UNITÉ DE RECHERCHE
INRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

1992
25ème anniversaire

N° 1729

Programme 6
Calcul Scientifique, Modélisation et
Logiciels numériques

CONVERGENCE OF TIKHONOV REGULARIZATION FOR CONSTRAINED ILL-POSED INVERSE PROBLEMS

Guy CHAVENT
Karl KUNISCH

Juillet 1992

RR-1729

Convergence of Tikhonov Regularization for Constrained Ill-posed Inverse Problems

G. Chavent Karl Kunisch[1]

May 1992

[1]Partially supported by Ministery of Foreign Affairs, France and the Fonds zur Förderung der wissenschaftlichen Forschung, Austria, under P-7869-PHY.

Convergence de la Régularisation de Tychonov pour les Problèmes Inverses mal posés avec Contrainte.

Convergence of Tikhonov Regularization for Constrained Ill-posed Inverse Problems.

Guy CHAVENT (*†) Karl KUNISCH(‡)

Résumé

On démontre des résultats de convergence et de taux de convergence pour des problèmes inverses mal posés formulés en terme de moindres carrés régularisés.

Abstract

In this paper convergence and rate of convergence results for nonlinear constrained ill-posed inverse problems formulated as regularized least squares problems are given.

Mots Clefs

Régularisation, problèmes inverses, moindres carrés non-linéaires, convergence.

Keywords

Regularization, inverse problems, non linear least squares, convergence.

(*) INRIA, Domaine de Voluceau-Rocquencourt, BP 105, 78153 Le Chesnay Cédex, France.
(†) CEREMADE, Université Paris Dauphine, 75775 Paris Cédex 16, France.
(‡) Technische Universität Graz, Institut für Mathematik, Kopernikusgasse 24, A-8010 Graz, Autriche.

Abstract

In this paper convergence and rate of convergence results for nonlinear constrained ill-posed inverse problems formulated as regularized least squares problems are given.

1 Introduction

The purpose of this note is to analyze the convergence of constrained regularized least squares problems as the regularization parameter and the error level converge to zero. Specifically, consider the problem

$$(\mathcal{P}) \qquad \min |\varphi(x) - z|^2 \text{ over } x \in C$$

where φ is a possibly nonlinear mapping, C is a closed convex set and z is fixed. Given an a-priori guess x_0 of the solution of $(\mathcal{P})$, we are interested in retrieving the solution $\hat{x}$ of $(\mathcal{P})$ which is closest to the a-priori guess x_0 ("x_0-minimum norm solution"). If the inverse of φ lacks favorable properties then one may choose a regularized least squares formulation to obtain approximate solutions to the above problem. Thus for small ϵ one introduces

$$(\mathcal{P}_\epsilon) \qquad \min |\varphi(x) - z_\epsilon|^2 + \epsilon^2 |x - x_0|^2 \text{ over } x \in C.$$

In the formulation of $(\mathcal{P}_\epsilon)$ we also allow the situation where the data z contain error. The error corrupted data are denoted by z_ϵ and they are assumed to satisfy $|z - z_\epsilon| \leq \delta$. The question of convergence of the solutions x_ϵ^δ of $(\mathcal{P}_\epsilon)$ as $\delta \to 0, \epsilon \to 0$ arises. The main results of this paper can be summarized as follows:

- Provided that a qualification hypothesis holds, the x_0-minimum norm solution $\hat{x}$ satisfies necessarily

  $$(\mathbf{N}) \qquad x_0 - \hat{x} \in \overline{\text{Range } \varphi'(\hat{x})^* + T(C, \hat{x})^-}.$$

  Here $\varphi'(\hat{x})^*$ denotes the adjoint of the linearization of φ at $\hat{x}$ and $T(C, \hat{x})^-$ is the polar cone of the tangent cone $T(C, \hat{x})$ to C at $\hat{x}$.

- If (N) is satisfied and φ is linear, then $x_\epsilon^\delta \to \hat{x}$ as soon as $\epsilon \to 0$ and $\frac{\delta}{\epsilon} \to 0$.

- If the regularity hypothesis

  $$(\mathbf{R}) \qquad x_0 - \hat{x} \in \text{Range } \varphi'(\hat{x})^* + T(C, \hat{x})^-$$

  is satisfied (with additional assumptions in the nonlinear case), then $\| x_\epsilon^\delta - \hat{x} \|_E = O(\sqrt{\delta})$ as soon as $\epsilon \to 0$ and $\delta \sim \epsilon^2$.

The proofs involve only "hard analysis", no weak convergence arguments are used.

Concerning related investigations we mention the books by Badeva and Morozov, and by Groetsch [BM,G], and the references given there. In [G] the rate of convergence question is treated for linear inverse problems using singular value decomposition. Rate of convergence results for general nonlinear inverse problems can be found in [EKN, N4]. In these investigations constraints are not considered explicity and the regularity condition (R) for the rate of convergence results mentioned above is $x_0 - \hat{x} \in \text{Range } \varphi'(\hat{x})^*$. - In [N1, N2] Neubauer has given a thorough analysis of Tikhonov regularization for constrained linear inverse problems. In fact the result on $O(\sqrt{\delta})$-convergence is also obtained in [N1] for linear inverse problems under a regularity condition which is formally different from (R). We shall show that the regularity conditions in [N1] and (R) are equivalent. Our proof is different from that in [N1] and it is given in such a way that the extension to nonlinear problems is rather straightforward.

