Local stability of the output least square parameter estimation technique

Guy Chavent

CENTRE DE ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.:954 90 20

Rapports de Recherche

N° 136

LOCAL STABILITY OF THE OUTPUT LEAST SQUARE

PARAMETER ESTIMATION TECHNIQUE

Guy CHAVENT

Juin 1982

LOCAL STABILITY OF THE OUTPUT LEAST SQUARE PARAMETER ESTIMATION TECHNIQUE

Guy CHAVENT

ABSTRACT

We investigate the well posedness of the OLS (output least square) setting of parameter estimation problems :

find $\hat{x} \in C$ which minimizes $\|\phi(x)-z\|^2$ over C

where x is the parameter, C the convex of admissible parameters, z the measurement and ϕ the parameter $\rightarrow$ output mapping.

When $\|\phi'(x)y\| \geq \alpha\|x\|$ with $\alpha > 0$, and C samll enough compared to the "curvature" of ϕ, we obtain the (lipschitz) continuity of the mapping $z \rightarrow \hat{x}$ on a neighborhood of $\phi(C)$.

As "practical" application of this result, we get that for finite dimensional parameters, the OLS technique yields an estimated parameter converging towards the exact one when model and measurement errors tends to zero, as soon as the derivative of the parameter $\rightarrow$ output mapping is injective for the exact parameter and the size of the set of admissible parameter is small enough.

Two examples of application are given.

RESUME

On étudie la stabilité du problème d'estimation de paramètre par la méthode des moindres carrés.

$$\text{trouver } \hat{x} \in C \text{ qui minimise } \|\phi(x)-z\|^2 \text{ sur } C$$

(x est le paramètre, C le convexe des paramètres admissibles, z l'observation et ϕ l'application paramètre $\rightarrow$ sortie).

Lorsque $\|\phi'(x)y\| \geq \alpha\|x\|$ $(\alpha > 0)$, est suffisamment petit devant la "courbure" de ϕ, nous obtenons la continuité lipschitz, sur un voisinnage de $\phi(C)$, de l'application $z \rightarrow \hat{x}$.

Comme application "pratique" de ce résultat, on trouve dans le cas d'un nombre fini de paramètres, que le paramètre estimé par la méthode des moindres carrés converge vers le paramètre réel quand les erreurs de mesure et de modèle tendent vers zéro, dès lors que la dérivée de l'application paramètres $\rightarrow$ sortie est injective pour la valeur exacte des paramètres et que la taille du convexe des paramètres admissible est suffisamment petite.

Deux exemples d'application à des problèmes d'estimation de paramètre sont donnés.

LOCAL STABILITY OF THE OUTPUT LEAST SQUARE PARAMETER ESTIMATION TECHNIQUE

G. CHAVENT

A typical situation for parameter estimation is the following :

- the set C of admissible parameters x is a convex and bounded set of a Banach space E
- the measurement z is made in a (pre)-Hilbert space F with norm $\| \quad \|$
- the mapping $\phi : x \to y$ defined by the state equation + observation operator is never linear, but usually regular (at least twice continuously differentiable).

A parameter $x \in C$ is usually said identifiable (cf [2], [3], [4]) if :

(I) $\quad \phi(x') \neq \phi(x) \quad \forall\, x' \in C,\ x' \neq x$

which is satisfied as soon as ϕ is injective on C.

Pratically, the parameter estimation problem is solved by minimization of the output least square criterion :

(OLS) $\quad$ find $x \in C$ which minimizes $J(x) = \|\phi(x)-z\|^2$ over C.

In order to take in account the measurement and model errors (which causes z never to belong to $\phi(C)$!), the output least square identifiability was introduced in [1] :

(OLSI) $\left\{ \begin{array}{l} \exists \text{ a neighbourhood } V \text{ of } \phi(c) \text{ in F such that, for every } z \in V, \\ \text{the (OLS) problem has a unique solution x depending continuously} \\ \text{on z.} \end{array} \right.$

A sufficient condition for OLSI, requiring that $\phi'(x)$ be invertible, was given in the same paper together with an application to the stability of the inverse seismic problem.

The aim of the present paper is to give a weaker sufficient condition for OLSI, which requires only that $\phi'(x)$ is injective and "coercive". Such a condition is expected to have a larger field of applications, especially for the case of finite dimensional parameters. We shall give two such examples of application, including estimation of coefficients in O.D.E.

THEORY

The following hypothesis are valid throughout this paragraph (and shall not be repeated in the lemmas and theorems) :

$$(1)\qquad \begin{cases} (E,\|\ \|_E) = \text{normed vector space} \\ C \subset E \qquad \text{convex, bounded} \\ (F,\|\ \|_F) = \text{pre-Hilbert space, with scalar product } (\ ,\ ) \end{cases}$$

$$(2)\qquad \phi : C \to F \text{ twice Gateaux-derivable, s.t.}$$

$$(3)\qquad \begin{cases} \exists\ \alpha,\ \beta,\ \delta > 0 \text{ such that} \\ \alpha\|x\|_E \le \|\phi'(c).x\|_F \le \delta\|x\|_E & \forall c \in C,\ \forall x \in E \\ \|\phi''(c)(x,x)\|_F \le \beta\|x\|_E^2 & \forall c \in C,\ \forall x \in E \end{cases}$$

and we are looking for conditions on C and ϕ such that :

$$(4)\qquad \exists\ \gamma > 0,\ d(a,\phi(C)) < \gamma \Rightarrow \begin{cases} \exists\,!\ x \in C \text{ solution of (OLS)} \\ \text{depending continuously on } z. \end{cases}$$

If (4) is satisfied, then ϕ is necessarily injective over C. A sufficient condition for the injectivity of ϕ is given by the

Lemma 1 : If

(5) $\qquad \beta \text{ diam } C < 2\alpha$

then ϕ is injective.

