}_2^T \\ 0 & Y_n^T \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 y_{11} & 0 \\ D_n \vec{c}_2 & D_n.Y_n \end{pmatrix} = \begin{pmatrix} \lambda_1 y_{11}^2 + \vec{c}_2^T.D_n.\vec{c}_2 & \vec{c}_2^T.D_n.Y_n \\ Y_n^T.D_n.\vec{c}_2 & Y_n^T.D_n.Y_n \end{pmatrix}.$$

Notant $E = \begin{pmatrix} \mu_1 & 0 \\ 0 & E_n \end{pmatrix}$, de $E = Y^T.D.Y$ on déduit que :

$$\mu_1 = \lambda_1 y_{11}^2 + \vec{c}_2^T.D_n.\vec{c}_2, \quad Y_n^T.D_n.\vec{c}_2 = 0, \quad E_n = Y_n^T.D_n.Y_n.$$

Donc $D_n.\vec{c}_2 = 0$ (car Y_n est inversible), d'où $\mu_1 = \lambda_1 y_{11}^2$ et μ_1 et λ_1 ont même signe, et on a $E_n = Y_n^T.D_n.Y_n$, et on applique l'hypothèse de récurrence. \blacksquare

C Annexe. Point selle

C.1 Définition

On considère une fonction $\mathcal{M} : X \times \Lambda \rightarrow \mathbb{R}$ où X et Λ sont deux espaces vectoriels.

Définition C.1 Un point $(x, \lambda) \in X \times \Lambda$ est un point-selle de la fonction \mathcal{M} ssi :

$$\inf_{y \in X} \mathcal{M}(y, \lambda) = \mathcal{M}(x, \lambda) = \sup_{\mu \in \Lambda} \mathcal{M}(x, \mu), \quad (\text{C.1})$$

i.e. au point (x, λ) , $\mathcal{M}(x, \lambda)$ est "min en x " et est "max en λ ".

Exemple C.2 La fonction $\mathcal{M}(y, \mu) = y^2 - \mu^2$ a pour point selle le point $(0, 0)$. Faire un dessin. \blacksquare

Notons, pour $\mu \in \Lambda$ et pour $y \in X$:

$$F(y) \stackrel{\text{d\u00e9f}}{=} \sup_{\mu \in \Lambda} \mathcal{M}(y, \mu) \quad \text{et} \quad G(\mu) \stackrel{\text{d\u00e9f}}{=} \inf_{y \in X} \mathcal{M}(y, \mu). \quad (\text{C.2})$$

Alors (x, λ) est un point-selle de \mathcal{M} ssi :

$$G(\lambda) = F(x) \quad (= \mathcal{M}(x, \lambda)). \quad (\text{C.3})$$

D\u00e9finition C.3 La fonction G est appel\u00e9e la fonction duale de la fonction F .

Proposition C.4 Si $\mathcal{M} \in C^1(X \times \Lambda; \mathbb{R})$ admet un point selle, alors pour $\mu \in \Lambda$, on a :

$$G(\mu) = \mathcal{M}(y(\mu), \mu) \quad \text{pour } y(\mu) \text{ v\u00e9rifiant } \frac{\partial \mathcal{M}}{\partial x}(y(\mu), \mu) = 0,$$

et pour $y \in X$, on a :

$$F(y) = \mathcal{M}(y, \mu(y)) \quad \text{pour } \mu(y) \text{ v\u00e9rifiant } \frac{\partial \mathcal{M}}{\partial \lambda}(y, \mu(y)) = 0.$$

Preuve. A μ fix\u00e9, l'inf en y de $\mathcal{M}(y, \mu)$ est atteint en un point $y = \text{not\u00e9 } y(\mu)$ t.q. $\frac{\partial \mathcal{M}}{\partial x}(y, \mu) = 0$. M\u00eame d\u00e9marche pour F . \blacksquare

Exemple C.5 On prend $X = \mathbb{R}^n$, $\Lambda = \mathbb{R}^p$, $\mathcal{M} = \mathcal{L}$ (lagrangien) o\u00f9 :

$$\mathcal{L}(\vec{y}, \vec{\mu}) = \frac{1}{2} \vec{y}^T \cdot A \cdot \vec{y} + \vec{\alpha}^T \cdot \vec{y} + \vec{\mu}^T \cdot (B \cdot \vec{y} - \vec{c}),$$

o\u00f9 A est une matrice $n \times n$ sym\u00e9trique d\u00e9finie positive et B une matrice $p \times n$ de rang p , $\vec{\alpha} \in \mathbb{R}^n$ et $\vec{c} \in \mathbb{R}^p$. Donc \mathcal{L} est quadratique en \vec{y} et lin\u00e9aire en $\vec{\mu}$.

