Agrégation. Épreuve de Modélisation : option Calcul scientifique

Suggestions d'illustrations de la notion de conditionnement sur les trois problèmes ci-dessous.

1- **Résolution des systèmes linéaires.**
   1.1- Comment Maple calcule le conditionnement d'une matrice.
   1.2- Idem avec Matlab.
   1.3- Mettre en œuvre la méthode de la puissance voir par exemple L'ouvrage de N. J. Higham, Accuracy and Stability of Numerical Algorithm.
   1.4- Montrer les variations du conditionnement avec la dimension sur des classes de matrices.

2- **Calcul d'une racine simple d'un polynôme d'une variable**
   2.1- Comment Matlab calcule des racines de polynômes.
   2.2- Illustrer le théorème de Rouché.

3- **Problème des valeurs et vecteurs propres.**
   3.1 - Comment Matlab calcule des valeurs propre et des vecteurs propres.
   3.2- Illustrer le théorème 5.1.

Conditionnement : le problème des erreurs en analyse numérique.

1. **Qu'est-ce qu'un problème.**

Soient E et F deux ensembles. L'ensemble E est l'ensemble des données et F l'ensemble des sorties. Un problème est un élément du sous ensemble $G \subset E \times F$.

On donne comme exemples les trois problèmes suivants

1.1. **Résolution des systèmes linéaires.** Ici $E = \mathcal{M}_{n\times n}(\mathbb{C}) \times \mathbb{C}^n$ et $F = \mathbb{C}^n$. Un problème est défini par

$$G = \{(A, b, x) \in E \times F \ : \ Ax - b = 0 \ \}.$$

1.2. **Racine d'un polynôme.** Ici $E = \mathbb{C}_d[x]$ et $F = \mathbb{C}$. Un problème est défini par

$$G = \{(f, x) \in E \times F \ : \ f(x) = 0 \ \}.$$

1.3. **Problèmes des valeurs propres et des vecteurs propres.** Ici $E = \mathcal{M}_{n\times n}(\mathbb{C}^n)$ et $F = \mathbb{C}^n \times \mathbb{C}$. Un problème est défini par $G = \{(A, b, x, \lambda) \in E \times F \ : \ Ax - \lambda x, ||x|| = 1 \ \}$.

1.4. **Modélisation abstraite dans cadre vectoriel.** On se limite au cadre abstrait d'espace vectoriel de dimension finie. En fait dans les exemples ci-dessus, G est un ensemble qui est défini par :

$$G = \{(x, y) \in E \times F \ : \ f(x, y) = 0\}$$

où

$$f \ : \ (x, y) \in E \times F \to f(x, y) \in H.$$

Dans les trois exemples ci-dessus la fonction f est analytique et on supposera ceci dans la suite. On suppose que E,F et H sont des espaces hermitiens (euclidiens) de dimension finie:

$$m = dim\ H = dim\ F \leq n = dim\ E.$$

Chaque espace est munie d'une norme. On notera indifférement par $||x||$ la norme d'un élément ou par $||A||$ la norme d'un opérateur (multi-)linéaire. Le contexte est suffisament clair pour éviter toute confusion.

2. Analyse directe des erreurs

Il s'agit du problème suivant : quelle perburbation sur la sortie y produit une perturbation de la donnée x?

La réponse est donnée par le théorème des fonctions implicites.

Théorème 2.1. *Soit une application f analytique définie d'un ouvert O de $E \times F$ dans G.*

Soit $(x_0, y_0) \in O$ tel que

1- $f(x_0, y_0) = 0$,
2- $D_2f(x_0, y_0)$ *soit un isomorphisme (rang de* $D_2f(x_0, y_0) = m$*.)*

Alors il existe un ouvert U_{x_0} *et une seule application* φ *analytique définie sur un ouvert* $U_{x_0} \subset E$ *dans un ouvert* $V_{y_0} \subset F$ *vérifiant :*

1- $\varphi(x_0) = y_0$
2- *Pour tout* $x \in U_{x_0}$ *on a* $(x, \varphi(x)) \in U_{x_0} \times V_{y_0}$ *et* $f(x, \varphi(x)) = 0$.
3- $\varphi'(x) = -D_2f(x, \varphi(x))^{-1} D_1f(x, \varphi(x))$.

L'application de ce théorème pour quantifier l'influence d'une perturbation de la donnée x sur la sortie y est la suivante. On écrit

$$\begin{aligned} y - y_0 = \varphi(x) - \varphi(x_0) &= D\varphi(x_0)(x - x_0) + o(x - x_0) \\ &= -D_2f(x_0, y_0)^{-1} D_1f(x_0, y_0)(x - x_0) + o(x - x_0). \end{aligned}$$

Au premier ordre on a

$$||y - y_0|| \leq ||D_2f(x_0, y_0)^{-1} D_1f(x_0, y_0|| \, ||x - x_0|| + o(||x - x_0||).$$

On peut maintenant préciser la notion de conditionnement d'un problème.

Définition 2.2. *Soit* G *un problème et* $(x, y) \in G$.

1- *Un problème est dit bien posé si* $D_2f(x, y)^{-1}$ *existe.*
2- *L'opérateur* $K(f; x, y) = -D_2f(x, y)^{-1} D_1f(x, y)$ *est appelé l'opérateur de conditionnemment.*
3- *Le conditionnement* $C(f; x, y)$ *est la norme de* $K(f; x, y)$.

Cette définition du conditionnement est celle de l'erreur absolue au premier ordre. Dans les applications, c'est la notion d'erreur relative qui est significative. Il s'agit dans ce cas de quantifier $\dfrac{||y - y_0||}{||y_0||}$ à l'aide de $\dfrac{||x - x_0||}{||x_0||}$.

3. Conditionnement de la résolution des systèmes linéaires

Soit $f : (A, b, x) \in \mathcal{M}_{n \times n}(\mathbb{C}) \times \mathbb{C}^n \times \mathbb{C}^n \to Ax - b \in \mathbb{C}^n$.

Définition 3.1. *Le conditionnement d'une matrice est la quantité*

$$K(A) = ||A|| \, ||A^{-1}||.$$

On va voir que le conditionnement d'une matrice est relié à l'erreur relative d'un problème de résolution d'un système linéire.

Théorème 3.2. *Le problème de la résolution linéaire est bien posé si* A *est inversible. Supposons* A *inversible et considérons une matrice* M *telle que* $||M - A|| \leq \dfrac{1}{||A^{-1}||}$.

1- M *est inversible et* $||M^{-1}A|| \leq \dfrac{1}{1 - ||A^{-1}|| \, ||M - A||}$.

2- *Soient* (A,b,x) *et* (M,p,y) *tels que* $Ax-b=My-p=0$. *Alors*
$$\frac{||y-x||}{||x||} \le \frac{K(A)}{1-K(A)\frac{||M-A||}{||A||}} \left(K(A)\frac{||M-A||}{||A||}+\frac{||p-b||}{||b||}\right).$$

Preuve. Le problème est bien posé si $D_2f(A,b,x)$ est inversible. Or $D_2f(A,b,x)=A$.

Montrons que M est inversible. On a $A^{-1}M=I+A^{-1}(M-A)$.Donc $||A^{-1}M|| \ge 1-||A^{-1}||\,||M-A||>0$. On en déduit le 1.

Majorons l'erreur relative. Puisque $My-p=Ax-b=0$ on a

$$\begin{aligned} 0=My-p &= (A+M-A)(x+y-x)-b-(p-b) \\ &= Ax-b \quad +(M-A)x-(p-b)+A(y-x) \quad +(M-A)(y-x) \\ &= (M-A)x-(p-b)+M(y-x). \end{aligned}$$

Donc $y-x=M^{-1}A\,A^{-1}\left((M-a)x-(p-b)\right)$. D'autre part $Ax=b$ et $||A||\,||x|| \le ||b||$. On a donc

$$\begin{aligned} \frac{||y-x||}{||x||} &\le ||M^{-1}A|| \quad ||A^{-1}||\,||A|| \left(||M-A||\,||A^{-1}||+\frac{||p-b||}{||b||}\right) \\ &\le \frac{K(A)}{1-K(A)\frac{||M-A||}{||A||}} \left(K(A)\frac{||M-A||}{||A||}+\frac{||p-b||}{||b||}\right). \end{aligned}$$

□

4. Conditionnement d'une racine d'un polynôme

Dans cette partie p et q désignent des polynômes de degré d et x et y des nombres complexes. On note $|x|$ la norme de x et

$$|x|_1=(1+|x|^2)^{1/2}.$$

Lemme 4.1.

1- *L'application*

$$(p,q)\to\sum_{k=0}^{d}\binom{d}{k}^{-1}p_k\bar{q}_k$$

est un produit scalaire. On le note $<p,q>$. *Soit* $||p||$ *la norme associée.*

2- *Pour tout* k *on a*

$$\frac{p^{(k)}(x)-q^{(k)}(x)}{k!}=\binom{d}{k}<p-q,(1+z\bar{x})^{d-k}>.$$

3- $$\frac{\left|p^{(k)}(x)-q^{(k)}(x)\right|}{k!} \le \binom{d}{k}||p-q||\,|x|_1^{d-k}.$$

Soit $p = \sum_{k=0}^{d} p_k x^k$ un polynôme de $\mathbb{C}_d[x]$ et

$$f \,:\, (p,x) \in \mathbb{C}_d[x] \times \mathbb{C} \to f(p,x) = \sum_{k=0}^{d} p_k x^k \in \mathbb{C}.$$

Théorème 4.2. *Soit p un polynôme de $\mathbb{C}_d[x]$ et x une racine de p. Le problème (p,x) tel que $f(p,x) = 0$ est bien posé si la racine x de p est simple. De plus*

$$C(p,x) = \frac{|x|_1^d}{|p^{'}(x)|}.$$

Preuve. La dérivée $D_2 f(p,x) = p^{'}(x)$. Donc le problème (p,x) est bien posé si $f^{'}(x) \neq 0$, i.e. si x si une racine simple de p.

L'opérateur de conditionnement $K(p,x)$ est égal à :

$$K(p,x)(q) = -D_2 f(p,x) D_1 f(p,x)(q) = -\frac{\sum_{k=0}^{d} q_k x^k}{p^{'}(x)}.$$

Du lemme 4.1 on déduit

$$\begin{aligned} C(p,x) &= \max_{||q||=1} \frac{\left|\sum_{k=0}^{d} q_k x^k\right|}{|p^{'}(x)|} \\ &= \frac{|x|_1^d}{|p^{'}(x)|}. \end{aligned}$$

□

On montre également

Théorème 4.3. *On note par S^2 la sphère de centre $(0,1/2)$ et de rayon $1/2$. Soit $V_d = \max_{x_1,\dots,x_d \in S^2} \prod_{1\leq i<j\leq d} ||x_i - x_j||$. on considère un polynôme de degré d ayant pour zéros $(z_1,\dots,z_d)$ et ζ l'un de ses zéros. On introduit*

$$\bar{z}_i = \frac{(z_i,1)}{1+||z_i||^2}.$$

Alors

$$C(f,\zeta) \leq \sqrt{d(d+1)} \frac{V_d}{\prod_{1\leq i<j\leq d} ||\bar{x}_i - \bar{x}_j||}.$$

4.1. **Majoration de l'erreur absolue.** Soient (p,x) et (q,y) tels que $p(x) = q(y) = 0$. Le but est d'obtenir une majoration de $|y - x|$ quand q est proche de p.

Théorème 4.4. *Soit* $a = (1 + |x|_1)^d$, $\epsilon = \dfrac{a\gamma}{|p'(x)|}||q - p||$ *avec*

$$\gamma := \sup_{k\geq 2} \left(\frac{|p^{(k)}(x)|}{|p'(x)|} \right)^{\frac{1}{k-1}}.$$

Alors si $\epsilon \leq 3 - 2\sqrt{2}$ *on a :*

$$|y - x| \leq \frac{1 + \epsilon - \sqrt{\epsilon^2 - 6\epsilon + 1}}{4\gamma}.$$

Preuve. On écrit $0 = q(y) = (q - p)(y) + p(y)$. La formule de Taylor en x montre que

$$y - x = -\sum_{k=0}^{d} \frac{p^{(k)}(x) - q^{(k)}(x)}{k!p'(x)} - \sum_{k=2}^{d} \frac{p^{(k)}(x)}{k!}(y - x)^k.$$

En utilisant le 3 du lemme 4.1 et les notations a,ϵ et γ on obtient successivement :

$$\begin{aligned}\gamma|y - x| &\leq \frac{\gamma||q - p||}{|p'(x)|} \sum_{k=0}^{d} \binom{d}{k} |x|_1^{d-k} + \frac{\gamma^2|y - x|^2}{1 - \gamma|y - x|} \\ &\leq \epsilon + \frac{\gamma^2|y - x|^2}{1 - \gamma|y - x|}.\end{aligned}$$

Finalement $u = \gamma|y - x|$ vérifie

$$2u^2 - (1 + \epsilon)u + \epsilon \geq 0.$$

La condition $\epsilon \leq 3 - 2\sqrt{2}$ permet d'avoir la majoration

$$u \leq \frac{1 + \epsilon - \sqrt{\epsilon^2 - 6\epsilon + 1}}{4}.$$

□

Maintenant on étudie comment quantifier une perturbation q d'un polynôme p de telle sorte que q le même nombre de racine que p dans une boule donnée. Rappelons le théorème de Rouché.

Théorème 4.5. *Soient* p *et* q *deux polynômes de* $\mathbb{C}[z]$. *Si*

$$\forall x \in \partial B(x_0, r), \quad ||f(x) - g(x)|| < ||f(x)||$$

alors les polynômes p *et* q *ont le même nombre de racines chacune comptée avec sa multiplicité respective dans* $B(x_0, r)$.

Le lemme suivant donne le rayon d'une boule qui sépare une racine simple d'un polynôme de ses autres racines.

Lemme 4.6. *Soit* x *une racine simple de* p. *Alors* x *est l'unique racine de* p *dans la boule ouverte* $B\left(x, \dfrac{1}{2\gamma}\right)$.

Théorème 4.7. *Soient p un polynôme de $\mathbb{C}_d[x]$ et x une racine simple de p. On considère $r < \dfrac{1}{2\gamma}$. Tout polynôme q vérifiant*

$$||p-q||(1+(|x|+r)^2)^d < |p^{'}(x)|\left(1-\frac{1}{2\gamma}\right)$$

a une unique racine dans la boule ouverte $B(x,r)$

Preuve. Soit z tel que $|z-x| = r$. Du lemme 4.1 on a

$$|p(z)-q(z)| \leq ||p-q||\ |z|_1^d \leq ||p-q||(1+(|x|+r)^2)^d.$$

D'autre part une minoration de la formule de Taylor de p en x donne

$$|p(z)| \geq |p^{'}(x)|\left(1-\sum_{k\geq 2}\gamma^{k-1}r^k\right) \geq |p^{'}(x)|\frac{\gamma r^2}{1-\gamma r} \geq |p^{'}(x)|\left(1-\frac{1}{2\gamma}\right).$$

On sait que p a une seule racine dans la boule $B(x,r)$. Si pour tout $z \in \partial B(x,r)$ on $|p(z)-q(z)| < |p(z)|$ alors q a une seule racine dans $B(0,r)$. Ceci est réalisé si

$$||p-q||(1+(|x|+r)^2)^d < |p^{'}(x)|\left(1-\frac{1}{2\gamma}\right).$$

□

5. Conditionnement du problème des vecteurs propres et des valeurs propres

Soit $f\ :\ (A,\lambda,x) \in \mathcal{M}_{n,n}\times\mathbb{C}\times\mathbb{C}^n \to (Ax-\lambda x, ||x||^2/2-1) \in \mathbb{C}^n\times\mathbb{R}$. Dans cette partie on note M^H la transposée conjuguée d'une matrice M.

Théorème 5.1. *Le problème (A,λ,x) est bien posé si la valeur propre λ de la matrice A est simple. Le conditionnement $C(A,\lambda,x)$ est égal à*

$$C(A,\lambda,x) \leq 2||I-AU_1R^{-1}U_1^H|| + \sqrt{2}||U_1R^{-1}U_1^H||,$$

où

1- *U_1 est une matrice de taille $n\times n-1$ telle que $U=(x,U_1)$ soit une matrice unitaire vérifiant $U_1^Hx=0$.*
2- *$R = U_1(A-\lambda I)U_1^H$.*

Preuve. La dérivée $D_2f(A,\lambda,x)$ est égale à

$$D_2f(A,\lambda,x) = \begin{pmatrix} -x & A-\lambda I \\ 0 & x^H \end{pmatrix}$$

Soit $U = (x, U_1)$ une matrice unitaire tel que $U_1^H x = 0$. Alors $AU = (\lambda x, AU_1)$ et $U^H(A-\lambda I)U = \begin{pmatrix} 0 & x^H A U_1 \\ 0 & U_1^H(A-\lambda I)U_1 \end{pmatrix}$. Il s'ensuit

$$\begin{pmatrix} U^H & 0 \\ 0 & 1 \end{pmatrix} D_2 f(A,\lambda,x) \begin{pmatrix} 0 & 1 \\ U & 0 \end{pmatrix} = \begin{pmatrix} 0 & x^H A U_1 & -1 \\ 0 & U_1^H(A-\lambda I)U_1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Le déterminant de la matrice précédente est celui de la matrice $U_1^H(A-\lambda I)U_1$. Il est non nul si et seulement si λ est valeur propre simple de A.

Calculons l'inverse de $D_2 f(A,\lambda,x)$. Posons $z = x^H A U_1$ et $R = U_1^H(A-\lambda I)U_1$. L'inverse de la matrice $S = \begin{pmatrix} z & -1 \\ R & 0 \end{pmatrix}$ est égale à $S^{-1} = \begin{pmatrix} 0 & R^{-1} \\ -1 & zR^{-1} \end{pmatrix}$.

De plus $\begin{pmatrix} 0 & S \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 \\ S^{-1} & 0 \end{pmatrix}$. On en déduit :

$$\begin{aligned} D_2 f(A,b,\lambda,x) &= \begin{pmatrix} 0 & 1 \\ U & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ S^{-1} & 0 \end{pmatrix} \begin{pmatrix} U^H & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 1 \\ x & U_1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & R^{-1} & 0 \\ -1 & zR^{-1} & 0 \end{pmatrix} \begin{pmatrix} x^H & 0 \\ U_1^H & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -x^H + zR^{-1}U_1^H & 0 \\ U_1 R^{-1} U_1^H & x \end{pmatrix} \end{aligned}$$

L'opérateur de conditionnement est donc égal à:

$$-D_2 f(A,\lambda,x) D_1 f(A,\lambda,x)(B,\mu,y) = \begin{pmatrix} x^H(I - AU_1R^{-1}U_1^H Bx \\ U_1R^{-1}U_1^H Bx \end{pmatrix}.$$

Le conditionnement $C(A,\lambda,x)$ est la norme de l'opérateur précédent. Puisque $||x||^2 = 2$, il vient :

$$C(A,\lambda,x) \leq 2||I - AU_1R^{-1}U_1^H|| + \sqrt{2}||U_1R^{-1}U_1^H||.$$

□

6. Prolongement : analyse retrograde des erreurs

Considérons un problème bien posé : $(x_0, y_0) \in G$. Soit $y \in Y$ proche de y_0. La plupart du temps le cardinal de l'ensemble des $x \in X$ qui satisfont $(x, y) \in G$ est grand. Le problème est de choisir parmi ces x celui qui est le plus proche de x_0.

Précisons cette idée. L'outil pour traiter cette question est l'inverse de Moore-Penrose.

Définition 6.1. *Soit un opérateur linéaire sujectif L défini de E dans F. L'inverse de Moore-Penrose de L est l'opérateur $L^\dagger$ défini de F dans E par :*

$$L^\dagger = \left(L|_{(Ker\, L)^\perp}\right)^{-1}.$$

On montre que

Proposition 6.2. 1- $LL^\dagger = id_F$.
2- *Pour tout $x \in (Ker\, L)^\perp$ on a $L^\dagger L(x) = x$.*

L'analyse rétrograde des erreurs consiste à définir un opérateur σ dont la dérivée est l'inverse de Moore-Penrose de la dérivée $K(x, y)$ de $\varphi(x)$.

Théorème 6.3. *Si $K(x_0, y_0)$ est surjectif il existe un opérateur σ défini de $V_{y_0} \subset F$ dans E tel que pour tout $y \in V_{y_0}$ on ait :*
1- $\sigma(y_0) = x_0$.
2- $\varphi(\sigma(y)) = y$.
3- $D\sigma(y) = K(x, y)^\dagger$.
On appelle opérateur de conditionnement inverse du problème $(x_0, y_0) \in G$ l'inverse de Moore-Penrose $K(x_0, y_0)^\dagger$.

La norme de cet opérateur est le conditionnement inverse. On le note

$$C(x_0, y_0)^\dagger = \sup_{y \in Y\ ||y||=1} \frac{||K(x_0, y_0)^\dagger y||}{||y||}.$$

On a

Preuve. La fonction implicite φ est définie sur un ouvert U_{x_0} dans U_{y_0}. La fonction $\varphi_1(z) = \varphi(z + x_0) - y_0$, définie sur un voisinage de 0, vérifie $\varphi_1(0) = 0$ et $D\varphi_1(0) = K(x_0, y_0)$. Puisque cette dérivée est surjective, φ_1 est une submersion. On sait alors qu'il existe un difféomorphisme h défini dans un voisinage de 0 de E tel que :
1- $h(0) = 0$,
2- $Dh(0) = id_E$,
3- $\varphi_1 = K(x_0, y_0)\, o\, h$.
L'application $\sigma_1 = h^{-1}\, oK(x_0, y_0)^\dagger$ vérifie :
1- $\sigma_1(0) = 0$,
2- $\varphi_1(\sigma_1(w)) = w$ pour tout w dans un voisinage du zéro de F.
3- $D\sigma_1(y_0) = K(x_0, y_0)^\dagger$.
Maintenant la fonction $\sigma(w + y_0) = x_0 + \sigma_1(w)$ est définie dans V_{y_0} et vérifie :
1- $\sigma(y_0) = x_0$,
2- $\varphi(\sigma(y)) = \varphi(\sigma(w + y_0)) = \varphi(x_0 + \sigma_1(w)) = y_0 + \varphi_1(\sigma_1(w)) = y_0 + w = y$ pour tout w dans un voisinage du zéro de F.
3- $D\sigma(y_0) = D\sigma_1(y_0) = K(x_0, y_0)^\dagger$.

□

Théorème 6.4. *Pour tout opérateur δ défini sur V_{y_0} dans U_{x_0} vérifiant $\delta(y_0) = x_0$ et $\varphi(\delta(y)) = y$ pour tout $y \in V_{y_0}$ on a :*

$$C(x_0, y_0)^\dagger = ||K(x_0, y_0)^\dagger|| \leq ||D\delta(y_0)||.$$

Preuve. Pour tout $w \in Y$ on a $D\delta(y_0)w = (D\delta(y_0) - D\sigma(y_0))w + D\sigma(y_0)w$. Si $D\sigma(y_0)w \in Ker\, K(x_0, y_0)^\perp$ et $(D\delta(y_0) - D\sigma(y_0))w \in Ker\, K(x_0, y_0)$ alors

$$||D\delta(y_0)w||^2 = ||(D\delta(y_0) - D\sigma(y_0))w||^2 + ||D\sigma(y_0)w||^2 \geq ||D\sigma(y_0)w||^2.$$

Et le résultat suit.

Puisque $D\sigma(y_0) = K(x_0, y_0)^\dagger$ on a $D\sigma(y_0)w \in Ker\, K(x_0, y_0)^\perp$.

Il reste à montrer que $(D\delta(y_0) - D\sigma(y_0))w \in Ker\, K(x_0, y_0)$. Pour tout $y \in V_{y_0}$ on sait que : $y = \phi(\delta(y)) = x_0 + K(x_0, y_0)(h(\delta(y) - x_0))$. Donc $id_Y = K(x_0, y_0)\, o\, Dh(0)\, o\, D\delta(y_0) = K(x_0, y_0)\, o\, D\delta(y_0)$. Or $K(x_0, y_0)D\sigma(y_0) = K(x_0, y_0)\, o\, K(x_0, y_0)^\dagger = id_Y$. On en déduit $K(x_0, y_0)(D\delta(y_0) - D\sigma(y_0))w = 0$. Ce qui implique $(D\delta(y_0) - D\sigma(y_0))w \in Ker\, K(x_0, y_0)$. $\square$

Théorème 6.5. *On suppose que E et F sont des espaces euclidiens. On note par $u^T v$ le produit scalaire sur E ou F suivant le contexte. On considère pour y donné le problème d'optimisation :*

$$\min_{(x,y)\in G} \frac{1}{2}||x - x_0||^2$$

Il existe une fonction δ définie sur un voisinage de V_{y_0} dans E telle que :

1- $\delta(y_0) = x_0$.
2- *$\delta(y)$ est solution du problème d'optimisation.*
3- $D\delta(y_0) = K(x_0, y_0)^\dagger$.

Preuve. Considérons le lagrangien

$$L(x, \lambda) = \frac{1}{2}||x - x_0||^2 + (\varphi(x) - y)^T\lambda$$

Alors la solution x du problème d'optimisation est aussi solution du système :

$$\begin{aligned} &\varphi(x) - y = 0 \\ &x - x_0 + D\varphi(x)^T\lambda = 0. \end{aligned}$$

L'image de $D\sigma(x)^T$ est égal à l'orthogonal du noyau de $D\varphi(x)$. Donc $x - x_0 \in (Ker\, D\varphi(x))^\perp$ $\square$