Proof : Let $0,\ x \in C$ such that $\phi(0) = \phi(x)$ and defined $v(\theta) = \phi(\theta x)$ $0 \le \theta \le 1$.

A second order Taylor formula yields :

$$v(1) = v(0) + v'(0) + \int_0^1 (1-\theta)\ v''(\theta)\ d\theta$$

which, using (3), gives as $v(1) = v(0)$:

$$\alpha\|x\| \leq \frac{\beta}{2}\|x\|^2$$

which, if $x \neq 0$, implies $\|x-0\| \geq \operatorname{diam} C$, which is in contradiction with (5). Hence $x = 0$ and ϕ is injective. ∎

Remark 1 : The best coefficient in (5) lies between 2 and 2Π, as it results from lemma 1 and from the special case $C = [0,X] \subset \mathbb{R}$, $\phi(x) = (\sin x, \cos x) \in \mathbb{R}^2$ where $\alpha = \beta = 1$ and where ϕ is obviously injective as long as $0 < X < 2\Pi$. ∎

Lemma 2 : For any 0, $x \in C$, $x \neq 0$, and $a \in F$ define a function $f : [01] \to \mathbb{R}$ by

(7) $$f(\theta) = \|\phi(\theta x)-a\|_F^2 \qquad 0 \leq \theta \leq 1.$$

Then one has the following implication :

(8) $$\left.\begin{array}{l} f(0) = f(1) = d^2 \\ \exists\, \theta_0 \in]01[\text{ s.t. } f''(\theta_0) \leq 0 \end{array}\right\} \Longrightarrow d \geq \frac{\alpha^2}{\beta} - \frac{\beta}{8}\|x\|^2.$$

Proof : Derivation of f yields :

$$f'(\theta) = 2(\phi(\theta x)-a,\phi'(\theta x).x)$$

$$f''(\theta) = 2\|\phi'(\theta x).x\|^2 + 2(\phi(\theta x)-a,\phi''(\theta x)(x,x)).$$

Hence

$$|f'(\theta)| < 2\sqrt{f(\theta)}\;\|\phi'(\theta x).x\|$$

which together with (3) yields :

$$f''(\theta) \geq \frac{f'(\theta)^2}{2f(\theta)} - 2\beta\sqrt{f(\theta)}\;\|x\|_E^2.$$

This shows that the second derivative of $\theta \to 2\sqrt{f(\theta)} + \beta(\theta^2-\theta)\|x\|_E^2$ is positive, hence this function is convex :

$$2\sqrt{f(\theta)} + \beta(\theta^2-\theta)\|x\|_E^2 \leq 2d$$

and

(9) $$\sqrt{f(\theta)} \leq d + \frac{\beta}{8}\|x\|_E^2 \qquad \forall\, \theta \in [0,1].$$

But from the second hypothesis in (8) we get, using (3) :

$$2\alpha^2\|x\|_E^2 \leq \|\phi'(\theta_0 x).x\|_F^2 \leq -2(\phi(\theta_0 x)-a,\phi''(\theta_0 x)(x,x)) \leq 2\beta\sqrt{f(\theta_0)}\;\|x\|_E^2$$

i.e.

(10) $$\frac{\alpha^2}{\beta} \leq \sqrt{f(\theta_0)}$$

and the result follows from (9) and (10). ∎

Theorem 1 : (Uniqueness) If :

(11) $\beta \operatorname{diam} C < 2\sqrt{2}\,\alpha$

define

(12) $\gamma = \frac{\alpha^2}{\beta} - \frac{\beta}{8} (\operatorname{diam} C)^2 > 0.$

Then for every $a \in F$ such that $d(a,\phi(c)) < \gamma$, there exists at most one $x \in C$ solution to the OLS problem (in particular ϕ is injective on C).

Proof : Suppose there exist two distinct solutions 0 and x to the OLS problem. The function $f(\theta)$ defined by (7) in Lemma 2 satisfies here :

$$f(0) = f(1) \leq f(\theta) \quad \forall\, \theta \in [0,1].$$

Either $f(\theta)$ is identically equal to $f(0) = f(1)$ (and then $f''(\theta) \equiv 0$) or $f(\theta)$ is for some $\theta \in]01[$ strictly greater than $f(0) = f(1)$ (and then $f''(\theta_0) \leq 0$ for the $\theta_0 \in]01[$ which maximizes f over [01]). In both cases the hypothesis in (8) are fullfilled, which shows that :