Soit $\vec{\mu} \in \mathbb{R}^p$ donn\u00e9. On a $G(\vec{\mu}) = \mathcal{L}(\vec{z}, \vec{\mu})$ en un point \vec{z} v\u00e9rifiant $\frac{\partial \mathcal{L}}{\partial \vec{x}}(\vec{z}, \vec{\mu}) = \vec{0} = A \cdot \vec{z} + B^T \cdot \vec{\mu} - \vec{\alpha}$.

Soit $\vec{y} \in \mathbb{R}^n$ donn\u00e9. On a $F(\vec{y}) = \mathcal{L}(\vec{y}, \vec{v})$ en un point \vec{v} v\u00e9rifiant $\frac{\partial \mathcal{L}}{\partial \lambda}(\vec{y}, \vec{v}) = \vec{c} = B \cdot \vec{y}$.

Donc s'il y a un point selle $(\vec{x}, \vec{\lambda})$ ce point selle doit v\u00e9rifier :

$$\begin{cases} A \cdot \vec{x} + B^T \cdot \vec{\lambda} = -\vec{\alpha}, \\ B \cdot \vec{x} = \vec{c}. \end{cases}$$

Comme A est sym\u00e9trique d\u00e9finie positive et B est de rang p , ce syst\u00e8me a une unique solution, c'est notre point selle.

Calcul de G : $\frac{\partial \mathcal{L}}{\partial \vec{x}}(\vec{y}, \vec{v}) = 0$ donne on a $\vec{y} = -A^{-1} \cdot B^T \cdot \vec{\mu} - A^{-1} \cdot \vec{\alpha}$, et donc $G(\vec{\mu}) = \mathcal{L}(-A^{-1} \cdot B^T \cdot \vec{\mu} - A^{-1} \cdot \vec{\alpha}, \vec{\mu})$, d'o\u00f9 G donn\u00e9e par (3.9), fonction quadratique concave sur \mathbb{R}^p .

Calcul de F : $\frac{\partial \mathcal{L}}{\partial \lambda}(\vec{y}, \vec{v}) = \vec{c} = B \cdot \vec{y}$, et donc $F(\vec{y})$ n'est r\u00e9el (n'est pas infini) qu'aux points \vec{y} dans l'espace affine $\vec{c} + \text{Ker} B$:

$$F(\vec{y}) = \frac{1}{2} \vec{y}^T \cdot A \cdot \vec{y} \quad \text{pour les } \vec{y} \text{ t.q. } B \cdot \vec{y} = \vec{c}, \quad (\text{C.4})$$

comme en (2.11). \blacksquare

C.2 “sup inf ≤ inf sup”

On se place dans $\bar{\mathbb{R}}$ pour ne pas exclure $\pm\infty$.

Proposition C.6 (“sup inf ≤ inf sup”). Dans $\bar{\mathbb{R}}$, quelle que soit la fonction $\mathcal{M} : X \times \Lambda \rightarrow \mathbb{R}$, on a toujours :

$$\forall x \in X, \forall \lambda \in \Lambda, \quad \inf_{y \in X} \mathcal{M}(y, \lambda) \leq \sup_{\mu \in \Lambda} \mathcal{M}(x, \mu), \quad (\text{C.5})$$

i.e. on a toujours “sup inf ≤ inf sup” :

$$\sup_{\mu \in \Lambda} \inf_{y \in X} \mathcal{M}(y, \mu) \leq \inf_{y \in X} \sup_{\mu \in \Lambda} \mathcal{M}(y, \mu). \quad (\text{C.6})$$

