OUTPUT LEAST SQUARES STABILITY FOR ESTIMATION OF THE DIFFUSION COEFFICIENT IN AN ELLIPTIC EQUATION

F. Colonius*

Universität Frankfurt, The Federal Republic of Germany

K. Kunisch*

Technische Universität Graz, Austria

ABSTRACT

The estimation of unknown coefficients in partial differential equations is frequently studied as an output least squares problem involving an "observation" of the system for which the model is derived and the solution of the model equation as a function of the unknown parameter. We study the continuous dependence of the output least squares formulation on the observation of the system. There is no a-priori assumption on the uniqueness of the output least squares solutions.

OUTPUT LEAST SQUARES STABILITY

We study estimation of the diffusion coefficient $q = \mathrm{col}(q_1, \dots, q_n)$ in the elliptic equation

$$\begin{cases} -\sum_{i=1}^{n} (q_i u_{x_i})_{x_i} + c\,u = f \quad \text{in} \quad \Omega \subset \mathbb{R}^n \\ u \mid \partial\Omega = 0 \end{cases} \tag{1}$$

where $f \in L^2$ and $c \in L^2$ with $c \geq 0$. We assume that $n = 2$ or 3 but with the appropriate changes the case of arbitrary n can be treated with the same techniques. All function spaces are taken over the bounded domain Ω, which is assumed to have a smooth $(C^{2,-})$ boundary or to be a parallelepiped. Let $u = u(q)$ denote the solution of (1) corresponding to the diffusion coefficient q, and let $z^o \in L^2$ be

* Both authors acknowledge support from the Fonds zur Förderung der wissenschaftlichen Forschung, under grant S 3206.

an observation of the (e.g. physical) system for which (1) is a proposed model equation. Due to model and observation error there may or may not exist q^* in a set of admissible parameters Q_{ad} to be defined below, which satisfies $u(q^*) = z^o$. To estimate the unknown parameter q so that the corresponding solution $u(q)$ best fits the data we adopt the output least squares method

$$(OLS)_{z^o} \qquad \min \; |u(q) - z^o|^2$$

where q is chosen from the following set Q_{ad} of admissible parameters:

$$Q_{ad} = \{q \in Q : 0 < k_i(x) \leq q_i(x),\ x \in \Omega,\ i=1,\dots,n,\ |q|_Q \leq \gamma\}$$

where $k_i \in H^2$ and $\gamma > |col(k_1,\dots,k_n)|_Q$ are given. Here $Q = \bigotimes_{i=1}^{n} H^2$ is endowed with the Hilbert space product topology. Recall that $H^2 \subset C$ is a continuous embedding for $n \leq 3$.

Our objective in this note is to study the stability of the solution of $(OLS)_{z^o}$ on z^o. Solving $(OLS)_{z^o}$ involves the inversion of $\Phi : q \to u(q)$. Without strong assumptions on the problem data, Φ is not injective. Moreover, considering Φ^{-1} as a multivalued mapping, this inverse is not continuous unless the parameter space is endowed with a sufficiently weak topology. Along with $(OLS)_{z^o}$ we consider therefore the following regularized output least squares problem for $\beta \in \mathbb{R}$, $\beta > 0$:

$$(ROLS)^{\beta}_{z^o} \qquad \min \; |u(q) - z^o|^2 + \beta|q|^2_Q \quad \text{over } Q_{ad}.$$

Here and below we drop the index for the inner product and norm in L^2. The behaviour of the solutions of $(ROLS)^{\beta}_{z}$ as $\beta \to 0^+$ has been studied e.g. in [3-6].- We point out the fact that $(ROLS)^{\beta}_{z^o}$ has a solution q_{z^o} for every $\beta > 0$ without the necessity of a norm constraint in the definition of Q_{ad}, whereas $(OLS)_{z^o}$ may have no solution unless Q_{ad} is a bounded set. The results of this paper remain unchanged if the norm constraint in the definition of Q_{ad} is replaced by the assumption that $(OLS)_{z^o}$ has a solution.