(13) $d \geq \frac{\alpha^2}{\beta} - \frac{\beta}{8} \|x\|_E^2 .$

But, as $d(a,\phi(C)) < \gamma$, we have $d = \sqrt{f(0)} = \sqrt{f(1)} < \gamma$, which together with (12) (13) shows that $\|x\|_E > \operatorname{diam} C$ which is impossible. ∎

Lemma 3 : Let hypothesis (11) and notation (12) of theorem 1 hold. Then we have :

(14) $$\left\{\begin{array}{l} a \in F,\ 0,x \in C \\ \|\phi(0)-a\|_F < \gamma \\ \|\phi(x)-a\|_F < \gamma \end{array}\right. \Longrightarrow \left\{\begin{array}{l} f(\theta) \text{ defined by (7) satisfies} \\ f''(\theta) > \frac{\beta^2}{4} (\operatorname{diam} C)^2 \ \|x\|^2 \quad \forall\, \theta \in [01] \\ \text{(hence f is convex on [01]).} \end{array}\right.$$

Proof : The implication is obvious if $x = 0$. Hence let us suppose that $x \neq 0$ and that, for example, $f(0) \geq f(1)$. We prove first that $f(\theta) \leq f(0)\ \forall\, \theta \in]01[$. If not, there would exist $\theta_1 \in]01[$ and $\theta_0 \in]0,\theta_1[$ such that $f(0) = f(\theta_1)$ and $f''(\theta_0) \leq 0$. Using then lemma 2 between 0 and θ_1 yields :

(15) $\sqrt{f(0)} = \sqrt{f(\theta_1)} \geq \frac{d^2}{\beta} - \frac{\beta}{8} \|\theta_1 x\|^2 .$

But from (14) we know that $\sqrt{f(0)} < \gamma$ which together with (5) shows that :

$$\theta_1^2 \ \|x\|^2 > (\operatorname{diam} C)^2$$

which is impossible. We prove then that f is convex on]01[. Using the expression for $f''(\theta)$ given in the proof of lemma 2 and the hypothesis (3) gives :

$$f''(\theta) \geq 2(\alpha^2 - \beta\sqrt{f(\theta)})\,\|x\|^2 \qquad \forall\ \theta \in [0,1]$$

which implies, as we have proved that $\sqrt{f(\theta)} \leq \sqrt{f(0)} < \gamma$

$$f''(\theta) > 2(\alpha^2 - \beta\gamma)\ \|x\|^2 \qquad \forall\ \theta \in [01]$$

which gives the sought result using (12). ∎

Theorem 2 : (Existence and Uniqueness) Let hypothesis of theorem 1 hold and :

(16) C closed

(17) E = Banach space .

Then for every $a \in F$ such that $d(a,\phi(C)) < \gamma$, there exists one and only one x solution to the OLS problem. Moreover, any minimizing sequence is converging towards this unique solution.

Proof : Let $\{x_n,\ n \in \mathbb{N}\}$ be a minimizing sequence of the OLS problem :

$$\|\phi(x_n)-a\| \to d(\phi(C),a) < \gamma .$$

Hence there exists $N \in \mathbb{N}$ s.t., $m,p \geq N$ implies

$$\|\phi(x_m)-a\| < \gamma,\quad \|\phi(x_p)-a\| < \gamma$$

which through lemma 3 implies :

(18) $$f''(\theta) > c\|x_p-x_m\|_E^2 \text{ with } c = \frac{\beta^2}{4}(\text{diam } C)^2$$

where f is defined, as in lemma 2, by :

(19) $$f(\theta) = \|\phi(\theta x_p+(1-\theta)x_m)-a\|_F^2 \geq d(a,\phi(C))^2.$$

From (18) we deduce :

(20) $$\begin{cases} f(\theta) \leq \theta f(1) + (1-\theta)f(0) + \frac{c}{2}(\theta^2-\theta)\|x_p-x_m\|_E^2 \\ \forall\ \theta \in [01] \end{cases}$$

which together with (19) gives, for $\theta = \frac{1}{2}$:

$$d(a,\phi(C))^2 \leq \frac{1}{2}(f(1) + f(0)) - \frac{c}{8}\|x_p-x_m\|_E^2$$

or equivalently

(21) $$\frac{c}{8}\|x_p-x_m\|_E^2 \leq \frac{1}{2}\Big[d(x_p,\phi(C))^2 + d(x_m,\phi(C))^2\Big] - d(a,\phi(C))^2$$

which proves that $\{x_n\}$ is a Cauchy sequence as the right-hand side of (21) tends to zero as $m,p \to \infty$. Hence there exists, as C is closed, an $x \in C$ such that

$$x_n \to x \text{ in E as } n \to \infty$$

and such that x is a solution of the OLS problem. The uniqueness follows from theorem 1. ∎

Remark 2 : If $a_n \to a$ in F (with $d(a,\phi(C)) < \gamma$), then the sequence $\{x_n\}$ of the solutions of the OLS problems with a_n is a minimizing sequence for the OLS problem with a. Then we get immediatly from theorem 2 that $x_n \to x$ where x is the solution of the OLS problem with a : the mapping $a \to x$ defined by theorem 2 is continuous. In fact we shall see in theorem 3 that this mapping is lipschitzean. ∎

We first give a generalization of lemma 2.

