1) $J(x) = \| Ax - b \|^2$

$$\nabla J(x) = 2\,{}^tA(Ax - b)$$

2) $D^2 J = 2\,{}^tA A$

$\hookrightarrow$ Symétrique réelle $\underline{D.P}$ $(*)$

$(*)$ Car $\left({}^tx\,(2\,{}^tAA)\,x = 2\|Ax\|^2 \geq 0 \right.$ avec égalité ssi $x \in \mathrm{Ker}\,A = \{0\}$

donc J est elliptique.

$(\alpha = 2 \times \min Sp({}^tAA)$
$= 2\,(\sigma^*)^2$, avec σ^* plus petit valeur singulier de A

Decomp en Val Singulier :

$A = {}^tU \Sigma V$ $\qquad \Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ & & \sigma_p \end{pmatrix}$

$U \in M_n(\mathbb{R})$

$\Sigma \in M_{n,p}(\mathbb{R})$

$V \in M_p(\mathbb{R})$

les colonnes de U forment une famille $\underline{\underline{U}}$

V $\qquad \underline{\underline{U}}$

3)

$\left.\begin{array}{l} x \in \mathbb{R}^p \\ \mu \in \mathbb{R}_+ \end{array}\right\} L(x,\mu) = \|Ax - b\|^2 + \mu \langle c, x \rangle$

4) $\left\{ \begin{array}{l} J \text{ est elliptique}, \ C \text{ convexe fermé non vide} \\ C^1 \end{array} \right.$

$\Longrightarrow$ existence + unicité

```
function m = J(f)
    m = 0
    for i = 1 : (size(f,1) - 2)
        m = m + (f(i+2) - 2 * f(i+1) + f(i))^2
    end
endfunction
```

~~function G = gradJ~~

on a $J(f) = \sum_{i=1}^{n-2} \left(f(i+2) - 2 f(i+1) - f(i) \right)^2$

$$dJ(f) = \sum_{i=1}^{n-1} 2\left(f(i+2) - 2 f(i+1) + f(i) \right) \left(dx_{i+2} - 2 dx_{i+1}, dx_i \right)$$

```
function G = gradJ(f)

    G = zeros(size(f))
    for i = 1 : (size(f,1) - 2)
        d = 2 * (f(i+2) - 2 * f(i+1) + f(i));

        G(i+2) = G(i+2) + d;
        G(i+1) = G(i+1) - 2 * d;
        G(i) = G(i) + d;
endfunction.
```

Problème continu équivalent :

minimiser $J(f) = \int_0^1 f''(t)^2 \, dt$

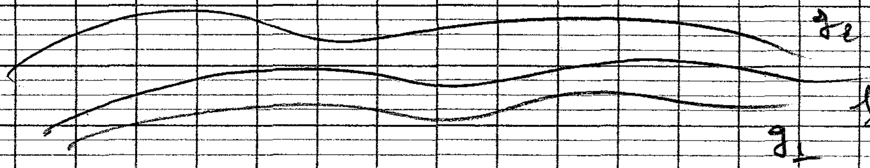
avec $g_1 \leq f \leq g_2$ $\quad g_1, f, g_2 : [0,1] \to \mathbb{R}$

Si on discrétise f par $f_d(i) = f(\frac{i}{n})$

$$f_d(i+2) - 2 f_d(i+1) + f_d(i)$$

$$= \underbrace{[f_d(i+2) - f_d(i+1)]}_{\simeq f'(\frac{i+1}{n}) \times \frac{1}{n}} - \underbrace{[f_d(i+1) - f_d(i)]}_{\simeq f'(\frac{i}{n}) \times \frac{1}{n}}$$

$$\simeq f''(\tfrac{i}{n}) \times \frac{1}{n} \times \frac{1}{n}$$



(P) $\quad \min_{x \in C} J(x)$

Lagrangien :

$$\underset{\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q}{L(x, \lambda, \mu)} = J(x) + \sum \lambda_i h_i(x) + \sum \mu_j g_j(x)$$