Ou encore, on a toujours :

$$\forall x \in X, \forall \lambda \in \Lambda, \quad G(\lambda) \leq F(x), \quad (\text{C.7})$$

i.e. on a toujours :

$$\sup_{\mu \in \Lambda} G(\mu) \leq \inf_{y \in X} F(y). \quad (\text{C.8})$$

Preuve. Pour (x, λ) donné, on a $\mathcal{L}(x, \lambda) \geq \inf_y \mathcal{L}(y, \lambda) = G(\lambda)$ et $\mathcal{L}(x, \lambda) \leq \sup_\mu \mathcal{L}(x, \mu) = F(x)$, d'où $G(\lambda) \leq F(x)$, ce pour tout $(x, \lambda) \in X \times \Lambda$, i.e. (C.7). D'où les autres relations. \blacksquare

Remarque C.7 Autre rédaction de la preuve pour démontrer directement (C.6) : il s'agit de montrer que :

$$\sup_{\mu \in \Lambda} (\inf_{y \in X} \mathcal{M}(y, \mu)) \leq \inf_{z \in X} (\sup_{\nu \in \Lambda} \mathcal{M}(z, \nu)) \quad \text{dans } \bar{\mathbb{R}}.$$

Par définition du sup et de l'inf, ce sera vrai si :

$$\forall \mu \in \Lambda, \quad \forall z \in X, \quad \inf_{y \in X} \mathcal{M}(y, \mu) \leq \sup_{\nu \in \Lambda} \mathcal{M}(z, \nu). \quad (\text{C.9})$$

Mais pas définition de l'inf et du sup, pour tout $\mu \in \Lambda$ et $z \in X$, on sait que :

$$\inf_{y \in X} \mathcal{M}(y, \mu) \leq \mathcal{M}(z, \mu) \quad \text{et} \quad \mathcal{M}(z, \mu) \leq \sup_{\nu \in \Lambda} \mathcal{M}(z, \nu).$$

D'où (C.9). \blacksquare

Exemple C.8 Soit $\mathcal{M}(y, \mu) = y^2 - \mu^2$. On pose alors $G(\mu) = \inf_{y \in \mathbb{R}} \mathcal{M}(y, \mu)$ et $F(y) = \sup_{\mu \in \mathbb{R}} \mathcal{M}(y, \mu)$. On a immédiatement :

$$G(\mu) = -\mu^2, \quad F(y) = y^2.$$

Et il est immédiat que $G(\mu) \leq F(y)$ pour tout $y, \mu \in \mathbb{R}$, car $\sup_{\mu \in \mathbb{R}} G(\mu) = 0 \leq 0 = \inf_{y \in \mathbb{R}} F(y)$. On a donc bien (C.6), et on a existence du point selle qui est $(0, 0)$. \blacksquare

C.3 L'inf-sup

Corollaire C.9 *Un point selle (x, λ) de \mathcal{M} (s'il existe) vérifie :*

$$\mathcal{M}(x, \lambda) = \sup_{\mu \in \Lambda} \inf_{y \in X} \mathcal{M}(y, \mu) = \inf_{y \in X} \sup_{\mu \in \Lambda} \mathcal{M}(y, \mu) \quad (\text{C.10})$$

(on peut inverser le sup et l'inf), et donc :

$$F(x) = G(\lambda) = \mathcal{M}(x, \lambda).$$

Et si $\mathcal{M} \in C^1(X \times \Lambda; \mathbb{R})$ alors on a :

$$\begin{cases} \frac{\partial \mathcal{M}}{\partial x}(x, \lambda) = 0, \\ \frac{\partial \mathcal{M}}{\partial \lambda}(x, \lambda) = 0. \end{cases} \quad (\text{C.11})$$

Preuve. Soit (x, λ) un point-selle ; alors :

$$\inf_y \sup_\mu \mathcal{M}(y, \mu) \leq \sup_\mu \underbrace{\mathcal{M}(x, \mu)}_{=F(x)} = \mathcal{M}(x, \lambda) = \inf_y \underbrace{\mathcal{M}(y, \lambda)}_{=G(\lambda)} \leq \sup_\mu \inf_y \mathcal{M}(y, \mu).$$

D'où "inf sup \leq sup inf" et la proposition C.6 donne l'inégalité inverse, d'où (C.10).