Lemma 4 : For any $0,x \in C$, $x \neq 0$, and for any $0,a \in F$ define a function $f : [01] \to \mathbb{R}$ by :

(22) $$f(\theta) = \|\phi(\theta x)-\theta a\|_F^2 \qquad 0 \le \theta \le 1 .$$

Then one has the following implication :

(23) $$\left.\begin{array}{l} f(0) = f(1) = d^2 \\ \exists\ \theta_0 \in\ ]01[\text{ s.t. } f''(\theta_0) \le 0 \end{array}\right\} \Longrightarrow d \ge \frac{\alpha^2}{\beta} - \frac{\beta^2}{8}\|x\|_E^2 - \frac{2\delta}{\beta}\frac{\|a\|_F}{\|x\|_E} .$$

Proof : It follows that of lemma 2 until (9) is proved. Then $f''(\theta_0) \le 0$ yields :

$$\|\phi'(\theta x)x-a\|_F^2 \le \beta\sqrt{f(\theta_0)}\ \|x\|_E^2$$

i.e.

$$\|\phi'(\theta x)x\|^2 - 2<\phi'(\theta x)x,a> + \|a\|^2 \le \beta\sqrt{f(\theta_0)}\ \|x\|_E^2$$

i.e., using (3) and the fact that $\|a\|^2 \ge 0$:

$$\alpha^2\|x\|_E^2 - 2\delta\|x\|_E\ \|a\|_F \le \beta\sqrt{f(\theta_0)}\ \|x\|_E^2$$

or

$$\sqrt{f(\theta_0)} \ge \frac{\alpha^2}{\beta} - 2\frac{\delta\|a\|_F}{\beta\|x\|_E}$$

with together with (9) gives the expected result. ∎

Lemma 5 : Let hypothesis (11) and notation (12) of theorem 1 hold. Then, for $0,x \in C$ and $0,a \in F$ and f defined by (22) we have :

(24) $$\left.\begin{array}{l} \sqrt{f(0)} = \sqrt{f(1)} < \gamma \\ \exists\ \theta_0 \in [0,1] \text{ s.t. } f''(\theta_0) \le 0 \end{array}\right\} \Longrightarrow \|x\|_E \le \frac{2\delta}{\beta(\gamma-\sqrt{f_0})}\|a\|_F .$$

Proof : From lemma 4 we get, if $x \neq 0$:

$$\sqrt{f_0} \geq \frac{\alpha^2}{\beta} - \frac{\beta^2}{8} \|x\|_E^2 - \frac{2\delta}{\beta} \frac{\|a\|_F}{\|x\|_E}$$

i.e.

$$\sqrt{f_0} \geq \frac{\alpha^2}{\beta} - \frac{\beta^2}{8} (\text{diam } C)^2 - \frac{2\delta}{\beta} \frac{\|a\|_F}{\|x\|_E}$$

i.e. using (12)

$$\sqrt{f_0} \geq \gamma - \frac{2\delta}{\beta} \frac{\|a\|_F}{\|x\|_E} .$$

■

Theorem 3 : Under hypothesis of theorem 2, the mapping $a \to x$ defined by theorem 2 is lipschitz-continuous. More precisely, let $a,b \in F$ such that

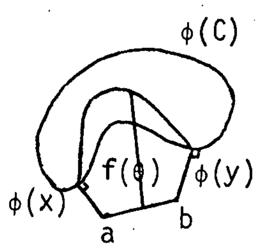
(25) $$d = \text{Max}\{d(a,\phi(C)),\ d(b,\phi(C))\} < \gamma$$

and let $x,y \in C$ be the corresponding unique solution to the OLS problem given by theorem 2. Define $f : [01] \to \mathbb{R}$ by

(26) $$f(\theta) = \|\phi(\theta x + (1-\theta)y) - \theta a - (1-\theta)b\|_F^2.$$

Then one has the following alternative :

i) either "a,b are on a concave side of $\phi(c)$", i.e.

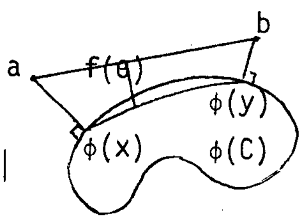


(27) $$\exists\, \theta \in]01[\text{ s.t. } f(\theta) > d^2$$

then :

(28) $$\|x-y\|_E \leq \frac{2\delta}{\beta(\gamma-d)} \|a-b\|_F .$$

ii) or "a,b are on a convex side of $\phi(c)$", i.e.



(29) $$f(\theta) \leq d^2 \quad \forall\, \theta \in [01]$$

then

(30) $$\|x-y\|_E \leq \frac{2\delta}{\alpha^2-\beta\gamma} \|a-b\|_F = \frac{16\delta}{\beta^2(\text{diam } C)^2} \|a-b\|$$

Proof : We prove first i). From (27) we get, if for example $f(0) \geq f(1)$, the existence of $\theta_1 \in]01[$ such that $f(0) = f(\theta_1) = d^2$, and of $\theta_0 \in]0,\theta_1[$ such that $f''(\theta_0) \leq 0$. Using then lemma 5 between y and $\theta_1 x + (1-\theta_1)y$ yields :

$$\|\theta_1(x-y)\|_E \leq \frac{2\delta}{\beta(\gamma-d)} \|\theta_1(a-b)\|$$

which gives (28) after division by $\theta_1 \neq 0$.