We define the following stability concept for the solutions of $(\mathrm{ROLS})^{\beta}_{z^o}$.

DEFINITION. The parameter q is called Output Least Squares stable by Regularization (ROLS-stable) at $z^o \in L^2$ in Q_{ad} for β in the interval $I \subset (0,\infty)$ if for every $\beta \in I$ and every global solution $q^{\beta}_{z^o}$ of $(\mathrm{ROLS})^{\beta}_{z^o}$ there exists a constant $d > 0$ and neighborhoods $V(z^o)$ of z^o in L^2 and $V(q^{\beta}_{z^o})$ of $q^{\beta}_{z^o}$ in Q such that for all $z \in V(z^o)$ there exists a local solution $q^{\beta}_{z} \in V(q^{\beta}_{z^o})$ of $(\mathrm{ROLS})^{\beta}_{z}$ and for every solution $q^{\beta}_{z} \in V(q^{\beta}_{z^o})$ we have $|q^{\beta}_{z} - q^{\beta}_{z^o}|_Q \leq d\ |z-z^o|^{1/2}_{L^2}$.

Remark 1. The method which we shall propose to study ROLS-stability will also allow to analyse continuous dependence on other parameters in the problem, as for example on k_i and γ (see [5]).

Remark 2. ROLS-stability requires that for every $\beta \in I$ every solution of $(\mathrm{ROLS})^{\beta}_{z}$ depends continuously on z. It does not presuppose injectivity of the mapping Φ or uniqueness of the solutions of $(\mathrm{OLS})_{z^o}$ or $(\mathrm{ROLS})^{\beta}_{z^o}$. However, once ROLS-stability is established it follows that the solutions of $(\mathrm{ROLS})^{\beta}_{z^o}$ are isolated.

First we summarize some properties of equation (1). Let

$$U = \{q \in Q : q_i(x) \geq k_i(x)/2,\ x \in \Omega,\ i=1,\dots,n\}.$$

It is then wellknown [7] that for every $f \in L^2$ and $q \in U$ there exists a unique solution $u(q) \in H^2$ of (1). Moreover there exists a nondecreasing positive function $\tilde{c}_1(\cdot)$ depending only on c and k_i such that

$$|u(q)|_{H^2} \leq \tilde{c}_1(|q|_Q)\ |f|_{L^2} \tag{2}$$

for $q \in U$ and $f \in L^2$. For $q \in U$ we define operators $A(q)$ in L^2 by $D(A(q)) = H^2 \cap H^1_o$ and $A(q)u = -\sum_{i=1}^{n} (q_i u_{x_i})_{x_i} + cu$. These operators are densely defined, closed and selfadjoint. By the above considerations they also are homeomorphisms between $D(A(q))$ endowed with the H^2-norm and L^2 .

We further put

$$A_1(q)u = -\sum_{i=1}^{n} (q_i u_{x_i})_{x_i} \quad \text{for} \quad u \in D(A(q)) \quad \text{and} \quad q \in U .$$

Observe that $A_1(q)$ is selfadjoint for every $q \in U$ and that for some constant K_1

$$|A_1(q)u| \leq K_1 |q|_Q \, |u|_{H^2} \quad \text{for all} \quad u \in H^2 \cap H_o^1 \quad \text{and} \quad q \in Q . \qquad (3)$$

Finally one can show that

$$q \to A^{-1}(q)f = u(q)$$

from $U \subset Q$ to L^2 is continuous from the weak to the strong topology. In order to reformulate $(\mathrm{ROLS})^{\beta}_{z^o}$ in an abstract setting we define $g : Q \to \bigotimes_{i=1}^{n} C \times \mathbb{R}$ by

$$g(q) = (\mathrm{col}(k_1-q_1,\ldots,k_n-q_n) \;, \quad |q|_Q^2 - \gamma^2)$$

and a closed convex cone K in $\bigotimes_{i=1}^{n} C \times \mathbb{R}$ with vertex at the origin by

$$K = \bigotimes_{i=1}^{n} C^+ \times \mathbb{R}^+ ,$$

where C^+ is the natural nonnegative cone in C. Then $(\mathrm{ROLS})^{\beta}_{z^o}$ is equivalent to

$$\min |A^{-1}(q)f-z|^2 + \beta|q|^2 \quad \text{over} \quad q \in Q_{ad} = \{q \in Q : g(q) \in -K\} .$$