→ Problème

Exercice 24 :

$A \in M_{n,p}(\mathbb{R})$ de rang p $\quad n \geq p$

$rg(A) + \dim \ker A = p$

$\Rightarrow \dim \ker A = 0$

1) $J(x) = \| Ax - b \|^2$

$$\nabla J(x) = 2\, {}^tA (Ax - b)$$

2) $D^2 J = 2\, {}^tA A$

↳ Symétrique réelle. D.P (*)

(*) Car $\begin{cases} {}^t x (2\, {}^tAA) x = 2 \| Ax \|^2 \geq 0 \\ \text{avec égalité ssi } x \in \operatorname{Ker} A = \{0\} \end{cases}$

donc J est elliptique.

($\alpha = 2 \times \min \operatorname{Sp}({}^tAA)$
$= 2 (\sigma^*)^2$, avec σ^* plus petit valeur singulier de A

Decomp en Val Singulier :

$A = {}^tU \Sigma V$ $\qquad \Sigma = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_p \end{pmatrix}$

$U \in \mathcal{M}_n(\mathbb{R})$

$\Sigma \in \mathcal{M}_{n,p}(\mathbb{R})$

$V \in \mathcal{M}_p(\mathbb{R})$

le colonne de U forme une famille $\perp$

V $\perp$

3) $\left.\begin{array}{l} x \in \mathbb{R}^p \\ \mu \in \mathbb{R}_+ \end{array}\right\} L(x, \mu) = \| Ax - b \|^2 + \mu \langle c, x \rangle$

4) $\left\{\begin{array}{l} J \text{ est elliptique}, \ C \text{ convexe fermé non vide} \\ C^1 \end{array}\right.$

⟹ existence + unicité

Contrainte affine $\Rightarrow$ l'unique solut° $\bar{x}$ de P vérifie KKT

1) $\bar{\mu} \geq 0$

2) $\bar{\mu} \langle c, \bar{x} \rangle = 0$

3) $2\,{}^tA(A\bar{x} + b) + \bar{\mu} c = 0$

Cas 1) g_1 inactive, $\bar{\mu} = 0$

$$\bar{x} = ({}^tAA)^{-1}\,{}^tAb \quad \text{et} \quad \langle c, \bar{x} \rangle < 0$$

Cas 2 : g_1 active :

$$\begin{cases} 2\,{}^tAA\bar{x} + \bar{\mu} c = 2\,{}^tAb \\ \langle c, \bar{x} \rangle = 0 \end{cases}$$

$$\Leftrightarrow \begin{pmatrix} \boxed{2\,{}^tAA} & \boxed{c} \\ {}^tc & 0 \end{pmatrix} \begin{pmatrix} \boxed{\bar{x}} \\ \bar{\mu} \end{pmatrix} = \begin{pmatrix} \boxed{2\,{}^tAb} \\ 0 \end{pmatrix}$$

et $\bar{\mu} \geq 0$ (2)

Algorithme :

Calculer $\bar{x} = ({}^tAA)^{-1}\,{}^tAb$

Si $\langle c, \bar{x} \rangle < 0$ c'est fini

Sinon résoudre (2)

i.e. : $\begin{pmatrix} \bar{x} \\ \mu \end{pmatrix} = B^{-1} \begin{pmatrix} 2\,{}^tAb \\ 0 \end{pmatrix}$ et la solution est $\bar{x}$

4) Cas 1 :

$${}^tAA = \begin{pmatrix} 1 & -1 & 1 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 3 & 11 \end{pmatrix}$$

$${}^tAb = \begin{pmatrix} -6 \\ -4 \end{pmatrix}$$

$\bar{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ $\quad \begin{cases} 3x + 3y = -6 \\ 3x + 11y = -4 \end{cases} \Rightarrow \begin{cases} y = 1/4 \\ x = -9/4 \end{cases}$

$\langle c, x \rangle = \frac{19}{4} \geq 0$ contradict°

Cas 2 :

$$\begin{cases} 3x + 3y - \mu = -6 \\ 3x + 11y + \frac{\mu}{2} = -4 \\ -2x + y = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x = -14/59 \\ y = -28/59 \\ \mu = \frac{228}{59} \geq 0 \end{cases}$$