We prove now ii). Using (3), (29) one gets, as in the proof of lemma 3 :

$$\frac{1}{2} f''(\theta) \geq (\alpha^2-\beta\gamma) \; \|x-y\|_E^2 + \|a-b\|_F^2 - 2\delta\|x-y\|_E \; \|a-b\|_F .$$

Hence :

$$f(\theta) - (\theta^2-\theta)\Big[(\alpha^2-\beta\gamma)\|x-y\|_E^2 + \|a-b\|_F^2 - 2\delta\|x-y\|_E \|a-b\|_F\Big] \leq \theta f(1)+(1-\theta)f(0).$$

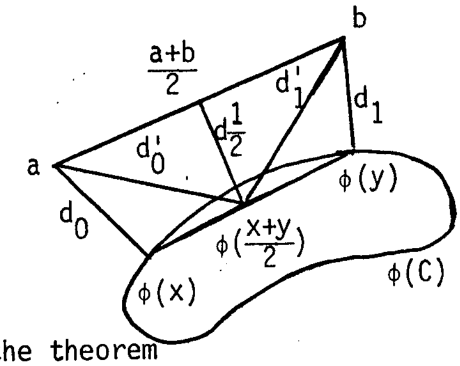
For $\theta = \frac{1}{2}$ one gets with the notation $d_\theta = \sqrt{f(\theta)}$

$$(31) \qquad \frac{1}{4}\Big[(\alpha^2-\beta\gamma)\|x-y\|_E^2 + \|a-b\|_F^2 - 2\delta\|x-y\|_E \|a-b\|_F\Big] \leq \frac{1}{2}(d_0^2+d_1^2) - d_{1/2}^2.$$

But

$$d_0 \leq \|\phi(\tfrac{x+y}{2})-a\| = d_0'$$

$$d_1 \leq \|\phi(\tfrac{x+y}{2})-b\| = d_1'.$$



Hence

$$(32) \qquad \frac{1}{2}(d_0^2+d_1^2) - d_{1/2}^2 \leq \frac{1}{2}(d_0'^2+d_1'^2) - d_{1/2}^2.$$

But as F is a pre-Hilbert space one can use the theorem of the mediane :

$$(33) \qquad \frac{1}{2}(d_0'^2+d_1'^2) - d_{1/2}^2 = \frac{\|a-b\|^2}{4}.$$

From (31), (32), (33), we get :

$$\frac{1}{4}\Big[(\alpha^2-\beta\gamma)\|x-y\|_E^2 - 2\delta\|x-y\|_E \|a-b\|_F\Big] \leq 0$$

which proves (30). ∎

Remark 3 : On the "concave side" of $\phi(C)$ one can intuitively expect that the Lipschitz constant of the projection blows-up to infinity, when the points approach to the outer rand of the "security interval" surrounding $\phi(C)$, which corresponds to the γ-d denominator of the Lipschitz constant in (28). Conversely, one expects, on the "convex side" of $\phi(C)$, a Lipschitz constant independant of the position of the points in the "security interval", which is the case in (30). ∎

APPLICATION TO OUTPUT LEAST SQUARE IDENTIFIABILITY

The main difficulty in the use of the above theory is to satisfy the "coercivity" property on the derivative of ϕ. When the parameter space E is finite dimensional, the unit sphere is compact, hence the injectivity of $\phi'(c)$ implies its coercivity. So the main field of application for this theory may consist in proving that finite dimensional parameter are OLSI whenever ϕ' is injective.

More precisely, let :

(34) $\begin{cases} E = \text{finite dimensional space for the parameter } x \\ \phi : E \to F \text{ with continuous 1rst and 2nd derivatives} \\ F = \text{pre-Hilbert space for the measurements} \end{cases}$

(35) $\begin{cases} \mathcal{X}_{inj} \subset E \text{ an (open) set such that } \phi'(c) \text{ injective for every } c \in \mathcal{X}_{inj} \\ K \subset \mathcal{X}_{inj} \text{ a (convex) \underline{compact set} resuming all the \underline{a-priori}} \\ \quad \text{\underline{information} concerning the unknown parameter.} \end{cases}$

We denote by :

$S = \{x \in E \mid \|x\| = 1\}$ the (<u>compact</u>) unit sphere of E.

Then the mapping $(c,x) \in K\times S \to \|\phi'(c).x\|_F \in \mathbb{R}$ is continuous over the compact set $K\times S$ and attains its maximum and minimum value :

(36) $\begin{cases} \exists\ (c_j,x_j) \in K\times S \quad j = 0,1,2 \quad \text{such that } \forall\ (c,x) \in K\times S : \\ \alpha = \|\phi'(c_0)x_0\|_F \le \|\phi'(c)x\|_F \le \|\phi'(c_1')x_1\|_F = \delta \\ \qquad\qquad \|\phi''(c)(x,x)\|_F \le \|\phi''(c_2)(x_2,x_2)\|_F = \beta \end{cases}$

wherenecessarily $\alpha > 0$ as $\phi'(c_0)$ is injective.