In order to discuss ROLS-stability we use a stability result for optimization problems in infinite dimensional spaces. Let Q and W be Hilbert spaces, and Y a Banach space with an ordering induced by a closed convex cone K with vertex at the origin. Suppose that $f : D \times W \to \mathbb{R}$ and $g : X \times W \to Y$ and that $D \subset Q$ is an open set satisfying

$$Q_{ad} = \{q \in Q : g(q,w) \in -K\} \subset D .$$

Let w^o be a fixed reference parameter and for arbitrary $w \in W$ consider

$$(P)_w \qquad \min f(q,w) \text{ such that } g(q,w) \in -K .$$

A functional $\lambda^* \in Y^*$, the dual of Y , is called Lagrange multiplier for $(P)_{w^o}$ at q^o if

$$f_q(q^o,w^o) + \lambda^* g_q(q^o,w^o) = 0 , \quad \lambda^* \in K^+ \quad \text{and} \quad \lambda^* g(q^o,w^o) = 0 \qquad (4)$$

and $F : Q \to \mathbb{R}$ given by $F(q,w^o) = f(q,w^o) + \lambda^* g(q,w^o)$ is called the associated Lagrange functional. Here K^+ is the dual cone of K. If q^o is a solution of $(P)_{w^o}$, then there exists a Lagrange multiplier, provided that q^o is a regular point with respect to the constraint, i.e.

$$0 \in \operatorname{int} \{g(q^o,w^o) + g'(q^o,w^o)Q + K\} .$$

PROPOSITION 1. [1]. Let q^o be a regular point and suppose that f and g are twice continuously Fréchet-differentiable with respect to q at (q^o,w^o) and that there exist constants $\nu > 0$ and $\gamma > 0$ such that for a Lagrangian functional F

$$F_{qq}(q^o,w^o)(h,h) \geq \gamma |h|^2$$

holds for all $h \in g_q^{-1}(-K + \mathbb{R}\, g(q^o)) \cap \{h : \lambda^* g_q(q^o,w^o)h \leq \nu |h|\}$.
Moreover assume that there exists a neighborhood $U = U_q \times U_w$ of (q^o,w^o) and constants L_q and L_g such that

$$|f(q,w) - f(q',w^o)| \leq L_f(|q-q'| + |w-w^o|)$$

$$|g(q,w) - g(q,w^o)| \leq L_g |w-w^o|$$

for all $(q,w) \in U$ and $q' \in U_q$.
Then there exist $r > 0$, $d > 0$ and a neighborhood V of w^o such that:

(i) The local extremal value function

$$\mu_r(w) = \inf \{f(q,w) : g(q,w) \in -K, |q-q^o| \leq r\}$$

is Lipschitz continuous at w^o.

For every $w \in V$ the following additional statements hold:

(ii) For every sequence $\{q_n\}$ with $g(q_n,w) \in -K$, $|q-q_n| \leq r$ and $\lim_n f(q_n,w) = \mu_r(w)$ it follows that $|q_n-q^o| < r$ for all n sufficiently large.

(iii) If there exists q_w with $g(q_w,w) \in -K$, $|q_w-q^o| \leq r$ and $f(q_w,w) = \mu_r(w)$, then $|q_w-q^o| < r$ and $|q_w-q^o| \leq d\ |w-w^o|^{1/2}$.

Clearly, $(ROLS)_z^\beta$ can be considered as a special case of $(P)_w$. Furthermore existence of solutions q_z^β as required in (iii) is guaranteed.

