Exercice 14

L'égalité matricielle $Dx = v$ est une manière ramassée d'écrire r contraintes linéaires $L_i x = v_i$ où L_i est le $i^{ème}$ vecteur ligne de D et v_i la $i^{ème}$ coordonnées de v.

Les gradients des fonctions $x \longmapsto L_i x - v_i$ qui définissent les contraintes sont les vecteurs L_i

Par hypothèse D est de rang r donc les r vecteurs L_i sont linéairement indépendants : la condition du théorème des extrema liés (à plusieurs contraintes (p 82)) est satisfaite.

En un point de l'ensemble $\{x / Dx = v\}$ où J est extremale il existe des nombres $\lambda_1 \ldots \lambda_r$ tels que

$$\nabla J = \sum_{i=1}^{r} \lambda_i L_i \quad (*)$$

Une autre façon d'écrire la somme $\sum_{i=1}^{r} \lambda_i L_i$ est

$$\sum_{i=1}^{r} \lambda_i L_i = (\lambda_1 \ldots \lambda_r) D$$
$$= \lambda D \qquad \text{où } \lambda = (\lambda_1 \ldots \lambda_r)$$

Calculons maintenant le gradient de J. Pour cela écrivons le développement limité de J en un point x

Comme nous avons aussi de la matrice hessienne de J

faisons un développement à l'ordre 2.

$$J(x+h) = \frac{1}{2}\langle A(x+h), x+h\rangle - \langle b, x+h\rangle$$

$$= \frac{1}{2}\langle Ax+Ah, x+h\rangle - \langle b, x\rangle - \langle b, h\rangle$$

$$= \frac{1}{2}\left(\langle Ax, x\rangle + \langle Ax, h\rangle + \langle Ah, x\rangle + \langle Ah, h\rangle\right) - \langle b, x\rangle - \langle b, h\rangle$$

$\langle h, Ax\rangle$ car A est symétrique

$$= \underbrace{\left(\frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle\right)}_{= J(x)} + \frac{1}{2}\left(\langle Ax, h\rangle + \langle Ah, x\rangle\right) - \langle b, h\rangle + \frac{1}{2}\langle Ah, h\rangle$$

$$= J(x) + \underbrace{\langle Ax - b, h\rangle}_{\text{linéaire en } h} + \frac{1}{2}\underbrace{\langle Ah, h\rangle}_{\text{quadratique en } h}.$$

Le DL à l'ordre 2 de J en x s'écrit de manière formelle

$$J(x+h) = J(x) + \nabla J(x).h + \frac{1}{2}\,{}^t h\, \mathrm{Hess}\, J(x)\, h + o(\|h\|^2)$$

Par unicité du développement limité on en déduit

$$\nabla J(x) = {}^t(Ax - b).$$

$$\mathrm{Hess}\, J(x) = A.$$

Le théorème des extrema liés donne donc le système

$$\begin{cases} {}^t(Ax-b) = \lambda D & (1) \quad (\text{c'est } (*)) \\ Dx = v & (2) \quad (\text{c'est dire que la contrainte est respectée}) \end{cases}$$

Transformons un peu (1)

$$ {}^t(Ax-b) = \lambda D \iff Ax - b = {}^tD\,{}^t\lambda $$

$$ \iff Ax = {}^tD\,{}^t\lambda + b $$

$$ \iff x = A^{-1}\,{}^tD\,{}^t\lambda + A^{-1}b \quad (***) \quad \left(\text{car } A \text{ est inversible déf.} > 0\right) $$

En multipliant cette dernière égalité par D on obtient

$$ Dx = DA^{-1}\,{}^tD\,{}^t\lambda + DA^{-1}b $$

Mais (2) assure alors

$$ v = DA^{-1}\,{}^tD\,{}^t\lambda + DA^{-1}b $$

soit $\quad DA^{-1}\,{}^tD\,{}^t\lambda = v - DA^{-1}b \quad (**)$

La matrice $DA^{-1}\,{}^tD$ est définie positive (donc inversible)

[En effet $\langle DA^{-1}\,{}^tD w, w\rangle$

$$ = \langle A^{-1}\,{}^tD w, {}^tD w\rangle $$

et ce dernier produit est ≥ 0 et nul seulement si ${}^tD w = 0$

car A^{-1} est définie positive comme inverse de A définie positive.

Or comme D est de rang n, ${}^tD w = 0$ si et seulement si $w = 0$]

On déduit donc de $(**)$ que

$$ {}^t\lambda = (DA^{-1}\,{}^tD)^{-1} v - (DA^{-1}\,{}^tD)^{-1} DA^{-1}b $$

En remplaçant ${}^t\lambda$ par cette valeur dans (***) on obtient

$$x = A^{-1}b + A^{-1}{}^tD(DA^{-1}{}^tD)^{-1}v - A^{-1}{}^tD(DA^{-1}{}^tD)^{-1}DA^{-1}b$$

C'est l'unique point où le système fourni par le théorème des extrema liés est satisfait..

La hessienne de J est A, définie positive, donc J est une fonction strictement convexe. La restriction de J à l'ensemble convexe (c'est un sous espace affine) $\{x / Dx = v\}$ est elle aussi convexe et elle admet un minimum là où les conditions du théorème des extrema liés sont satisfaites.

<u>Conclusion</u>: La fonction J admet un unique minimum sur l'ensemble $\{x / Dx = v\}$ atteint au point

$$x = A^{-1}b + A^{-1}{}^tD(DA^{-1}{}^tD)^{-1}v - A^{-1}{}^tD(DA^{-1}{}^tD)^{-1}DA^{-1}b.$$

Si on remplace la condition $Dx = v$ par la condition $Dx \leq v$, alors plusieurs cas se présentent.

* Si l'unique point critique de J ($A^{-1}b$) appartient à $Dx \leq v$ (autrement dit si $DA^{-1}b \leq v$) alors le minimum de J sur $\{x / Dx \leq v\}$ est le minimum de J sur $\mathbb{R}^d$ (atteint en $A^{-1}b$).

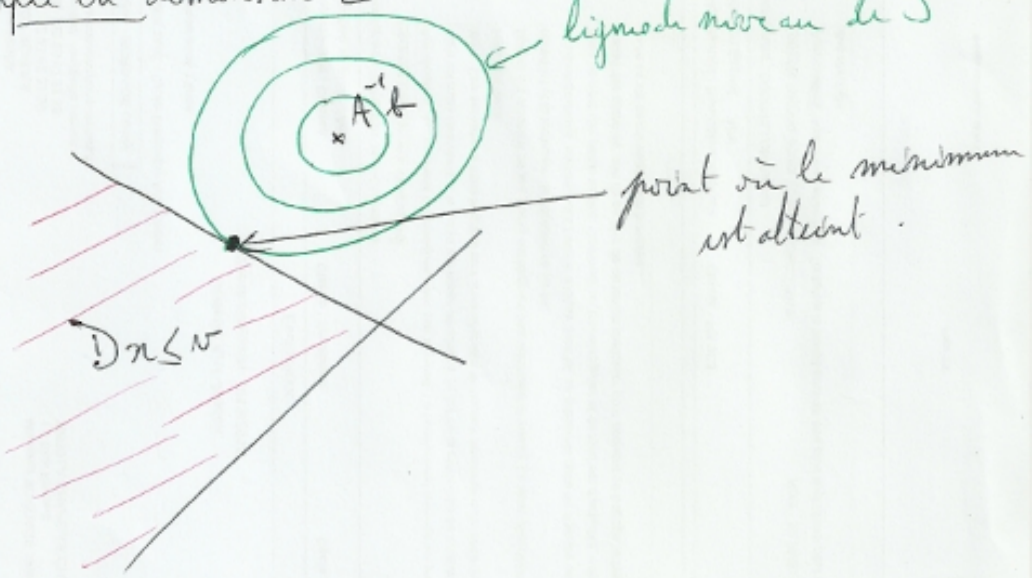
* Sinon il faut déterminer en quel point du bord de $\{x / Dx \leq v\}$ les (hyper) surfaces de niveau de J qui sont des ellipsoïdes centrés en $A^{-1}b$ touchent $\{x / Dx \leq v\}$ pour la première fois. Cela peut-être en un point de $Dx = v$ (auquel cas c'est un point trouvé dans la première partie) mais pas nécessairement. C'est en un point de la forme