Hence we have proven <u>the existence of α, β, δ</u> (<u>depending on K</u>) such that <u>hypothesis (3)</u> of the theory <u>is satisfied over K</u>.

From theorems 2 and 3 we get then the

Theorem 4 : Under hypothesis (34), (35), <u>the parameter $x \in E$ is OLSI on every convex compact subset C of K satisfying</u> :

(37) $\quad \operatorname{diam} C < 2\sqrt{2}\,\frac{\alpha}{\beta}$

where α, β are defined by (37). More precisely, the (OLS) problem has a unique solution depending in a Lipschitz-continuous way on z as soon as :

(38) $$d(z,\phi(C)) < \gamma = \frac{\alpha^2}{\beta} - \frac{\beta}{8}(\text{diam } C)^2 .$$

If $2\sqrt{2}\,\frac{\alpha}{\beta}$ turns out to be greater than diam K, this means that the a-priori information is large enough to ensure the well-posedness of the OLS Identification on C = K, provided of course that the model and measurement error can be kept strictly smaller than γ.

If $2\sqrt{2}\,\frac{\alpha}{\beta}$ is smaller than diam K, the a-priori information on the parameter is not sufficient to ensure the well posedness of (OLS) problem on C = K : one has to find additional information to reduce the diameter of K to C.

We see from (38) that there is a balance between the a-priori information on the parameter and the admissible model and measurement arror : the more a-priori information, the smaller the diameter of C and hence the larger the allowed upper bound γ on the model and measurement error.

We give now two simple examples where theorem 4 applies.

Example 1 : Let

(39) $$\begin{cases} \Omega \subset \mathbb{R}^n \text{ bounded} \\ F \in W^{3,\infty}(\mathbb{R}) \quad \text{i.e. } F, F', F'', F''' \quad L^\infty(\mathbb{R}) . \end{cases}$$

We consider then the following partial differential equation :

(40) $$\begin{cases} \frac{\partial y}{\partial t} - \Delta y + F(y) = v \\ y|_{\partial\Omega} = 0 \\ y|_{t=0} = 0 \end{cases}$$

for which we recall

Lemma 6 : The mapping $v \to y$ defined by (40) is Gateaux-Derivable from $L^2(Q)$ in W(OT) weak and $L^2(Q)$ strong, where

(41) $$\begin{cases} Q = \Omega \times]0T[\\ W(0T) = \{y \in L^2(0T\,;\,H_0^1(\Omega))\ \frac{dy}{dt} \in L^2(0T\,;\,H^{-1}(\Omega))\} . \end{cases}$$

The Gateaux-derivative at v is the mapping $\delta v \to \delta y$ defined by

$$(42)\quad \begin{cases} \dfrac{\partial \delta y}{\partial t} - \Delta \delta y + F'(y(v))\ \delta y = \delta v \\ \delta y|_{\partial\Omega} = 0 \\ y|_{t=0} = 0. \end{cases}$$

Proof : For a given $v \in L^2(Q)$, the existence of a unique solution $y(v)$ to (40) can be proved by standard techniques. We just give a sketch of the proof of the Gateaux-derivability.

For any $\lambda > 0$ let y_λ be the solution of (40) with right-hand side $v+\lambda\delta v$, and let $w_\lambda = \dfrac{y_\lambda - y_0}{\lambda}$. Obviously w_λ satisfies :

$$(43)\quad \begin{cases} \dfrac{\partial w_\lambda}{\partial t} - \Delta w_\lambda + F'(y_0+\theta\lambda w_\lambda)\ w_\lambda = \delta v \qquad (\theta \in [0,1]) \\ w_\lambda|_{\partial\Omega} = 0 \\ w_\lambda|_{t=0} = 0. \end{cases}$$

Multiplying (43) by w_λ and integrating over Ω yields :

$$(44)\quad \frac{1}{2}\frac{d}{dt}\ |w_\lambda(t)|^2 + \int_\Omega |\nabla w_\lambda(t)|^2 + \int_\Omega F'(y_0+\theta\lambda w_\lambda) w_\lambda^2 = \int_\Omega \delta v . w_\lambda .$$

Hence, as F' is bounded :

$$\begin{cases} \dfrac{1}{2}\dfrac{d}{dt}\ |w_\lambda(t)|^2 \le (\dfrac{1}{2} + \|F'\|_\infty)\ |w_\lambda(t)|^2 + \dfrac{1}{2}\ |\delta v(t)|^2 \\ w_\lambda(0) = 0 \end{cases}$$

which using Gronwall's lemma shows that :

$$\|w_\lambda\|_{L^\infty(OT\ ;\ L^2(\Omega))} \le C\|\delta v\|_{L^2(Q)}$$

and using (44) and the equation (43) :

$$(45)\quad \begin{cases} \|w_\lambda\|\ _{L^2(OT\ ;\ H_0^1(\Omega))} \le C\|\delta v\|\ _{L^2(Q)} \\ \|\dfrac{dw_\lambda}{dt}\|\ _{L^2(OT\ ;\ H^{-1}(\Omega))} \le C\|\delta v\|\ _{L^2(Q)}. \end{cases}$$

Taking first $\lambda = 1$ shows that the mapping $v \to y$ defined by (40) is <u>Lipschitz-continuous from $L^2(Q)$ into $W(OT)$</u>.