Puis, pour (x, λ) un point-selle, $\sup_\mu \mathcal{M}(x, \mu) = \mathcal{M}(x, \lambda)$, cf. (C.1), et Λ étant un espace vectoriel et \mathcal{M} étant C^1 , le sup est donné quand $\frac{\partial \mathcal{M}}{\partial \lambda}(x, \mu) = 0$ pour tout μ . Idem $\frac{\partial \mathcal{M}}{\partial x}(x, \lambda) = 0$. \blacksquare

C.4 Cas d'un lagrangien

On regarde le cas où $\mathcal{M} = \mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}$ est un lagrangien :

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \vec{\lambda}^T \cdot \vec{g}(\vec{x}), \quad (\text{C.12})$$

où $\vec{x} \in \mathbb{R}^n$ et $\vec{\lambda} \in \mathbb{R}^p$.

On note encore $P = \{\vec{x} : g(\vec{x}) = 0\}$.

Proposition C.10 Un point-selle $(\vec{x}, \vec{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p$ de \mathcal{L} (s'il en existe un) est tel que $f(\vec{x})$ réalise le minimum de f dans l'espace P :

$$\vec{g}(\vec{x}) = 0 \quad \text{et} \quad f(\vec{x}) = \min_{\vec{y} \in P} f(\vec{y}), \quad (\text{C.13})$$

i.e. en particulier \vec{x} est solution du problème primal (2.11).

Et si $f \in C^1(\mathbb{R}^n; \mathbb{R})$ et $\vec{g} \in C^1(\mathbb{R}^n; \mathbb{R}^p)$ alors $\mathcal{L} \in C^1(\mathbb{R}^n \times \mathbb{R}^p; \mathbb{R})$, et un point-selle $(\vec{x}, \vec{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^p$ de \mathcal{L} (s'il en existe un) est solution de :

$$\begin{cases} \vec{\nabla} f(\vec{x}) + \sum_{i=1}^p \lambda_i \vec{\nabla} g_i(\vec{x}) = 0, \\ g(\vec{x}) = 0. \end{cases} \quad (\text{C.14})$$

Preuve. Par définition d'un point-selle $(\vec{x}, \vec{\lambda})$, on a $\mathcal{L}(\vec{x}, \vec{\mu}) \leq \mathcal{L}(\vec{x}, \vec{\lambda})$ pour tout $\vec{\mu} \in \mathbb{R}^p$, i.e. :

$$\vec{\mu}^T \cdot \vec{g}(\vec{x}) \leq \vec{\lambda}^T \cdot \vec{g}(\vec{x}), \quad \forall \vec{\mu} \in \mathbb{R}^p,$$

i.e. $(\vec{\mu} - \vec{\lambda})^T \cdot \vec{g}(\vec{x}) \leq 0$ pour tout $\vec{\mu}$, i.e. $\vec{v}^T \cdot \vec{g}(\vec{x}) \leq 0$ pour tout \vec{v} . Ce n'est possible que si $\vec{g}(\vec{x}) = \vec{0}$, i.e. si $\vec{x} \in P$.

Puis, par définition d'un point-selle $(\vec{x}, \vec{\lambda})$, on a $\mathcal{L}(\vec{x}, \lambda) \leq \mathcal{L}(\vec{y}, \lambda)$ pour tout $\vec{y} \in \mathbb{R}^n$, donc en particulier pour tout $\vec{y} \in P$, et donc, comme on vient de voir que $\vec{x} \in P$:

$$f(\vec{x}) + 0 \leq f(\vec{y}) + 0, \quad \forall \vec{y} \in P.$$

Et donc \vec{x} est un minimum de f sur P .

Puis par dérivation en \vec{x} puis en $\vec{\lambda}$, on obtient (C.14). ▀

Références

- [1] Ciarlet P.G. : *Introduction à l'analyse matricielle et à l'optimisation*. Masson.
- [2] Faure : *Notes d'optimisation*. Polycopié de l'X, 1982.
- [3] Fortin M., Glowinski R. : *Méthodes de lagrangien augmenté*. Dunod, 1982.
- [4] Lascaux P., Théodor R. : *Analyse matricielle appliquée à l'art de l'ingénieur*.
- [5] Strang G. : *Linear Algebra and its applications*.