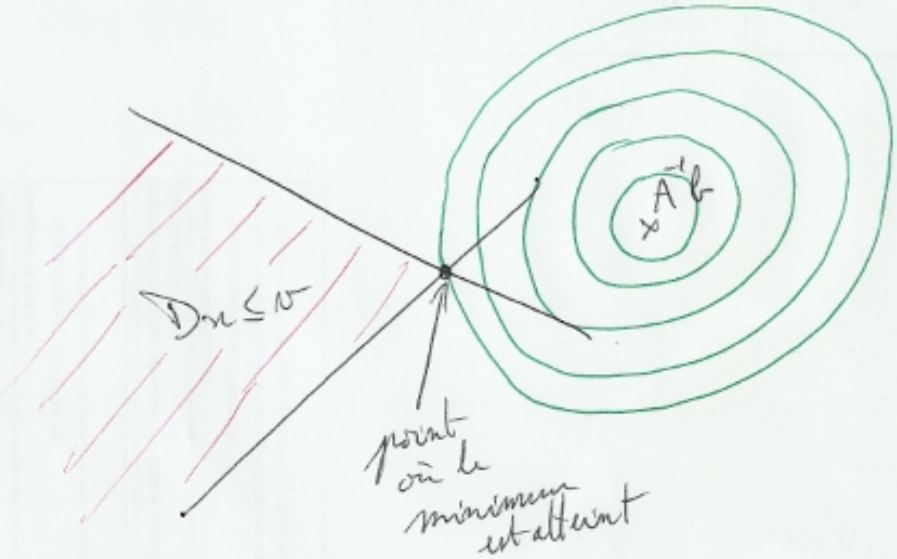
$$
\begin{aligned}
L_1 x &= v_1 \\
L_2 x &< v_2 \\
L_3 x &= v_3 \\
L_4 x &= v_4 \\
&\vdots \\
L_n x &< v_n
\end{aligned}
$$

où on peut distribuer $=$ ou $<$ de toutes les façons possibles. (il y en a 2^n).

Ici $n=d=2$

Exemple en dimension 2





- Si $A^{-1}b$ appartient à $\{x / Dx \leq v\}$ alors c'est en ce point que le minimum est atteint

- Il y a une quatrième possibilité (tangence sur la demi-droite)

Une justification du fait que la restriction de J à Dx=v a bien un minimum au point trouvé.

CENTRE
HENRI LEBESGUE
CENTRE DE MATHÉMATIQUES

La hessienne de J est A, définie positive, donc J est strictement convexe. La restriction de J à l'ensemble $Dx = v$ (sous-espace affine de dimension $d-r$ de $\mathbb{R}^d$) est aussi strictement convexe.

Si x et $x+h$ vérifie tous les deux $Dx = v$, $D(x+h) = v$ on a $Dh = 0$.

Le développement de Taylor de J à l'ordre 2 s'écrit

$$J(x+h) = J(x) + \langle Ax - b, h \rangle + \frac{1}{2} \langle Ah, h \rangle.$$

Si (1) est vérifiée $Ax - b = {}^tD\lambda$ le produit scalaire $\langle Ax-b, h \rangle$ s'écrit

$$\langle Ax-b, h \rangle = \langle {}^tD\lambda, h \rangle = \langle \lambda, Dh \rangle = 0$$

si $Dh = 0$.

Pour x et $x+h$ tq $Dx = v = D(x+h)$ on a donc

$$J(x+h) = J(x) + \frac{1}{2} \langle Ah, h \rangle.$$

Le gradient de la restriction de J à $Dx = v$ est donc nul en un point où (1) et (2) sont vérifiées.