We deduce then from (45) that, for $\lambda > 0$, w_λ is bounded in $W(OT)$. Hence there exists $w \in W$ and a subsequence $\{\lambda_n, n \in N\}$ such that $w_{\lambda_n} \to w$ weakly in W as $n \to \infty$. But we have seen that $\lambda w_\lambda = y_\lambda - y_0$ tends strongly to zero in W. Hence there exists a subsequence λ_μ of λ_n such that $\lambda_\mu W_{\lambda_\mu} \to 0$ a.e. in Q, and so does $\theta\lambda_\mu W_{\lambda_\mu}$ (as θ, function of λ and x, always belongs to $[o1]$). So we can pass to the limit in (43) for the subsequence $w_{\lambda\mu}$, using the Lebesgue Convergence Theorem, which shows that the weak limit w of the subsequence w_{λ_μ} is a solution of (42). As (42) has a unique solution δy, the whole sequence w_λ converges weakly towards $w = \delta y$, which proves the Gateaux-derivability of $v \mapsto y$ into $W(oT)$ weak. ∎

<u>We define now a parameter estimation problem associated with equation (40)</u> :

Let $v_1 \dots v_m$ be given such that :

(46) $\quad v_i \in L^\infty(Q) \quad i = 1,\dots,m, \ \{v_i,\dots,v_m\}$ linearly independant

and suppose we are looking for a right-hand side v of (40) of the form

(47) $\quad v = \sum_{i=1}^{m} a_i v_i \qquad a_i \in \mathbb{R}$

in order to obtain a desired solution $z \in L^2(Q)$.

The corresponding function ϕ of the theory is here

(48) $\quad \begin{cases} \phi : \mathbb{R}^m \to L^2(Q) \\ a = (a_1 - a_m) \to y \text{ sol of (40) with } v \text{ given by (47)} \end{cases}$

Lemma 7 : The function ϕ defined by (48) is twice continuously differentiable and $\phi'(a)$ is injective $\forall a \in \mathbb{R}^m$.

Proof : We get from lemma 6 that ϕ is Gateaux-derivable, and that $\chi_i = \frac{\partial \phi}{\partial a_i}$ is solution of :

$$(49) \begin{cases} \frac{\partial \chi_i}{\partial t} - \Delta\chi_i + F'(y)\chi_i = v_i \\ \chi_i|_{\partial\Omega} = 0 \\ \chi_i|_{t=0} = 0 \end{cases}$$

and hence the χ_i are linearly independant, i.e. $\phi'(a)$ is injective. As the v_i are taken in $L^\infty(Q)$ and $F' \in L^\infty(\mathbb{R})$ one gets, using a maximum principle, that $\chi_i \in L^\infty(Q)$. Let then χ_i and χ_i' be two solutions of (49) corresponding to two functions y and y' of $L^2(Q)$, and let $w = \chi_i - \chi_i'$, one has :

$$(50) \begin{cases} \frac{\partial w}{\partial t} - \Delta w + F'(y')w = -F''(\theta y + (1-\theta)y')(y-y')\chi_i \\ w|_{\partial\Omega} = 0 \\ w|_{t=0} = 0 \end{cases}$$

which shows, as $\|F'(y')\|_{L^\infty(Q)}$, $\|F''(\ y + (1- \ )y')\|_{L^\infty(Q)}$, $\|\chi_i\|_{L^\infty(Q)}$ are bounded independently of y', that the mapping $y \to \chi_i$ defined by (49) is continuous from $L^2(Q)$ into $L^2(Q)$. Hence the mapping $a \to \chi_i(a)$ defined by (47), (40), (49) is continuous from $\mathbb{R}^m$ into $L^2(Q)$, which proves that ϕ is continuously differentiable.

Similarly, the second partial derivations $\mu_{ij} = \frac{\partial^2 \phi}{\partial a_i \partial a_j}(a)$ exist and are given by :

$$(51) \begin{cases} \frac{\partial \mu_{ij}}{\partial t} - \Delta\mu_{ij} + F'(y)\mu_{ij} = -F''(y)\chi_i\chi_j \\ \mu_{ij}|_{\partial\Omega} = 0 \\ \mu_{ij}|_{t=0} = 0 \end{cases}$$

(this can be proved by the same kind of techniques as in the proof of lemma 6). From (51) one deduces, as F" and F"' are bounded and χ_i, $\chi_j \in L^\infty(Q)$, that μ_{ij} depends continuously, in $L^2(Q)$, on y in $L^2(Q)$, and hence on $a \in \mathbb{R}^m$, which ends the proof of the lemma. ∎

Hence theorem 4 applies for this parameter estimation problem.