5)

Function x = minimix_LxT (A, b, c)

linsolve (A'*A , - A'*b)

if c'* x ≥ 0 then
B = [2*A'*A, c ; c', 0]; d = [2*A'*b;0]

x = linsolve (B, - d)
x ($) = [0]
└> dernière coordonnée

end
endfunction;

(*) $\underset{x \in \mathbb{R}^p}{\text{Min}} \quad \|Ax - b\|^2 + \mu \langle c, x \rangle$

$\nabla_x = 2\,{}^tA(Ax - b) + \mu c = 0$

$x = (2\,{}^tAA)^{-1}(2\,{}^tAb - \mu c)$

Uzawa

initialisation $\mu^0 = 0$

itéré pour $k = 0, 1 \ldots$

$$x^k = (2\,{}^tAA)^{-1}(2\,{}^tAb - \mu^k c)$$
$$\mu^{k+1} = \max(0, \mu^k + \rho \langle c, x^k \rangle)$$

Converge pour $0 < \rho < \dfrac{2 \times \min \mathrm{Sp}({}^tAA) \times 2}{M_g^2 = \|c\|^2}$

$(\text{car } \|g(x) - g(y)\| \leq \|c\|.\|x-y\|$

+ function x = minimux Uzawa (A, b, c, rho, N)

```
    mu = 0
    B = inv (2 * A' * A)
    d = 2 * A' * b
    for i = 1 : N
        x = B * (d - mu * c)
        mu = max (0, mu + rho * c' * x)
    end
endfunction
```

8)

+ En utilisant KKT
il faut 2^{100} for le test (if --- end)
pas traitable.

+ Pour UZAWA
l'algorithme est le même mais

$$mu' = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_{100} \end{pmatrix}$$

(P) Minimiser $J(x)$
$x \in C$

Lagragien : $L(x, \lambda, \mu) = J(x) + \sum \lambda_i h_i(x) + \sum \mu_j g_j(x)$

- Pas dual

(P^*) $\underset{(\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q_+}{\text{Max}} \varphi(\lambda, \mu)$ où $\varphi(\lambda, \mu) = \underset{x \in \mathbb{R}^n}{\text{Min}} L(x, \lambda, \mu)$

Expl : $c, s \in \mathbb{R}^n$, $s \neq 0$

Minimiser $J(x) = \frac{1}{2}\|x\|^2 + \langle c, x \rangle$

Sous la contrainte $\langle s, x \rangle \leq 0$ $p = 0$, $q = 1$, $g_1(x) = \langle s, \cdot$

$L(x, \mu) = \frac{1}{2}\|x\|^2 - \langle c, x \rangle + \mu \langle s, x \rangle$

$\varphi(\mu) = L(x^*, \mu)$
où x^* vérifie $\nabla_x(x^*, \mu) = 0$

$x \mapsto L(x, \mu)$ est convexe et elliptique
donc admet un unique minimiseur sur $\mathbb{R}^n$ caractérisé par $\nabla_x L(x, \mu) = 0$

$\nabla_x L(x, \mu) = x - c + \mu s$
donc $x^* = c - \mu s$

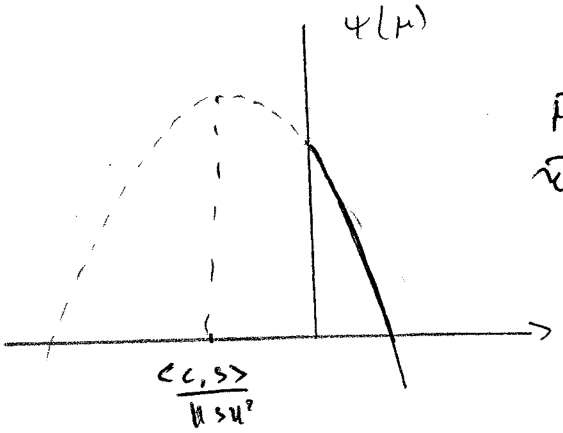
$\varphi(\mu) = -\frac{1}{2}\|c - \mu s\|^2$ concave

$\varphi(\mu) = -\frac{1}{2}\langle c - \mu s, c - \mu s \rangle = \frac{1}{2}(\|c\|^2 - 2\mu\langle s, c \rangle + \mu^2\|s\|^2)$

$\varphi'(\mu) = -\langle c - \mu s, -s \rangle$

$\varphi'(\mu) = 0 \Leftrightarrow \mu = \frac{\langle c, s \rangle}{\|s\|^2}$

1^{er} cas : $\langle c, s\rangle \leq 0$

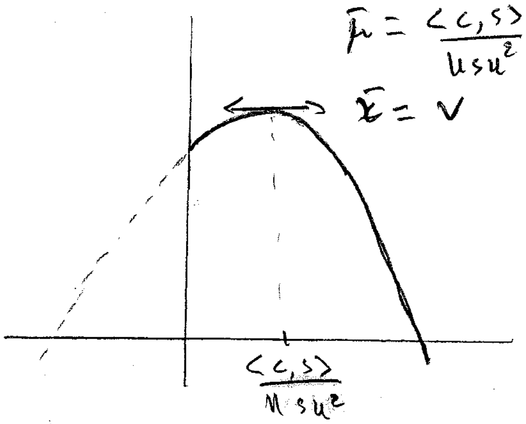


$\bar{\mu} = 0$

$\bar{x} = c - \bar{\mu} s = c$

$\max\limits_{\mu \in \mathbb{R}_+} \psi(\mu) = \psi(0) = -\frac{1}{2}\|c\|^2$

2^{ie} cas : $\langle c, s\rangle > 0$



$\bar{\mu} = \frac{\langle c,s\rangle}{\|s\|^2}$

$\bar{x} = v$

$\max\limits_{\mu \in \mathbb{R}^+} \psi(\mu) = \psi\left(\frac{\langle c,s\rangle}{\|s\|^2}\right)$

$= -\frac{1}{2}\|\underbrace{c - \frac{\langle c,s\rangle}{\|s\|^2} s}_{v}$

Si on veut $\underset{x \in C}{\operatorname{argmin}}\, J(x)$

on calcule $\underset{x \in \mathbb{R}^n}{\operatorname{argmin}}\, L(x, \bar{\mu}) = \bar{x}$ où $\bar{\mu} = \operatorname{argmin}$

or $\operatorname{argmin} L(\hat{x}, \mu) = \overbrace{c - \mu s}^{x^*}$

- Comparaison avec la résolution de (P) par KKT

$\exists\, \bar{\mu}$ tq
1) $\bar{\mu} \geq 0$
2) $\bar{\mu}\langle s, \bar{x}\rangle = 0$
3) $\nabla J(\bar{x}) + \bar{\mu}\, g_1(\bar{x}) = 0$

i.e $\bar{x} - c + \bar{\mu} s = 0$

1^{ie} cas : g_1 inactive $(\langle s, \bar{x}\rangle < 0)$

alors $\bar{\mu} = 0$ et $\bar{x} = c$

Compatible si $\langle c, \bar{x}\rangle < c$

2^{ie} cas : g_1 active $\begin{cases} \langle s, \bar{x}\rangle = 0 \\ \bar{x} = c - \bar{\mu} s \end{cases}$

$0 = \langle \bar{x}, s\rangle = \langle c, s\rangle - \bar{\mu}\|s\|^2 \Rightarrow \bar{\mu} = \frac{\langle c,s\rangle}{\|s\|^2}$

et $\bar{x} = c - \frac{\langle c,s\rangle}{\|s\|^2} s$

non contradictoire si $\bar{\mu} \geq 0$, i.e $\langle c, s\rangle \geq$

Algorithme d'Uzawa

Principe :

On suppose que l'on sait calculer explicitement $\psi(\lambda,\mu)$ et on resout le probleme dual (P^*) par la methode du gradient Projeté $\left(\max_{(\lambda,\mu)\in\mathbb{R}^p\times\mathbb{R}^q} \psi(\lambda,\mu)\right)$

algorithme :

- initialisation : (λ^0,μ^0) quelconque ds $\mathbb{R}^p\times\mathbb{R}^q$ (en pratique souvent $\lambda^0=0$, $\mu^0=0$)

$\rho>0$ • iterer pour $k=0,1,\dots$

1) $x^k = \operatorname{argmin}_{x\in\mathbb{R}^n} L(x,\lambda^k,\mu^k)$

2) $\begin{cases} \lambda_i^{k+1} = \lambda_i^k + \rho\, h_i(x^k) & \text{pour } 1\le i\le p \\ \mu_j^{k+1} = \max\left(0, \mu_j^k + \rho\, g_j(x^k)\right) & \text{pour } 1\le j\le q \end{cases}$

$\mu^{k+1} = \Pi_{\mathbb{R}}$

interpretation : gradient projeté appliqué à la maximisation de ψ

etape 2) $(\lambda,\mu)^{k+1} = \Pi_{\mathbb{R}^p\times\mathbb{R}^q_+}\left((\lambda,\mu)^k + \rho\nabla_{\lambda,\mu} L(x^k,\lambda^k,\mu^k)\right)$

Rappel :

$$\Pi_{\mathbb{R}^p\times\mathbb{R}^q_+}(\lambda_1,\dots\lambda_p,\mu_1,\dots\mu_q) = \begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_p\\ \mu_1^+\\ \vdots\\ \mu_q^+\end{pmatrix} \quad \text{où } x^+ = \max(x,0)$$

Théorème :

- J C^1, h_i affines, g_j C^1 convexes
  J α-elliptique, $h=\begin{pmatrix}h_1\\ \vdots\\ h_p\end{pmatrix}$ M_h-lipschitz, $g=\begin{pmatrix}g_1\\ \vdots\\ g_q\end{pmatrix}$ M_g-lipschi[tz]

Si le Lagrangien L associé à (P) admet un point selle $(\bar{x},\bar{\lambda},\bar{\mu})$ alors on a $x^k \xrightarrow[k\to+\infty]{} \bar{x}$ dès que $0<\rho<\frac{2\alpha}{M_h^2+M_g^2}$

($\bar{x}$ est solut° de (P))

Preuve:

1) $(\bar{x}, \bar{\lambda}, \bar{\mu})$ point selle $\Rightarrow$ $(\bar{x}, \bar{\lambda}, \bar{\mu})$ verifie KKT i), ii), iii)

donc $\bar{\lambda}_i = \bar{\lambda}_i + \rho\, h_i(\bar{x}) \quad \forall i$

(1) $\bar{\mu}_j = \max(0, \bar{\mu}_j + \rho\, g_j(\bar{x})) \quad \forall j$

car soit $\bar{\mu}_j = 0$ et $\bar{\mu}_j + \rho\, \underbrace{g_j(\bar{x})}_{\leq 0} \leq 0$

soit $\bar{\mu}_j \geq 0$ et $\bar{\mu}_j(\bar{\mu}_j + \rho\, g_j(\bar{x})) = \bar{\mu}_j^2$

donc $\bar{\mu}_j \max(0, \bar{\mu}_j + \rho\, g_j(\bar{x})) = \bar{\mu}_j^2 \Rightarrow (1)$

2) $\cancel{\|\lambda^{k+1} - \bar{\lambda}\|^2 = \sum_{i=1}^{p} \rho^2 (h_i(x^k) - h_i(\bar{x}))^2 = \rho^2 \|h_i(x^k) - h(\bar{x})\|^2 \leq \rho^2 M_h^2 \|x^k - \bar{x}\|^2}$

on sait que:

$$\begin{cases} \mu^{k+1} = \Pi_{\mathbb{R}^q_+}(\mu^k + \rho\, g(x^k)) \\ \bar{\mu} = \Pi_{\mathbb{R}^q_+}(\bar{\mu} + \rho\, g(\bar{x})) \end{cases}$$

(C) $\|\mu^{k+1} - \bar{\mu}\|^2 \underset{\Pi_{\mathbb{R}^q_+}\ 1\text{-lipsch}}{\leq} \|\mu^k - \bar{\mu} + \rho(g(x^k) - g(\bar{x}))\|^2$

$\leq \|\mu^k - \bar{\mu}\|^2 + 2\rho \langle \mu^k - \bar{\mu}, g(x^k) - g(\bar{x}) \rangle + \rho^2 M_g^2 \|x^k - \bar{x}\|^2$

pareille pour les λ:

(D) $\|\lambda^{k+1} - \bar{\lambda}\|^2 \leq \|\lambda^k - \bar{\lambda}\|^2 + 2\rho \langle \lambda^k - \bar{\lambda}, h(x^k) - h(\bar{x}) \rangle + \rho^2 M_h^2 \|x^k - \bar{x}\|^2$

3) $\forall y \quad L(y, \lambda^k, \mu^k) \geq L(x^k, \lambda^k, \mu^k)$

$$J(y) - J(x^k) + \sum_{i=1}^{p} \lambda_i^k (h_i(y) - h_i(x^k)) + \sum_{j=1}^{q} \mu_j^k (g_j(y) - g_j(x^k)) \geq 0$$

$y = x^k + t(x - x^k) = tx + (1-t)x^k \Rightarrow g_j(y) \leq t g_j(x) + (1-t) g_j(x^k)$

$\Rightarrow g_j(y) - g_j(x^k) \leq t(g_j(x) - g_j(x^k))$

egalité pour h_i (car affine de convexe) donc $\dfrac{J(x^k+t(x-x^k))-J(x^k)}{t} + \sum_{i=1}^{p} \lambda_i^k (h_i(x)-h_i(x^k)) + \sum_{j=1}^{q} \mu_j^k (g_j(x)-g_j(x^k))$

(A) $t \to 0^+$ $\quad \langle \nabla J(x^k), x-x^k \rangle + \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \geq 0$

on a pareille pour $\bar{x}$ car $\forall y \; L(y, \bar{\lambda}, \bar{\mu}) \geq L(\bar{x}, \bar{\lambda}, \bar{\mu})$

(B) donc $\langle \nabla J(\bar{x}), x-\bar{x} \rangle + \sum_{i=1}^{p} \bar{\lambda}_i (h_i(x)-h_i(\bar{x})) + \sum_{j=1}^{q} \bar{\mu}_j (g_j(x)-g_j(\bar{x}))$;

Prenons dans (A) $x=\bar{x}$ et $x=x^k$ dans (B) et

Sommons : (A) + (B)

ona : $-\underbrace{\langle \nabla J(x^k)-\nabla J(\bar{x}), \bar{x}-x^k \rangle}_{\geq \alpha \|x^k-\bar{x}\|^2} \leq \langle \lambda^k-\bar{\lambda}, h(\bar{x})-h(x^k) \rangle + \langle \mu^k-\bar{\mu}, g(\bar{x})-g(x^k) \rangle$

(car α elliptique)

(C) + (D) $\Rightarrow$ $\quad S_{k+1} \leq S_k + \|x^k-\bar{x}\|^2 \, \varepsilon \big(\underbrace{\varepsilon (M_k^2+M_g^2)-2\alpha}_{<0} \big)$

$(S_k = \|\lambda^k-\bar{\lambda}\|^2 + \|\mu^k-\bar{\mu}\|^2)$

$S_{k+1} \leq S_k$ donc $S_k \searrow$ or $S_k \geq 0$

donc S_k converge

$S_{k+1} - S_k \leq \|x^k-\bar{x}\|^2 \, \varepsilon (\varepsilon (M_k^2+M_g^2)-2\alpha) \leq 0$

donc $S_{k+1} - S_k \to 0$

donc $\|x^k-\bar{x}\| \to 0$