LEMMA 1. The first and second derivatives $\eta = u_q(q)h$ and $\xi = u_{qq}(q)(h,h)$ of the solution u at q in directions $h \in Q$ are characterized by the property that η and $\xi \in H^2 \cap H_o^1$ and

$$A(q)\eta = -A_1(h)u(q),$$

$$A(q)\xi = -2A_1(h)\eta.$$

Here $q \to u(q)$ is taken as a mapping from Q to $H^2 \cap H_o^1$.

In order to formulate our first main result we introduce

$$Q^\beta = \{q^\beta \in Q : q^\beta \text{ is a solution of } (ROLS)^\beta_{z^o}\}$$

and the attainable set

$$V = \{u(q) : q \in Q_{ad}\}.$$

Here and below we drop the index z^o for the notation of the solutions of $(ROLS)^\beta_{z^o}$.

THEOREM 1. Let $\bar{\beta} > 0$ be chosen such that for a minimum norm solution q_m^o of $(OLS)_{z^o}$

$$|q_m^o|^2 - \sup_{Q^\beta} |q^\beta|^2 < (K_1 \tilde{c}_1(|q_m^o|))^{-2} \tag{5}$$

and define

$$\underline{\beta} = \text{dist } (z^o,V)^2 \ [(K_1 \tilde{c}_1(|q_m^o|))^{-2} - |q_m^o| + \sup_{Q^{\bar{\beta}}} |q^{\bar{\beta}}|]^{-1} \geq 0. \tag{6}$$

If $\underline{\beta} < \bar{\beta}$ then the diffusion coefficient q in (OLS) is ROLS-stable at z^o in Q_{ad} for $\beta \in (\underline{\beta},\bar{\beta})$. In particular, if $z^o \in V$, then q is ROLS-stable in Q_{ad} for all $\beta \in (0,\bar{\beta})$.

Proof. The second assertion is a direct consequence of the first one which we verify by means of Proposition 1. In the notation of Proposition 1 we take $Y = \bigotimes_{i=1}^{n} C \times \mathbb{R}$, $K = \bigotimes_{i=1}^{n} C^+ \times \mathbb{R}^+$ and $W = L^2$. Here C^+ denotes the cone of nonnegative functions in C. Since the regular point condition and the necessary smoothness requirements are quite easy to verify we concentrate on the second order sufficiency condition. Let λ^* be a Lagrange multiplier in the dual cone of $\bigotimes_{i=1}^{n} C^+ \times \mathbb{R}^+$ and consider the corresponding Lagrange functional

$$F(q) = F(q,z^o) = |u(q)-z^o|^2 + \langle\lambda^*, g(q,z^o)\rangle + \beta|q|_Q^2 .$$

We compute for $q^\beta \in Q^\beta$ and $\beta \in (\underline{\beta},\bar{\beta})$

$$\begin{aligned} F_{qq}(q^\beta)(h,h) = {} & 2|\eta|^2 + 2(u(q^\beta)-z^o,\xi) + \\ & + \langle\lambda^*, g_{qq}(q^\beta)(h,h)\rangle + 2\beta|h|_Q^2 , \end{aligned} \tag{7}$$

where $\eta = u_q(q^\beta)(h)$ and $\xi = u_{qq}(q^\beta)(h,h)$. Using Lemma 1 and the fact that $g_{qq}(q^\beta)(h,h) \in K$ we find

$$\begin{aligned} F_{qq}(q^\beta)(h,h) & \geq 2|\eta|^2 - 4(u(q^\beta)-z^o,\ A^{-1}(q^\beta)A_1(h)\eta) + 2\beta|h|_Q^2 \\ & = 2|\eta|^2 - 4(A_1(h)A^{-1}(q^\beta)(u(q^\beta)-z^o),\eta) + 2\beta|h|_Q^2 \\ & \geq 2\beta|h|^2 - 2\ |A_1(h)A^{-1}(q^\beta)(u(q^\beta)-z^o)|^2 . \end{aligned} \tag{8}$$

By (2) and (3)

$$|A_1(h)A^{-1}(q^\beta)(u(q^\beta)-z^o)| \le K_1|h|_Q\, \tilde{c}_1(|q^\beta|)\; |u(q^\beta)-z^o| \quad .$$

This estimate is now used in (8) and we find

$$F_{qq}(q^\beta)(h,h) \ge 2|h|^2\, [\beta - K_1^2\, \tilde{c}_1(|q^\beta|)^2\, |u(q^\beta)-z^o|^2] \quad .$$

On the other hand, we have for every minimum norm solution q_m^o of $(OLS)_{z^o}$ and every $q^\beta \in Q^\beta$

$$-|u(q^\beta)-z^o|^2 \ge \beta(|q^\beta|^2 - |q_m^o|^2) - \text{dist}\,(z^o,V) \;,$$

see [5]. Therefore with $K_2 = K_1\, \tilde{c}_1(|q^\beta|)$

$$\begin{aligned} F_{qq}(q^\beta)(h,h) &\ge 2|h|^2\, K_2^2\, [\beta(K_2^{-2} + |q^\beta|^2 - |q_m^o|^2) - \text{dist}\,(z^o,V)^2] \\ &\ge 2|h|^2\, K_2^2\, \text{dist}\,(z^o,V)^2\, [\underline{\beta}^{-1}\, \beta-1] \;. \end{aligned}$$

This implies $F_{qq}(q^\beta_{z^o})(h,h) \ge \text{const}\, |h|^2$ for every $\beta \in (\underline{\beta},\overline{\beta})$ and the proof is finished.

Theorem 1 gives results that guarantee continuous dependence of the solutions of $(ROLS)^\beta_{z^o}$ on the observation z^o. In the first part of the assertion there is no attainability assumption, but $\underline{\beta}$ may be greater than $\overline{\beta}$, in which case the assertion is void. In this case decreasing dist (z^o,V) either by a more accurate measurement z^o or an improvement of the model should lead to success. - Thus consider $z_n^o \to z^o$ in L^2 with $z^o \in V$. We denote the solutions of $(OLS)_{z_n^o}$ and $(ROLS)^\beta_{z_n^o}$ by $q^o_{z_n^o}$ and $q^\beta_{z_n^o}$. The following stability property can be obtained for z_n^o sufficiently close to $z^o \in V$. (Recall that stability is investigated with respect to the upper index in z_n^o.) By Q_z^o we denote the set of solutions of the unregularized problem $(OLS)_z$.

THEOREM 2. Let the assumptions of Theorem 1 hold. Choose a sequence of observations z_n^o with $z_n^o \to z^o$ in L^2 and $z^o \in V$. Then there exists

$\tilde{\beta}>0$ with the following property: For all $\beta^* \in (0,\tilde{\beta})$ there exists an index $N(\beta^*) \in \mathbb{N}$ and a neighborhood $I(\beta^*)$ of β^* such that for all $n \geq N(\beta^*)$ the parameter q is ROLS-stable in Q_{ad} at z_n^o for all $\beta \in I(\beta^*)$.

For the proof we refer to [5].

DISCUSSION

1) We recall that G. Chavent [2,3] has introduced a stability concept for parameter estimation problems which is different from ours. A parameter is called output least squares identifiable (OLSI), if there exists a neighborhood $\tilde{V}$ of the attainable set V such that for every element $z \in \tilde{V}$ there exists a unique solution $q \in Q_{ad}$ of $(OLS)_z$ depending continuously on z. Chavent derives general sufficient conditions involving $\mathrm{dist}\,(z, V)$, $\mathrm{diam}\,(Q_{ad})$, lower and upper bounds on the first and second derivative of $q \to u(q)$ which imply OLSI. The main distinction between the stability concept introduced in this paper and OLSI is the fact that OLSI requires uniqueness of the solution of the output least squares problem whereas ROLS-stability (or also OLS-stability, see 4)) does not.

2) In a recent paper, Kravaris and Seinfeld [6] have studied the use of regularization for parameter estimation in parabolic partial differential equation. Their approach is based on a variant of Tikhonov's lemma which states that if a continuous function f between metric spaces X and Y is injective on a precompact subset $K \subset X$, then f is continuously invertible on $f(K)$. In applications to $(OLS)_{z^o}$ this requires that the regularization term in $(ROLS)^{\beta}_{z^o}$ is replaced by $\beta|q|_{Q_c}$, where Q_c is a compactly embedded subspace of Q with norm $|\cdot|_{Q_c}$. In computations this leads to some inconvenience, since it is more involved to implement the Q_c- than the Q-norm. Moreover the Tikhonov approach of [6] requires uniqueness of the solution of the unregularized problem, which is not needed in our analysis, where stability is checked at each solution of $(ROLS)_{z^o}$. If ROLS-stability can be guaranteed then the solutionsof $(ROLS)_{z^o}$ are isolated.

3) The second order sufficient condition (see (7) and (8) in the proof of Theorem 1) can be used as a convenient tool to obtain some insight into the specific features of parameter estimation problems.

   (a) If $z^o \in V$, then in view of the term $\langle u(q^\beta) - z^o, \xi \rangle$ and the convergence of $u(q^\beta)$ to z^o as $\beta \to 0^+$ [5], the lower bound on $F_{qq}(q^\beta)(h,h)$ is easier to obtain than in case that $z^o \notin V$.

   (b) The advantage of a regularization term is obvious from (7),(8). It helps to achieve strict positivity of $F_{qq}(q^\beta)$ and the required second order sufficiency condition.

   (c) To explain the next observation, suppose that $z^o \in V$ and take $\beta = 0$. Then the second derivative of the Lagrange functional reduces to $|\eta| = |A^{-1}(q^\beta)\nabla(h\nabla u(q^\beta))|$ and it is apparent that continuous dependence of q in H^2 on z cannot hold, since $|A^{-1}(q^\beta)\nabla(h\nabla u(q^\beta))|$ can be bounded below by $|\nabla h \nabla(u(q^\beta))|_{H^{-2}}$ only and since, moreover, our conditions do not exclude the case meas $\{x : \nabla u(q^\beta)(x) = 0\} > 0$. Hence some kind of regularization is necessary.

4) If for $\beta = 0$ the same kind of stability of the solutions q_z of $(OLS)_z$ on z as required for ROLS-stability holds, then we call q OLS-stable. Special cases of OLS-stability are studied in [5].

REFERENCES

[1] W. Alt: Lipschitzian perturbations of infinite optimization problems; in: Mathematical Programming with Data Perturbations II, ed. A.V. Fiacco, Lecture Notes in Pure and Applied Mathematics 85, Marcel Dekker, New York, 1983, 7-21.

[2] G. Chavent: Local stability of the output least square parameter estimation technique, Matematica Applicada e Computacional, 2 (1983), 3-22.

[3] G. Chavent: On parameter identifiability, Proceedings of the 7-th IFAC Symposium on Identification and System Parameter Estimation, York, England, July 1985.

[4] F. Colonius and K. Kunisch: Stability for parameter estimation in two point boundary value problems, to appear in Journal Reine Angewandte Mathematik.

[5] F. Colonius and K. Kunisch: Output least squares stability in elliptic systems, submitted to Appl. Math. Optimization.

[6] C. Kravaris and J.H. Seinfeld: Identification of parameters in distributed systems by regularization, SIAM J. Control and Optimization 23 (1985), 217-241.

[7] O.A. Ladyzhenskaya and N.N. Ural'tseva: Linear and quasilinear elliptic equations, Academic Press 1986.

[8] D.L. Russell: Some remarks on numerical aspects of coefficient identification in elliptic systems, in: Optimal Control of Partial Differential Equations, ed. K.H. Hoffmann and W. Krabs, Birkhäuser 1984, 21o-228.