Example 2 : We consider a system of ordinary differential equations :

(52) $y' = A(y).a$

(53) $y'(0) = v$

where

$$(54)\quad \begin{cases} y(t) \in \mathbb{R}^n & \text{is the "state" at time } t \\ a \in \mathbb{R}^m & \text{is the vector of unknown coefficients} \\ A(y) \in \mathcal{M}_{n\times m} & \text{is a } n\times m \text{ matrix whose elements are } \mathcal{C}^\infty \text{ know functions (usually} \\ & \text{polynomials in the chemical applications) of the } y = (y_1,\ldots,y_n). \\ v \in \mathbb{R}^n & \text{is the known initial datum.} \end{cases}$$

and we want to estimate the parameter vector a from the measurement, say of one linear combination of the y_j's, continuously over the time interval $[0,T]$: the measured data z is here :

(55) $z : t \quad [0,T] \to z(t) \simeq \alpha\, y(t) \in \mathbb{R}$

where

(56) $\alpha \in \mathbb{R}^{n*}$ is a given vector (observation operator).

The mapping ϕ of the theory is here :

(57) $\phi : a \in \mathbb{R}^m \to (t \to \alpha\, y(t)) \in \mathcal{C}^1([0,T])$

This mapping is clearly $\mathscr{C}^\infty$ over $\mathbb{R}^m$, so that the upper bounds δ and β of ϕ' and ϕ'' over any given compact set K of $\mathbb{R}^m$ exist. In order to get the existence of α we have to study the injectivity of ϕ.

For $\delta a \in \mathbb{R}^m$, let $\delta y \in \mathscr{C}'([0,T] ; \mathbb{R}^n)$ and $\delta\phi \in \mathscr{C}^1([0,T])$ be the corresponding differentials, given by :

(58) $$\begin{cases} \delta y' = B(y,a)\delta y + A(y)\delta a \\ \delta y(0) = 0 \end{cases}$$

(59) $$\delta\phi(t) = \alpha\delta y(t) \qquad \forall t \in [0,T]$$

where :

(60) $$B(y,a) = \frac{\partial}{\partial y}[A(y).a] \text{ is a n n matrix.}$$

As (58) is a linear system of differential equations, its solution can be explicitely written as :

(61) $$\delta y(t) = \left\{ \int_0^t e^{C(t)-C(s)} A(y(s))ds \right\} \delta a$$

where

(62) $$C(t) = \int_0^t B(y(s),a)ds$$

Hence :

(63) $$\delta\phi(t) = \left\{ \int_0^t e^{C(t)-C(s)} A(y(s))ds \right\} \delta a \qquad \forall t \in [0,T]$$

which give immediately the :

Lemma 8 : A necessary and sufficient condition for $\phi'(a)$ to be injective is that there exists m times $t_j \in [0,T]$ $j = 1,2,\ldots,m$ such that the m vectors of $\mathbb{R}^{m*}$:

$$(64) \qquad \int_0^{t_j} e^{C(t_j)-C(s)} A(y(s))ds \qquad j = 1,2,\ldots,m$$

are linearly independent.

Of course, this condition is not very useful, as the vectors in (64) are not easy to calculate.

We give now a less accurate, but sufficient condition for the injectivity of $\phi'(a)$: we look for conditions on the initial value v and the parameter value a such that $\delta\phi \equiv 0 \Rightarrow \delta a = 0$.

From $\delta\Phi \equiv 0$ we see that :

$$(65) \qquad \psi(t) = \alpha B(y(t),a)\delta y(t) + \alpha A(y,t))\delta a = 0 \qquad \forall t \in [0,T]$$

which implies that :

$$(66) \qquad \psi^{(j)}(0) = 0 \qquad j = 0,1,\ldots,m-1$$

But the derivatives $\psi^{(j)}(t)$ can be calculated in terms of y, a and δa using (65) and (58) by a simple (but tedious as j increases !) calculation. This can be written as :

$$(67) \qquad \begin{cases} \psi^{(j)}(0) = d_j(v,a,\alpha).\delta a \\ \\ d_j(v,a) \in \mathbb{R}^{m*} \end{cases} \qquad j = 0,1,\ldots,m-1$$

Lemma 9 : A sufficient condition for $\phi'(a)$ to be injective is that the m vectors $d_j(v,a,\alpha) \in \mathbb{R}^{m*}$, $j = 0,1,\ldots,m-1$ defined in (57) are linearly independent.

Using lemma 8 or 9 one can define a subset $\mathcal{A}_{inj}$ of $\mathbb{R}^m$ on which $\phi'(a)$ is necessarily injective, and then apply theorem 4 on any (convex) compact set K of $\mathcal{A}_{inj}$.

ACKNOWLEDGEMENT

I am very much indebted to Luc TARTAR who provided the proof of theorems 1 and 2.

REFERENCES

[1] G. CHAVENT
"About the stability of the optimal control solution of inverse problems", in "Mathematical and Numerical Methods of Inverse and Improperly Posed Problems", Ed. by G. ANGER, Akademie Verlag, Berlin, (G.D.R.), 1979.

[2] K. MARTENSSON
"Least Square Identifiability of Dynamic Systems", Technical Report RB 7344, Dept. Electr. Eng., U.S.C. Los Angeles, 1973.

[3] S. KITAMURA a,d S. NAKAGIRI
"Identifiability of Spatially Varying and Constant Parameters in Distributed Systems of Parabolic Type", SIAM J. Control and Optimization, Vol. 15, n° 5, August 1977.

[4] C.S. KUBRUSLY
"Distributed Parameter System Identification, a Survey", Int. J. of Control, Vol. 26, n° 4, 509-535, 1977.

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique