WORKING PAPER

ON NONINVERTIBLE EVOLUTIONARY SYSTEMS: GUARANTEED ESTIMATES AND THE REGULARIZATION PROBLEM

A.B. Kurzhanski
I.F. Sivergina

November 1989 (Revised Version)
WP-89-058

IIASA
International Institute
for Applied Systems Analysis

ON NONINVERTIBLE EVOLUTIONARY SYSTEMS: GUARANTEED ESTIMATES AND THE REGULARIZATION PROBLEM

A.B. Kurzhanski
I.F. Sivergina

November 1989 (Revised Version)
WP-89-058

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Foreword

This paper deals with an *inverse problem*: the estimation of an initial distribution in the first boundary value problem for the heat equation through some biassed information on its solution. Numerically stable solutions to the inverse problem are normally achieved through various regularization procedures. It is shown that these procedures could be treated within a unified framework of solving guaranteed estimation problems for systems with unknown but bounded errors.

A. Kurzhanski
System and Decision Sciences Program.

On Noninvertible Evolutionary Systems: Guaranteed Estimates and the Regularization Problem

A.B. Kurzhanski

International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

I.F. Sivergina

Institute of Mathematics and Mechanics of the Ural Scientific Center
Academy of Sciences of the USSR, Sverdlovsk, USSR

This paper deals with the selection of an initial distribution in the first boundary-value problem for the heat equation in a given domain $[0,\theta] \times \Omega$, $\theta < \infty$ with zero values on its boundary S so that the deviation of the respective solution from a given distribution would not exceed a preassigned value $\gamma > 0$. The result is formulated here in terms of the "theory of guaranteed estimation" for noninvertible evolutionary systems. It also allows an interpretation in terms of regularization methods for ill-posed inverse problems and in particular, in terms of the quasiinvertibility techniques of J.-L. Lions and R. Lattes.

1. The Problem.

Assume Ω to be a compact domain in $\mathbf{R}^n$ with a smooth boundary S; $\theta > 0$, $\gamma > 0$ to be given numbers, functions $y(t,x)$, $z(x)(\mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^1)$, $(\mathbf{R}^n \to \mathbf{R}^1)$ to be given and such that $y(\cdot,\cdot) \in L_2([0,\theta] \times \Omega)$, $z(\cdot) \in L_2(\Omega)$.

Denote $u = u(t,x;\, w(\cdot))$ to be the solution to the boundary value problem

$$\frac{\partial u}{\partial t} - \Delta u = 0 \,, \quad 0 \le t \le \theta \,, \tag{1}$$

$$u\big|_{[0,\theta]\times S} = 0 \,,$$

$$u|_{t=0} = w(\cdot)\ ,$$

Also denote

$$J(w(\cdot)) = \alpha \int_0^\theta \int_\Omega (u(t,x;\ w(\cdot)) - y(t,x))^2 dxdt + \tag{2}$$

$$+ \beta \int_\Omega (u(\theta,x;\ w(\cdot)) - z(x))^2 dx$$

with $\alpha \geq 0$, $\beta \geq 0$.

Consider the following problem: among the possible initial distributions $w(\cdot) \in L_2(\Omega)$ specify a distribution $w^0(\cdot)$ that ensures

$$J(w^0(\cdot)) \leq \gamma\ . \tag{3}$$

The latter is an ***inverse problem*** [1]. With $\alpha = 0$ it was studied by J.-L. Lions and R. Lattes within the framework of the method of "quasiinvertibility" [2]. Numerical stability was ensured in this approach.

Let us now transform the previous problem into the following: among the distributions $w(\cdot) \in L_2(\Omega)$ determine *the set* $W^*(\cdot) = \{w^*(\cdot)\}$ *of all those distributions* $w^*(\cdot)$ *that yield the inequality*

$$J(w^*(\cdot)) \leq \gamma\ .$$

Assuming that the problem is solvable ($W^*(\cdot) \neq \phi$) we may describe its solution in terms of the theory of "guaranteed observation" [3]. Namely, assume $y(t,x)$, $z(x)$ to be the available measurements of the process (1), so that

$$y(t,x) = u(t,x;\ w(\cdot)) + \xi(t,x) \tag{4}$$

$$z(x) = u(\theta,x;\ w(\cdot)) + \sigma(x)$$

$$0 \leq t \leq \theta,\ \ x \in \Omega$$

where $\xi(t,x)$, $\sigma(x)$ stand for the *measurement noise* which is ***unknown*** in advance *but bounded* by the restriction

$$\alpha \int_0^\theta \int_\Omega \xi^2(t,x)\ dxdt + \beta \int_\Omega \sigma^2(x) dx \leq \gamma\ . \tag{5}$$

Then $W^*(\cdot)$ will be precisely the *set of all initial states* of system (1) *consistent with measurements* $y(t,x)$, $z(x)$ (4) *and with restriction* (5).

The aim of this paper will be to describe some stable schemes of calculating the sets $W^*(\cdot)$ and their specific elements. (A direct calculation of these may obviously lead to unstable numerical procedures.)

2. The Regularizing Problem (A General Solution)

Consider a rather general problem. Assume the values ξ, σ, w to be unknown in advance while satisfying a joint quadratic constraint

$$\langle w(\cdot),\, N(\varepsilon) w(\cdot)\rangle + \alpha \int_0^\theta \langle \xi(t,\cdot),\, M(\varepsilon)\xi(t,\cdot)\rangle\, dt \tag{6}$$

$$+ \beta\, \langle \sigma(\cdot),\, K(\varepsilon)\sigma(\cdot)\rangle \le \gamma + h_\varepsilon\,, \quad h_\varepsilon > 0\,,$$

Here $N(\varepsilon)$, $M(\varepsilon)$, $K(\varepsilon)$ are nonnegative self-adjoint operators from $\mathcal{L}_2(\Omega)$ into itself (with $N(\varepsilon)$ invertible) and such that each of them depends on a small parameter $\varepsilon>0$. The symbol $\langle\cdot,\cdot\rangle$ denotes a scalar product in $L_2(\Omega)$.

An *informational set* $W_\varepsilon(\cdot)$ of distributions $w(\cdot)$ consistent with measurements y and z will be defined as the variety of those and only those functions $w(\cdot) \in L_2(\Omega)$ for each of which there exists such a pair $\xi(\cdot,\cdot) \in \mathcal{L}_2([0,\theta] \times \Omega)$ and $\sigma(\cdot) \in \mathcal{L}_2(\Omega)$ that equalities (1), (4) would be fulfilled together with the inequality (6).

Lemma 2.1. The informational set $W_\varepsilon(\cdot)$ *consists of all those functions* $w(\cdot) \in \mathcal{L}_2(\Omega)$ *that satisfy the inequality*

$$\langle w(\cdot) - w_\varepsilon^0(\cdot), \mathbf{B}(\varepsilon)(w(\cdot) - w_\varepsilon^0(\cdot))\rangle \le \gamma + h_\varepsilon - \kappa_\varepsilon^2 \tag{7}$$

where

$$\mathbf{B}(\varepsilon) = N(\varepsilon) + U^* M(\varepsilon) U + U_\theta^* K(\varepsilon) U_\theta\,,$$

$$w_\varepsilon^0(\cdot) = \mathbf{B}^{-1}(\varepsilon)(U^* M(\varepsilon) y(\cdot,\cdot) + U_\theta^* K(\varepsilon) z(\cdot))\,,$$

$$\kappa_\varepsilon^2 = \langle z(\cdot),\, K(\varepsilon) z(\cdot)\rangle + \int_0^\theta \langle\, y(t,\cdot),\, M(\varepsilon) y(t,\cdot)\rangle\, dt\,,$$

$$(Uw(\cdot))(t,x) = u(t,x;\, w(\cdot))\,, \quad (U_\theta w(\cdot))(x) = u(\theta,x;\, w(\cdot))\,;$$

$$U : \mathcal{L}_2(\Omega) \rightarrow \mathcal{L}_2([0,\theta] \times \Omega);\ U_\theta : \mathcal{L}_2(\Omega) \rightarrow \mathcal{L}_2(\Omega)$$

and where U^* stands for the respective adjoint operator.

It is further assumed that h_ε is such that $W_\varepsilon(\cdot)$ is nonvoid.

If there exists an $\varepsilon_0 \geq 0$ such that

$$J(w_\varepsilon^0(\cdot)) \rightarrow \inf \{J(w(\cdot)) \mid w(\cdot) \in \mathcal{L}_2(\Omega)\}$$

$$\text{with } \varepsilon \rightarrow \varepsilon_0$$

then the problem of estimating the distributions $w(\cdot)$ due to the system (1), (4), (6) will be further referred to as the ***regularizing problem*** for problem (1), (3).

3. Quasiinvertibility

With $\alpha = 0$ in equation (2) we arrive at the problem investigated in [2] by means of the *quasiinvertibility* techniques. Following the latter consider an auxiliary boundary-value problem

$$\frac{\partial V_\varepsilon}{\partial t} - \Delta V_\varepsilon - \varepsilon \Delta^2 V_\varepsilon = 0\ ,\ 0 \leq t \leq \theta\ ,\ (\varepsilon > 0)$$

$$V_\varepsilon|_{[0,\theta] \times S} = \Delta V_\varepsilon|_{[0,\theta] \times S} = 0$$

$$V_\varepsilon|_{t=\theta} = z(\cdot)\ .$$

Then taking

$$w_\varepsilon(\cdot) = V_\varepsilon(0,\cdot) \tag{8}$$

we come to

$$J(w_\varepsilon(\cdot)) \rightarrow 0 \quad (\varepsilon \rightarrow 0)\ .$$

The following question does arise: is it possible to select the operators $N(\varepsilon)$, $M(\varepsilon)$, $K(\varepsilon)$ that define the quadratic constraint (6) in such a way that the center $w_\varepsilon^0(\cdot)$ of the informational ellipsoid $W_\varepsilon(\cdot)$ would coincide with the solution $V_\varepsilon(0,\cdot)$ of Lions and Lattes?

Assume $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_i \cdots$ to be the eigenvalues and $\{\varphi_i(\cdot)\}$ to be the respective complete system of orthonormal eigenfunctions in the first boundary-value problem for the operator $A = -\Delta$ in the domain Ω.

Assume

$$(N(\varepsilon)w)(\cdot) = \sum_{i=1}^{\infty} (1 - e^{\varepsilon\lambda_i^2\theta}) w_i\varphi_i(\cdot) \tag{9}$$

$$(K(\varepsilon)\sigma)(\cdot) = \sum_{i=1}^{\infty} e^{-(\varepsilon\lambda_i^2 - 2\lambda_i)\theta} \sigma_i\varphi_i(\cdot)$$

with w_i (respectively σ_i, z_i) being the Fourier coefficients for the expansion of functions $w(\cdot)$ (respectively $\sigma(\cdot)$, $z(\cdot)$) in a series along the system of functions $\{\varphi_i(\cdot)\}$.

Theorem 3.1. Assume $\alpha = 0$ and operators $N(\varepsilon)$, $M(\varepsilon)$, $K(\varepsilon)$ of inequality (6) to be defined as in (9) with $M(\varepsilon) = 0$. Then for all $\varepsilon > 0$ the center $w^0(\cdot)$ of the ellipsoid $W_\varepsilon(\cdot)$ (7) will coincide with the "Lions - Lattes" solution $w_\varepsilon(\cdot)$ (8). Namely

$$w_\varepsilon^0(\cdot) = w_\varepsilon(\cdot) = V_\varepsilon(0,\cdot)$$

and $w_\varepsilon^0(\cdot)$ will be represented as

$$w_\varepsilon^0(\cdot) = \sum_{i=1}^{\infty} e^{(-\varepsilon\lambda_i^2+\lambda_i)\theta} z_i\varphi_i(\cdot) .$$

The next theorem indicates that an appropriate selection of the operators $N(\varepsilon)$, $K(\varepsilon)$ in (6) (with $M(\varepsilon) = 0$) would allow to approximate the set

$$W^*(\cdot) = \{w^*(\cdot) | J(w^*(\cdot)) \le \gamma\}$$

with respective informational sets $W_\varepsilon(\cdot)$.

Theorem 3.2. Assume $\alpha = 0$, $\beta = 1$, $\varepsilon > 0$, $\nu > 0$ and the operators $N(\varepsilon)$, $M(\varepsilon)$, $K(\varepsilon)$ of inequality (6) to be defined as

$$(N(\varepsilon)w)(\cdot) = (N_{\varepsilon,\nu}w)(\cdot) =$$

$$\sum_{i=1}^{\infty} (e^{-2(1+\nu\lambda_i)^{-1}\lambda_i\theta} - e^{-(\varepsilon\lambda_i^2+2\lambda_i)\theta}) w_i\varphi_i(\cdot) ,$$

$$(K(\varepsilon)\sigma)(\cdot) = (K_\varepsilon\sigma)(\cdot) = \sum_{i=1}^{\infty} e^{-\varepsilon\lambda_i^2\theta} \sigma_i\varphi_i(\cdot) ,$$

$$M(\varepsilon) = 0 .$$

Then with $h_\varepsilon = 0$ there exists a pair $\varepsilon_0 > 0$, $\nu_0 > 0$ such that with $\varepsilon \le \varepsilon_0$, $\nu \le \nu_0$ the respective informational ellipsoidal set $W_\varepsilon(\cdot) = W_{\varepsilon,\nu}(\cdot) \ne \phi$. Its centers $w_{\varepsilon,\nu}^0$ converge:

$$\lim w_{\varepsilon,\nu}^0 = w_\varepsilon(\cdot) \quad (\nu \to 0)$$

and

$$\lim W_{\varepsilon,\nu}(\cdot) = W^*(\cdot) \quad (\varepsilon \to 0, \nu \to 0)$$

in the sense of Kuratowski [4].

4. Extremality and the General Regularization Scheme

Consider the minimization process for the functional (2). With $\alpha = 0$ a numerically stable scheme for calculating inf J is ensured by the quasiinvertibility method discussed above. We will now proceed with the construction of a respective algorithm for the general case, particularly for $\beta \geq 0$.

Theorem 4.1. The value

$$\inf_{w(\cdot)} J = \alpha \int_0^\theta \|y(t)\|^2\, dt + \beta\|z(\cdot)\|^2 - \sum_{i=1}^{\infty} v_i(\alpha p_i + \beta\, e^{-\lambda_i \theta}\, z_i)^2 ,$$

where

$$v_i = 2\lambda_i(\beta\, e^{-2\lambda_i\theta} + \alpha(1 - e^{-2\lambda_i\theta}))^{-1}$$

$y_i(t)$, p_i *are the Fourier coefficients for* $y(t,\cdot)$, $p(\cdot)$,

$$p(x) = \int_0^\theta u(t,x;y(t,x))\,dt$$

$$y(t) = \{y_1(t),\ldots,y_k(t),\ldots\} ,$$

is a sequence in ℓ_2. *The sequence*

$$w_\varepsilon(\cdot) = \sum_{i=1}^{\infty} v_i(\alpha\, e^{-\varepsilon\lambda_i}\, p_i + \beta\, e^{-(\varepsilon\lambda_i^2 + \lambda_i)\theta}\, z_i)\varphi_i(\cdot) \tag{10}$$

minimizes $J(w(\cdot))$ *with* $\varepsilon \to 0$.

Theorem 4.2. Suppose $\beta = 0$. *Then for* $w_\varepsilon(\cdot)$ *of* (10) *we will have*

$$w_\varepsilon(x) = 2(\Delta u(\varepsilon,x;p(\cdot)) + \sum_{k=1}^{\infty} (-1)^k \Delta u(2\theta k,x;p(\cdot)))$$

and consequently

$$J(w_\varepsilon(\cdot)) \to \inf_{w(\cdot)} J(w(\cdot)) \quad \textit{with} \quad \varepsilon \to 0 .$$

Remark 4.1. Once there exists a distribution $w(\cdot) \in L_2(\Omega)$ that ensures the equalities

$$y(t,x) \equiv u(t,x;w(\cdot))$$

$$z(x) \equiv u(\theta,x;w(\cdot))$$

the value

$$\inf_{w(\cdot)} J(w(\cdot)) = 0 \ .$$

The next question is whether the functions $w_\varepsilon(\cdot)$ of (10) could serve as centers of some "informational ellipsoids" $\mathbf{W}_\varepsilon$ that would correspond to an appropriate selection of operators $N(\varepsilon)$, $M(\varepsilon)$, $K(\varepsilon)$ in the restriction (6). The answer is affirmative and is given by the following theorem.

Theorem 4.3. Suppose the restriction (6) is defined through the operators

$$(M(\varepsilon)\xi)(t,x) = 2 \sum_{i=1}^{\infty} e^{-\varepsilon\lambda_i} \lambda_i (1 - e^{-2\lambda_i\theta})^{-1} \xi_i(t)\varphi_i(x) \tag{11}$$

with $N(\varepsilon)$, $K(\varepsilon)$ being the same as in (9). *Then the center $w_\varepsilon^0(\cdot)$ of the respective informational domain $\mathbf{W}_\varepsilon$ for equation* (1) *under restriction* (6), (9), (11) *will coincide with the distribution given by formula* (10): $w_\varepsilon^0(\cdot) = w_\varepsilon(\cdot)$.

Remark 4.2. Define a *minmax estimate* w^0 for a bounded convex set $\mathbf{W}$ as its *Chebyshev center*:

$$\sup\{\|w^0 - w\| \mid w \in \mathbf{W}\} = \min_{z \in \mathbf{W}} \sup \{\|z - w\| \mid w \in \mathbf{W}\} \ .$$

Then once $\mathbf{W}$ is an ellipsoid its Chebyshev center w^0 will coincide with its formal center. For an arbitrary bounded informational set that may appear in nonlinear nonconvex problems its Chebyshev center may be taken as a natural "guaranteed estimate" for the unknown parameter w.

5. Other Regularizing Procedures

Consider $\alpha = 0$. (a) Another regularizing procedure may be designed through the solution $v_\varepsilon(t,x)$ to the following problem:

$$\frac{\partial}{\partial t}(v_\varepsilon - \varepsilon \Delta v_\varepsilon) - \Delta v_\varepsilon = 0, \;\; 0 \le t \le \theta$$

$$v_\varepsilon \mid_{[0,\theta]\times S} = 0, \quad v_\varepsilon \mid_{t=\theta} = z(\cdot)$$

so that

$$w_\varepsilon(\cdot) = v_\varepsilon(0,\cdot) \; . \tag{12}$$

The system (12) was introduced in paper [5]. The function $w_\varepsilon(\cdot) = v_\varepsilon(0,\cdot)$ will be the center of the respective informational ellipsoid consistent with measurement $z(\cdot)$ if we assume

$$(N(\varepsilon)w)(\cdot) = \sum_{i=1}^{\infty} e^{-\lambda_i \theta}(1 - e^{-\varepsilon(1 + \varepsilon\lambda_i)^{-1}\lambda_i^2\theta})w_i\varphi_i(\cdot)$$

$$(K(\varepsilon)\sigma)(\cdot) = \sum_{i=1}^{\infty} e^{(1 + \varepsilon\lambda_i)^{-1}\lambda_i\theta}\sigma_i\varphi_i(\cdot) \; , \;\; M(\varepsilon) = 0 \; .$$

Here the center of the ellipsoid is defined in a formal way, through formula (7). The ellipsoid itself is however unbounded.

(b) With $z(\cdot)$ given, assume that there exists a solution to equation

$$U_\theta w(\cdot) = z(\cdot)$$

Consider the constraint (6) with

$$(N(\varepsilon)w)(\cdot) = n_\varepsilon w(\cdot), \;\; (K(\varepsilon)\sigma)(\cdot) = k_\varepsilon \sigma(\cdot) \; , \;\; M(\varepsilon) = 0$$

where $n_\varepsilon > 0$, $k_\varepsilon > 0$ are real numbers.

Then with $n_\varepsilon = \varepsilon^2$, $k_\varepsilon = 1$ the center $w_\varepsilon^0(\cdot)$ of the respective ellipsoid $W_\varepsilon(\cdot)$ will coincide with the quasisolution (in the sense of V.K. Ivanov [6]) to the equation

$$U_\theta w(\cdot) = z(\cdot) \; ,$$

on the set

$$M = \{w(\cdot) \mid \|w(\cdot)\| \le \|w_\varepsilon^0(\cdot)\|\} \text{ , i. e.}$$

$$w_\varepsilon^0(\cdot) = \arg\min \|U_\theta w(\cdot) - z(\cdot)\| \; , \;\; w(\cdot) \in M \; .$$

(c) Assuming $n_\varepsilon = 1$, $k_\varepsilon = \varepsilon^{-2}$ the function $w_\varepsilon^0(\cdot)$ will be an approximate solution to the equation

$$U_\theta w(\cdot) = z(\cdot)$$

by the "bias method" with bias

$$d(U_\theta w(\cdot), z(\cdot)) = J(w(\cdot)) \ .$$

So that

$$w_\varepsilon^0(\cdot) \text{ would solve the problem}$$

$$\min \{\|w(\cdot)\| : d(U_\theta w(\cdot),\ z(\cdot)) \le J(w_\varepsilon^0(\cdot))\}$$

In both cases (b), (c) we observe that $J(w_\varepsilon^0(\cdot)) \to 0$ with $\varepsilon \to 0$.

6. A Continuity Theorem

Taking the solution (10) present it as a linear maping

$$w_\varepsilon(\cdot) = F_\varepsilon(y(\cdot,\cdot), z(\cdot))$$

from $L_2([0,\theta]\times\Omega) \times L_2(\Omega)$ into $L_2(\Omega)$.

Suppose

$$y_\delta(t,x) = u(t,x;\ w^*(\cdot)) + \xi_\delta(t,x)$$

$$z_\delta(x) = u(\theta,x;\ w^*(\cdot)) + \sigma_\delta(x)$$

where

$$\|\xi_\delta(\cdot,\cdot)\| \le \delta_1, \quad \|\sigma_\delta(\cdot)\| \le \delta_2; \ \ \delta_1, \delta_2 > 0 \ .$$

Theorem 6.1. The mapping F_ε is uniformly continuous in $L_2([0,\theta]\times\Omega) \times L_2(\Omega)$. The following estimate is true

$$\|F_\varepsilon(y_\delta(\cdot,\cdot),\ z_\delta(\cdot)) - w^*(\cdot)\| \le (2\varepsilon^{-2}(4\delta_1^2 + \delta_2^2))^{1/2} \ .$$

With $\varepsilon \to 0$, $\delta_i \to 0$, $(\delta_i \varepsilon^{-1}) \to 0$, i=1,2, there is a strong convergence $F_\varepsilon(y_\delta(\cdot,\cdot),\ z_\delta(\cdot)) \to w^*(\cdot)$.

References

[1] Tikhonov A.N., Arsenin V.Ya. Methods of Solving Ill-posed Problems. Nauka, Moscow, 1986.

[2] Lions J.-L., Lattes R. Méthode de Quasi-Réversibilité et Applications. Dunod, Paris, 1967.

[3] Kurzhanski A.B. Control and Observation Under Uncertainty. Nauka, Moscow, 1977.

[4] Kuratowski R. Topology Vol. 1, 2. Academic Press, 1966.

[5] Gaewski H., Zacharias K. Zur Regularisierung einer Klasse nichtkorrekter Probleme bei Evolutionsgleichungen. J. Math. Anal. & Appl., V. 38, No. 3, 1972.

[6] Ivanov V.K., Vasin V.V., Tanana V.P. The Theory of Linear Ill-posed Problems and Its Applications. Nauka, Moscow, 1978.