The paper is organized as follows. In section 2 we derive the necessary condition (N) for the x_0-minimum norm solution $\hat{x}$ in the case where φ is a linear operator A. This justifies to call (R) a regularity condition. It will also be demonstrated that (R) is equivalent to the existence of a Lagrange multiplier for the constrained minimum norm problem defining $\hat{x}$. Sections 3 and 4 are devoted to convergence and rate of convergence analysis of linear and nonlinear problems respectively.

2 Preliminaries

Here we consider the linear case.

(i) **The unregularized problem.**
We shall use the following sets and mappings:

$$\begin{aligned} E, F &\cdots \text{ Hilbert spaces,} \\ C &\cdots \text{ closed convex subset of } E. \end{aligned}$$

$$A : E \to F \quad \cdots \quad \text{continuous linear operator,}$$
$$z \in F \quad \cdots \quad \text{error-free observation.}$$

It is assumed that z admits a projection $\hat{z}$ on $rg\ A(C)$, the range of A restricted to C. Due to convexity of C and linearity of A the projection $\hat{z}$ is unique. The unregularized problem is given by:

$$\textbf{(P)} \qquad \begin{cases} \text{find } \hat{x} \in C \text{ such that } J(x) = |A(x) - z|^2 \text{ is minimal} \\ \text{over } C \text{ at } \hat{x}. \end{cases}$$

This problem admits the following set of solutions:

$$X = \{x \in C : A(x) = \hat{z}\}.$$

It is a closed convex subset of E.

(ii) The regularized problem.
In addition to the above quantities let us denote by

$$x_0 \in E \quad \cdots \quad \text{a-priori guess,}$$
$$z_n \in F \quad \cdots \quad \text{sequence of measurements for } z \text{ with}$$
$$\delta_n = |z_n - z| \to 0,$$
$$\epsilon_n > 0 \quad \cdots \quad \text{sequence of regularization parameters with } \epsilon_n \to 0.$$

The n-th regularized problem is defined by:

$$\textbf{(P}_\textbf{n}\textbf{)} \qquad \begin{cases} \text{find } \hat{x}_n \in C \text{ such that} \\ J_n(x) = |Ax - z_n|_F^2 + \epsilon_n^2 |x - x_0|_E^2 \\ \text{is minimal over } C \text{ at } \hat{x}_n. \end{cases}$$

It is simple to establish:

Theorem 1. For every $n \in \mathbf{N}$ there exists a unique solution $\hat{x}_n$ of (P_n).

(iii) The minimum norm solution of the unregularized problem.
The choice of an a-priori guess $x_0 \in E$ implies the selection of a specific

solution $\hat{x}$ in the solution set X of the unregularized problem. It is the element closest to x_0 and is found by solving

$$\textbf{(Q)} \qquad \begin{cases} \text{find } \hat{x} \in X \text{ such that } |x - x_0|^2 \text{ is minimal over} \\ X \text{ at } \hat{x}. \end{cases}$$

Clearly (Q) has a unique solution which will be referred to as x_0-minimum norm solution. To exhibit some of its properties we shall utilize the following notions. For any convex set $K \subset E$ and $x \in \bar{K}$ the tangent cone to K at x is defined by

$$T(K, x) = \{y \in E : \exists x_n \in K, \lambda_n > 0 \text{ with } \lambda_n(x_n - x) \to y\},$$

and the negative polar cone $N(K, x)$ to K at x is given by

$$N(K, x) = T(K, x)^- = \{y \in E :< y, x >_E \leq 0 \text{ for all } x \in T(K, x)\}.$$

The tangent cone to a convex set and the negative polar cone are closed convex cones. The tangent cones to X and C satisfy:

Lemma 2 For any $x \in X$ we have

$$T(X, x) \subset \ker A \cap T(C, x).$$

Proof. Let $y \in T(X, x)$. Then there exist sequences $\{x_n\}$ in X and $\{\lambda_n\}$, with $\lambda_n > 0$, such that $y = \lim \lambda_n(x_n - x)$. It follows that $y \in \ker A$ and since $x_n \in X \subset C$ we also have that $y \in T(C, x)$.□

Concerning the converse to the inclusion in Lemma 2 we make the following

Definition. An element $x \in X$ is said to be qualified, if

$$T(X, x) = \ker A \cap T(C, x),$$

and X is called qualified if all of its elements are qualified.

The set X is obviously qualified if A is injective or if $C = E$. If

$$0 \in \text{ int } (C - \ker A) \tag{1}$$

then X is also qualified [A, pg 57]. Here int denotes the interior of the indicated set. To describe a situation when x is not qualified let us assume that C is a closed ball in E and $\ker A \cap C = \{x\}$. If $\ker A$ is not trivial, then x is not qualified. – For the case of box constraints we have the following result:

Proposition 3. Let C be defined by a finite number of linear constraints:

$$C = \{x \in E : M_i x \leq b_i, i \in I = \{1, \cdots, N_C\}\},$$

where M_i are bounded linear functionals on E and $b_i \in \mathbf{R}$. Then X is qualified.

Proof. Step 1. For any $\tilde{x} \in C$ we define

$$\begin{aligned} K &= \{y \in E : \exists x \in C, \lambda > 0 \text{ such that } y = \lambda(x - \tilde{x})\} \\ &= \bigcup_{\lambda > 0} \lambda(C - \tilde{x}), \end{aligned}$$

and prove that $T(C, \tilde{x}) = K$. By definition of $T(C, \tilde{x})$ we see that $T(C, \tilde{x}) = \bar{K}$ and hence it suffices to show that K is closed. Let $y_n \in K$ and $y \in E$ be such that $\lim y_n = y$. As $y_n \in K$ there exist $\lambda_n > 0$ and $x_n \in C$ such that $y_n = \lambda_n(x_n \tilde{x})$. Let us denote by $I(\tilde{x})$ the set of active indices at $\tilde{x}$: $I(\tilde{x}) = \{i \in I : M_i \tilde{x} = b_i\}$. Then we find $M_i y_n = \lambda_n(M_i x_n - M_i \tilde{x}) = \lambda_n(M_i x_n - b_i) \leq 0$ for all $i \in I(\tilde{x})$ and $n = 1, 2, \cdots$. Hence $M_i y \leq 0$ for all $i \in I(\tilde{x})$. Next we choose $\lambda > 0$ small enough so that $M_i \tilde{x} + \lambda M_i y \leq b_i$ for all $i \notin I(\tilde{x})$. It is simple to check that $\tilde{x} + \lambda y \in C$ and hence $y \in K$ and K is closed.

Step 2. We prove that $(\ker A \cap K) \subset T(X, \tilde{x})$. Let $y \in \ker A \cap K$ be given. Then $y = \lambda(x - \tilde{x})$, where $\lambda > 0$ and $x \in C$, and $Ay = 0$. Hence $Ax = A\tilde{x}$, so that $x \in X$ and $y = \lambda(x - \tilde{x}) \in T(X, \tilde{x})$. This ends the proof. □

The solution $\hat{x}$ of (Q) satisfies the following optimality condition.

Theorem 4. If the x_0-minimum-norm solution $\hat{x}$ is qualified then

$$x_0 - \hat{x} \in \overline{\text{rg } A^* + T(C, \hat{x})^-}. \tag{2}$$

Proof. The Euler condition for (Q) is given by

$$< x_0 - \hat{x}, x - \hat{x} > \leq 0 \text{ for all } x \in X.$$

It follows that

$$< x_0 - \hat{x}, y > \leq 0 \text{ for all } y \in T(X, \hat{x}),$$

and in trems of the negative polar cone

$$x_0 - \hat{x} \in T(X, \hat{x})^-.$$

The qualification hypothesis on $\hat{x}$ implies that

$$x_0 - \hat{x} \in (\ker A \cap T(C, \hat{x}))^-,$$

and, using a property of polar cones

$$x_0 - \hat{x} \in \overline{(\ker A)^- + T(C, \hat{x})^-}.$$

This further implies that

$$x_0 - \hat{x} \in \overline{\overline{\text{rg } A^*} + T(C, \hat{x})^-}.$$

Let now $\eta > 0$ be given arbitrarily. Then we can first find $x \in \overline{\text{rg } A^*} + T(C, \hat{x})^-$ such that

$$|x_0 - \hat{x} - x| \leq \frac{\eta}{2}.$$

Then $x = x_1 + x_2$ with $x_1 \in \overline{\text{rg } A^*}$ and $x_2 \in T(C, \hat{x})^-$, and we can find $x_3 \in \text{rg } A^*$ such that $|x_1 - x_3| \leq \frac{\eta}{2}$.

Hence we obtain

$$|x_0 - \hat{x} - (x_3 + x_2)| \geq |x_0 - \hat{x} - x| + |x - x_3 - x_2| \geq \eta,$$

which proves the announced result. □

It is a simple consequene of the poof of Theorem 4 that (2) is not only necessary but also sufficient for $\hat{x}$ to be the solution of (Q). We have

Corollary 5. If $\hat{x}$ satisfies (2) then it is the x_0-minimum norm solution of (P).

Under assumption (1) the conclusion of Theorem 4 can be sharpened:

Theorem 6. If $0 \in$ int $(C - \ker A)$, then the x_0-minimum-norm solution $\hat{x}$ satisfies

$$x_0 - \hat{x} \in \overline{\text{rg } A^*} + T(C, \hat{x})^-,$$

Proof. Under the assumption of the theorem we have

$$T(X, \hat{x})^- = \overline{\text{rg } A^*} + T(C, \hat{x})^-,$$

[A, pg. 57]. Referring back to the proof of Theorem 4 we obtain the desired result. □

(iv) The regularity condition.

In order the obtain convergence rates for the regularized problem as $\delta_n \to 0$ and $\epsilon_n \to 0$ we shall require the condition

$$x_0 - \hat{x} \in \text{ rg } A^* + T(C, \hat{x})^-, \tag{3}$$

which is referred to as regularity condition. We give now an interpretation of this condition in terms of the existence of a Lagrange multiplier. Condition (3) implies the existence of $\hat{w} \in F$ and $\hat{\mu} \in T(C, \hat{x})^-$ such that

$$x_0 - \hat{x} = A^* \hat{w} + \hat{\mu}.$$

It follows that

$$< x_0 - \hat{x}, x - \hat{x} >_E = < \hat{w}, A(x - \hat{x}) >_F + < \hat{\mu}, x - \hat{x} >_E \text{ for all } x \in C$$

and therfore

$$< \hat{x} - x_0, x - \hat{x} >_E + < \hat{w}, A(x - \hat{x}) >_F \geq 0 \text{ for all } x \in C.$$

This is equivalent to saying that the Lagrangian

$$\mathcal{L}(x, w) = \frac{1}{2}|x - x_0|_E^2 + < w, Ax - \hat{z} >_F$$

satisfies

$$\mathcal{L}(\hat{x}, \hat{w}) \leq \mathcal{L}(x, \hat{w}) \text{ for all } x \in C,$$

which means that $\hat{w} \in F$ is a Lagrange multiplier for the constraint $Ax = \hat{z}$. Hence we see that the regularity condition (3) is equivalent to the existence of a Lagrange multiplier $\hat{w} \in F$ for the problem of projecting the a-priori guess x_0 onto C under the constraint $Ax = \hat{z}$.

3 Convergence of regularized linear problems

Concerning convergence and rate of convergence of the solutions $\hat{x}_n$ of (P_n) we have the following result.

Theorem 7. (i) If the x_0-minimum norm solution $\hat{x}$ satisfies

(H1) $\quad x_0 - \hat{x} \in \overline{\text{rg } A^* + T(C, \hat{x})^-}$

(which is the case if $\hat{x}$ is qualified) and if $\epsilon_n \to 0, \frac{\delta_n}{\epsilon_n} \to 0$, then the solutions $\hat{x}_n$ of (P_n) converge strongly to $\hat{x}$.

(ii) If the regularity assumption

(H2) $\quad x_0 - \hat{x} \in \text{ rg } A^* + T(C, \hat{x})^-$

holds, then

$$|\hat{x}_n - \hat{x}|_E \leq \frac{\delta_n}{\epsilon_n} + \epsilon_n |\hat{w}|_F$$

and

$$|A\hat{x}_n - \hat{z}|_F \leq 2(\delta_n + \epsilon_n^2 |\hat{w}|_F),$$

where $\hat{w} \in F$ is any element satisfying

$$x_0 - \hat{x} = A^* \hat{w} + \hat{\mu} \text{ with } \hat{\mu} \in T(C, \hat{x})^-.$$

In particular, if $\epsilon_n \to 0$ and $\epsilon_n^2 \sim \delta_n$, then

$$|\hat{x}_n - \hat{x}|_E = O(\delta_n^{1/2}),$$

and

$$|A\hat{x}_n - \hat{z}|_F = O(\delta_n).$$

Proof. Step 1. We prove that (H1) implies the following a-priori estimate:

$$\begin{cases} \text{for all } \eta > 0 \text{ there exists } w \in F \text{ such that} \\ |A\hat{x}_n - \hat{z}|_F \leq \epsilon_n \eta + 2(\delta_n + \epsilon_n^2 |w|_F) \\ |\hat{x}_n - \hat{x}|_E \leq 2\eta + \frac{\delta_n}{\epsilon_n} + \epsilon_n |w|_F. \end{cases} \tag{4}$$

Let $\eta > 0$ be given. Then (H1) implies the existence of $w \in F$ and $\mu \in T(C, \hat{x})^-$ such that

$$|x_0 - \hat{x} - (A^* w + \mu)|_E \leq \eta.$$

Since $\hat{x}_n$ is a solution of (P_n) it follows that

$$J_n(\hat{x}_n) \leq J_n(x) \text{ for all } x \in C.$$

Choosing $x = \hat{x}$ yields:

$$|A\hat{x}_n - z_n|_F^2 + \epsilon_n^2 |\hat{x}_n - x_0|_E^2 \leq |A\hat{x} - z_n|_F^2 + \epsilon_n^2 |\hat{x} - x_0|_E^2.$$

Now we proceed similarly as in [EKN, Theorem 2.4]. Since the aim is to estimate $|A\hat{x}_n - \hat{z}|$ and $|\hat{x}_n - \hat{x}|$ we rewrite the above inequality as

$$\begin{aligned} |A\hat{x}_n - \hat{z}|_F^2 \quad &+ \quad \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 \leq |A\hat{x} - z_n|_F^2 + \epsilon_n^2 |\hat{x} - x_0|_E^2 \\ + |A\hat{x}_n - \hat{z}|_F^2 \quad &+ \quad \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 - |A\hat{x}_n - z_n|_F^2 - \epsilon_n^2 |\hat{x}_n - x_0|_E^2. \end{aligned}$$

Using the identity $a^2 + b^2 - (a+b)^2 = -2ab$ yields

$$|A\hat{x}_n - \hat{z}|_F^2 + \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 \leq 2 < A\hat{x}_n - \hat{z}, z_n - \hat{z} >_F + 2\epsilon_n^2 < \hat{x}_n - \hat{x}, x_0 - \hat{x} >_E .$$

This further yields

$$\begin{aligned} |A\hat{x}_n - \hat{z}|_F^2 \quad &+ \quad \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 \leq 2 < A\hat{x}_n - \hat{z}, z_n - z >_F + 2 < A\hat{x}_n - \hat{z}, z - \hat{z} >_F \\ &+ \quad 2\epsilon_n^2 < \hat{x}_n - \hat{x}, x_0 - \hat{x} - (A^* w + \mu) >_E + 2\epsilon_n^2 < \hat{x}_n - \hat{x}, A^* w >_E \\ &+ \quad 2\epsilon_n^2 < \hat{x}_n - \hat{x}, \mu >_E . \end{aligned}$$

Since by assumption $\hat{z}$ is the projection of z onto rg $A(C)$ and since $\mu \in T(C,\hat{x})^-$ the second and the last term on the right hand side of the last inequality are negative, and we obtain

$$\begin{aligned} |A\hat{x}_n - \hat{z}|_F^2 \;+\; & \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 \leq 2 < A\hat{x}_n - \hat{z}, z_n - z >_F \\ + \; & 2\epsilon_n^2 < \hat{x}_n - \hat{x}, x_0 - \hat{x} - (A^* w + \mu) >_E + 2\epsilon_n^2 < \hat{x}_n - \hat{x}, A^* w >_E . \end{aligned}$$

Rearranging terms yields

$$\begin{aligned} |A\hat{x}_n - \hat{z}|_F^2 \;+\; & \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 \leq 2 < A\hat{x}_n - \hat{z}, z_n - z + \epsilon_n^2 w >_F \\ + \; & 2\epsilon_n^2 < \hat{x}_n - \hat{x}, x_0 - \hat{x} - (A^* w + \mu) >_E . \end{aligned}$$

Since $|z_n - z|_F \leq \delta_n$ and $|x_0 - \hat{x} - (A^* w + \mu)| \leq \eta$ we obtain

$$\begin{aligned} |A\hat{x}_n - \hat{z}|_F^2 \;+\; & \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 \leq 2(\delta_n + \epsilon_n^2 |w|_F) |A\hat{x}_n - \hat{z}|_F \\ + \; & 2\epsilon_n^2 \eta |\hat{x}_n - \hat{x}|_E, \end{aligned}$$

which can be expressed as

$$a^2 + b^2 \leq 2\alpha a + 2\beta b,$$

with the obvious definition for a, b, α, β. This last inequality is equivalent to

$$(a - \alpha)^2 + (b - \beta)^2 \leq \alpha^2 + \beta^2,$$

which implies

$$\max(a - \alpha, b - \beta) \leq (\alpha^2 + \beta^2)^{1/2},$$

and since $(\alpha^2 + \beta^2)^{1/2} \leq \alpha + \beta$, we have

$$a \leq 2\alpha + \beta \text{ and } b \leq \alpha + 2\beta,$$

which gives the desired estimates (4).

Step 2. Let $\epsilon > 0$ and choose $\eta = \frac{\epsilon}{3}$. From Step 1 we know the existence of $w \in F$ such that

$$|\hat{x}_n - \hat{x}|_E \leq \frac{2\epsilon}{3} + \frac{\delta_n}{\epsilon_n} + \epsilon_n |w|_F.$$

Using the asymptotic assumptions on δ_n and ϵ_n this implies the existence of n_0 such that $|\hat{x}_n - \hat{x}|_E \leq \epsilon$ for all $n \geq n_0$, and therfore $\lim_n \hat{x}_n = \hat{x}$.

Step 3. The regularity assumption (H2) allows to take $\eta = 0$ in (4). Together with the assumptions on the asymptotic behavior of δ_n and ϵ_n, this implies part (ii) of the theorem. □

Remark 8. Concerning the first part of the theorem, we note that if $\hat{x}$ is qualified (e.g. if $C = E$, or if A is injective, or if C is defined by a finite number of linear constraints, or if $0 \in \text{int}\,(C - \ker A)$) then (H1) is satisfied due to Theorem 5 and strong convergence of $\hat{x}_n$ to $\hat{x}$ holds. Using weak subsequence arguments it can be shown [N2] that (H1) is in fact superfluous and strong convergence of x_n to $\hat{x}$ holds in any case. The advantage of (H1) are the explicit estimates (4).

Remark 9. The conclusion of the second part of Theorem 7 is also obtained in [N1] for the case $x_0 = 0$ under the condition that $\hat{x} \in \text{rg}\,(P_C A^*)$, where P_C denotes the metric projection in E onto C. The analogous condition in the case that a nontrivial a-priori guess x_0 is used is given by

(H2′) $\qquad \hat{x} \in P_C(\,\text{rg}\,A^* + \{x_0\})$.

We now argue that $(H2')$ and $(H2)$ are equivalent. In fact $(H2')$ is equivalent to the existence of $y \in \text{rg}\,A^*$ such that

$$< y + x_0 - \hat{x}, x - \hat{x} > \leq 0 \text{ for all } x \in C,$$

which in turn is equivalent to

$$< y + x_0 - \hat{x}, h > \leq 0 \text{ for all } h \in T(C, \hat{x}) = \overline{\bigcup_{\lambda > 0} \lambda(C - \{\hat{x}\})}.$$

The last fact can be expressed as

$$y + x_0 - \hat{x} \in T(C, \hat{x})^-$$

or equivalently as

$$x_0 - \hat{x} \in \text{ rg } A^* + T(C, \hat{x})^-,$$

which is (H2), as desired.

In the following result rate of convergence is obtained by assuming regularity of $x_0 - \hat{x}$ in terms of fractional powers of A^*A.

Theorem 10. If for some $\nu \in [0, \frac{1}{2}]$ the x_0-minimum norm solution $\hat{x}$ satisfies the regularity assumption

(H3) $\quad x_0 - \hat{x} \in \text{ rg } (A^*A)^\nu + T(C, \hat{x})^-,$

and $\epsilon = \epsilon(\delta)$ is choosen according to $c(\epsilon)\epsilon^{2\nu+1} \sim \delta$ where

$$c^2(\epsilon) = \int_0^\infty \frac{\epsilon^{2(1-2\nu)} t^{2\nu}}{t + \epsilon^2} dE_t |\hat{w}|^2,$$

$x_0 - \hat{x} = (A^*A)^\nu \hat{w} + \hat{\eta}$ for some $\hat{\eta} \in T(C, \hat{x})^-$ and E_t is the spectral family of A^*A, then

$$|\hat{x}_n - \hat{x}|_E = \begin{cases} o\left(\delta_n^{\frac{2\nu}{2\nu+1}}\right) \text{ for } \nu \in [0, \frac{1}{2}) \\ O\left(\delta_n^{1/2}\right) \text{ for } \nu = \frac{1}{2}, \end{cases}$$

as $\delta \to 0^+$.

Proof. Step 1 (Disguise of $c^2(\epsilon)$). Applying Hölder's inequality (with $p = \frac{1}{\nu}, q = \frac{1}{1-\nu}$) we find

$$\begin{aligned} c^2(\epsilon) &= \int_0^\infty \frac{\epsilon^{2(1-2\nu)}}{(t+\epsilon^2)^{1-2\nu}} \frac{t^{2\nu}}{(t+\epsilon^2)^{2\nu}} dE_t |\hat{w}|^2 \\ &\leq \left[\int_0^\infty \left(\frac{\epsilon^2}{t+\epsilon^2} \right)^{\frac{1-2\nu}{1-\nu}} dE_t |\hat{w}|^2 \right]^{1-\nu} \left[\int_0^\infty \left(\frac{t}{t+\epsilon^2} \right) dE_t |\hat{w}|^2 \right]^{\nu} \\ &\leq \left[\int_0^\infty \left(\frac{\epsilon^2}{t+\epsilon^2} \right)^{\frac{1-2\nu}{1-\nu}} dE_t |\hat{w}|^2 \right]^{1-\nu}, \end{aligned}$$

and hence $c(\epsilon) \to 0$ as $\epsilon \to 0$ if $\nu \in [0, \frac{1}{2})$ and $c(\epsilon)$ is bounded if $\nu = \frac{1}{2}$.

Step 2. The optimality condition for the convex problem (P_n) is given by

$$< (A^*A + \epsilon_n^2)\hat{x}_n - A^*z_n - \epsilon_n^2 x_0, c - \hat{x}_n >_E \geq \text{ for all } c \in C.$$

Introducing the inner product

$$< x, y >_n = < (A^*A + \epsilon_n^2)x, y >$$

on E the optimality condition can be equivalenty expressed by

$$\begin{cases} < \hat{x}_n - (A^*A + \epsilon_n^2)^{-1}(A^*z_n + \epsilon_n^2 x_0), c - \hat{x}_n >_n \geq 0 \text{ for all } x \in C \\ \hat{x}_n \in C. \end{cases}$$

Thus $\hat{x}_n = P_C^n((A^*A + \epsilon_n^2)^{-1}(A^*z_n + \epsilon_n^2 x_0))$, where P_C^n denotes the projection onto C with respect to the $< \cdot, \cdot >_n$ inner product. Since $A^*A\hat{x} = A^*\hat{z}$ we also find

$$< (A^*A + \epsilon_n^2)\hat{x} - \epsilon_n^2\hat{x} - A^*z - \epsilon_n^2\hat{\eta}, c - \hat{x} > \geq 0 \text{ for all } c \in C,$$

and equivalently

$$< \hat{x} - (A^*A + \epsilon_n^2)^{-1}(\epsilon_n^2\hat{x} + A^*z + \epsilon_n^2\hat{\eta}), c - \hat{x} >_n \geq 0 \text{ for all } c \in C,$$

and therefore

$$\hat{x} = P_C^n((A^*A + \epsilon_n^2)^{-1}(\epsilon_n^2\hat{x} + A^*z + \epsilon_n^2\hat{\eta})).$$

Since the projection onto a convex set is a contraction we obtain

$$|\hat{x}_n - \hat{x}|_n \leq |(A^*A + \epsilon_n^2)^{-1}A^*(z_n - z) + \epsilon_n^2(A^*A + \epsilon_n^2)^{-1}(x_0 - \hat{x} - \hat{\eta})|_n.$$

This further implies

$$\begin{aligned} \epsilon_n|\hat{x}_n - \hat{x}| &\leq |\hat{x}_n - \hat{x}|_n \leq |(A^*A + \epsilon_n^2)^{-1/2}A^*(z_n - z)| + \epsilon_n^2|(A^*A + \epsilon_n^2)^{-1/2}(A^*A)^\nu\hat{w}| \\ &= O(\delta_n) + \epsilon_n^{1+2\nu}\left[\int_0^\infty \frac{\epsilon^{2(1-2\nu)}t^{2\nu}}{t + \epsilon^2} dE_t|\hat{w}|^2\right]^{1/2} \\ &= O(\delta_n) + \epsilon_n^{1+2\nu}c(\epsilon), \end{aligned}$$

and therefore

$$|\hat{x}_n - \hat{x}| = O\Big(\frac{\delta_n}{\epsilon_n} + \epsilon_n^{2\nu} c(\epsilon)\Big) = O\Big(\delta^{\frac{2\nu}{1+2\nu}} \cdot c(\epsilon)^{\frac{1}{1+2\nu}}\Big).$$

In view of Step 1 this ends the proof. □

Remark 11. It can be seen from the last step of the proof that if ϵ is chosen according to $\epsilon(\delta) \sim \delta^{\frac{1}{2\nu+1}}$ then $|\hat{x}_n - \hat{x}|_E = O\Big(\delta_n^{\frac{2\nu}{2\nu+1}}\Big)$ for $\nu \in [0, \frac{1}{2}]$. In the unconstrained case the result analogous to Theorem 10 holds for all $\nu \in [0, 1]$ and the optimal rate $O(\delta^{2/3})$ is obtained for $\nu = 1$, see [G], for example. In the constrained case we failed to increase the rate beyond $O(\sqrt{\delta})$ by extending the range of ν beyond $\frac{1}{2}$. Some rather involved conditions for convergence of $\hat{x}_n$ to $\hat{x}$ like $O(\delta^{2/3})$ are given in [N1].

Remark 12. The regularity condition (H3) can be equivalently expressed as

$$(H3') \qquad \hat{x} \in P_C(\mathrm{rg}\ (A^*A)^\nu + \{x_0\}).$$

In the noise free case (i.e. $z_n = z$ for all n) and under additional conditions different from $(H3')$ the conclusion of Theorem 10 was also obtained in [N3].

In the unconstrained case the result analogous to Theorem 10 holds for all $\nu \in [0, 1]$ and the optimal rate is obtained for $\nu = 1$, see [G] for example.

4 Rate of convergence for regularized nonlinear problems

The following assumptions will be made throughout this section:
$\varphi : C \subset E \to F$,
φ is weakly sequentially closed and continuously Fréchet differentiable,
$z \in \varphi(C)$, i.e. $z = \hat{z}$ in the notation of section 2.

Here φ is called weakly sequentially closed if for any sequence $\{x_n\}$ in C, weak convergence of x_n to x in E and of $\varphi(x_n)$ to y in F imply $\varphi(x) = y$.

(i) Necessary optimality condition for x_0-minimum norm solution.

Due to the attainability assumption $z \in \varphi(C)$, the problem

(P) find $\hat{x} \in C$ such that $J(x) = |\varphi(x) - z|_F^2$ is minimal over C at $\hat{x}$

admits the solution set $X = \{x \in C : \varphi(x) = z\}$. The regularized problems are given by

$$(\mathbf{P_n}) \qquad \begin{cases} \text{find } \hat{x}_n \in C \text{ such that} \\ J_n(x) = |\varphi(x) - z_n|_F^2 + \epsilon_n^2 |x - x_0|_E^2 \\ \text{is minimal over } C \text{ at } \hat{x}_n, \end{cases}$$

where z_n, x_0 and ϵ_n are defined as in section 2. Weak closedness of φ implies existence of solutions $\hat{x}_n$ of (P_n). It will be shown that these solutions converge with rate $O(\sqrt{\delta_n})$ to an x_0-minimum norm solution of (P), i.e. to a solution of

(Q) find $\hat{x} \in X$ such that $|x - x_0|^2$ is minimal over X at $\hat{x}$,

provided $x_0 - \hat{x}$ satisfy an approriate regularity condition. Again weak sequential closedness of φ implies the existence of a solution to (Q) [EKN], but differently from the linear case, X is not convex and solutions to (Q) are not necessarily unique if φ is nonlinear. To state a necessary optimality condition for solutions to (Q) we define the sequential tangent cone to X at $x \in X$ as

$$T(X, x) = \{h \in E : \exists x_n \in X, \lambda_n > 0, \lambda_n \to 0, \text{ with } \lambda_n^{-1}(x_n - x) \to y\}.$$

In the case of convex X this cone coincides with the cone where the condition $\lambda_n \to 0$ is omitted. A short calculation using the fact that φ is Fréchet differentiable shows that

$$T(X, x) \subset \ker \varphi'(x) \cap T(C, x) \tag{5}$$

for any $x \in X$. Any x_0-minimum norm solution of (Q) clearly satisfies

$$x_0 - \hat{x} \in T(X, x)^-. \tag{6}$$

To further exploit this fact we again require a qualification condition, which generalizes the one given in the linear case. The element $\hat{x}$ is called qualified, if

$$T(X, \hat{x}) = \ker \varphi'(\hat{x}) \cap T(C, \hat{x}). \tag{7}$$

We next give a sufficient condition for $\hat{x}$ to be qualified in the case where

$$C = \{x \in E : g(x) \in K\},$$

and

$$\text{rg } (\varphi) \subset G \subset F,$$

where K is a closed convex cone in a real Banach space Y, g is a continuous mapping from E to Y, which is convex with respect to the ordering induced by K, and where G is a real Banach space. It is also assumed that g and $\varphi : C \to G$ are twice Fréchet differentiable at $\hat{x}$. The motivation for introducing G is given by the fact that the regular point condition below will require $\varphi'(\hat{x})$ to be surjective. If $\varphi'(\hat{x})$ was surjective onto F, however, then this would imply existence of a continuous (generalized) inverse of $\varphi'(\hat{x})$, an assumption which we clearly want to avoid.

Lemma 13. If in addition to the above assumptions on g and φ, $K + \mathbf{R}g(\hat{x})$ is closed and $\hat{x}$ is a regular point of $X = \{x \in E : \varphi(x) = z, g(x) \in K\}$, i.e.

$$\begin{pmatrix} \varphi'(\hat{x}) \\ g'(\hat{x}) \end{pmatrix} X - \begin{pmatrix} 0 \\ K + \mathbf{R}g(\hat{x}) \end{pmatrix} = \begin{pmatrix} G \\ Y \end{pmatrix},$$

then $\hat{x}$ is qualified.

Proof. In view of (5) we need to show that

$$\ker \varphi'(\hat{x}) \cap T(C, \hat{x}) \subset T(X, \hat{x}).$$

The regular point condition implies that

$$\ker \varphi'(\hat{x}) \cap L(C, \hat{x}) \subset T(X, \hat{x}),$$

where $L(C, \hat{x}) = \{x \in E : g'(\hat{x}) \in K + \mathbf{R}g(\hat{x})\}$ is the linearizing cone of C at $\hat{x}$ [MZ, Lemma 2.1]. Since $K + \mathbf{R}g(\hat{x})$ is assumed to be closed, the linearizing cone coincides with the tangent cone and we obtain

$$\ker \varphi'(\hat{x}) \cap L(C, \hat{x}) = \ker \varphi'(\hat{x}) \cap T(C, \hat{x}) \subset T(X, \hat{x}),$$

as desired.

Theorem 14. If an x_0-mininum norm solution $\hat{x}$ is qualified, then

$$x_0 - \hat{x} \in \overline{\text{rg } \varphi'(\hat{x})^* + T(C, \hat{x})^-}.$$

If, moreover, $0 \in$ int $(C - \ker \varphi'(\hat{x}))$, then

$$x_0 - \hat{x} \in \overline{\text{rg } \varphi'(\hat{x})^*} + T(C, \hat{x})^-.$$

Proof. Due to (6)

$$< x_0 - \hat{x}, h > \leq 0, \text{ for all } h \in T(X, \hat{x}).$$

The proof is now identical with that of Theorems 4 and 6 when A is replaced by $\varphi'(\hat{x})$.

(ii) Rate of convergence

Theorem 15. Let $\hat{x}$ be an x_0-minimum norm solution which satisfies

(H4)(i) $x_0 - \hat{x} \in$ rg $\varphi'(\hat{x})^* + T(C, \hat{x})^-$ (regularity condition)
(ii) there exists $L > 0$ such that $\| \varphi'(x) - \varphi'(\hat{x}) \| \leq L|x - \hat{x}|$ for all $x \in C$,
(iii) $L|\hat{w}|_F < 1$, where $\hat{w} \in F$ is an element satisfying

$$x_0 - \hat{x} = \varphi'(\hat{x})^* \hat{w} + \hat{\mu} \text{ with } \hat{\mu} \in T(C, \hat{x}),$$

and let $\{\hat{x}_n\}$ be any sequence of solutions to (P_n). Then

$$|\hat{x}_n - \hat{x}|_E \leq \frac{1}{\sqrt{1 - L|\hat{w}|_F}} \Big(\frac{\delta_n}{\epsilon_n} + \epsilon_n^2 |\hat{w}|_F \Big)$$

and

$$|\varphi(\hat{x}_n) - z|_F \leq 2(\delta_n + \epsilon_n^2 |\hat{w}|_F).$$

In particular, if $\epsilon_n \to 0$ and $\epsilon_n^2 \sim \delta_n$, then $|\hat{x}_n - \hat{x}|_E = O(\delta_n^{1/2})$ and $|\varphi(\hat{x}_n) - z|_F = O(\delta_n)$.

Proof. Due to (H4)(ii) there exist $r_n \in E$ such that

$$\varphi(\hat{x}_n) - \varphi(\hat{x}) = \varphi'(\hat{x})(\hat{x}_n - \hat{x}) + r_n. \tag{8}$$

and $|r_n|_F \leq \frac{1}{2} L |\hat{x}_n - \hat{x}|^2$. Following the argument in the proof to Theorem 7 we find

$$\begin{aligned} |\varphi(\hat{x}_n) - z|_F^2 \; &+ \; \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 \leq 2 < \varphi(\hat{x}_n) - z, z_n - z >_F \\ &+ \; 2\epsilon_n^2 < \hat{x}_n - \hat{x}, x_0 - \hat{x} >_E . \end{aligned}$$

By (H4)(i) this implies

$$\begin{aligned} |\varphi(\hat{x}_n) - z|_F^2 \; &+ \; \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 \leq 2 < \varphi(\hat{x}_n) - z, z_n - z >_F \\ &+ \; 2\epsilon_n^2 < \hat{x}_n - \hat{x}, \varphi'(\hat{x})^* \hat{w} + \hat{\mu} >_E \end{aligned}$$

with $\hat{w} \in F$ and $\hat{\mu} \in T(C, \hat{x})^-$. This further gives

$$\begin{aligned} |\varphi(\hat{x}_n) - z|_F^2 \; &+ \; \epsilon_n^2 |\hat{x}_n - \hat{x}|_E^2 \leq 2 < \varphi(\hat{x}_n) - z, z_n - z >_F \\ &+ \; 2\epsilon_n^2 < \varphi'(\hat{x})(\hat{x}_n - \hat{x}), \hat{w} >_F = 2 < \varphi(\hat{x}_n) - z, z_n - z + \epsilon_n^2 \hat{w} >_F \\ &+ \; \epsilon_n^2 L |\hat{w}|_F |\hat{x}_n - \hat{x}|_E^2, \end{aligned}$$

where we used (8). Rearranging terms we obtain

$$\begin{aligned} |\varphi(\hat{x}_n) - z|_F^2 \; &+ \; \epsilon_n^2 (1 - L|\hat{w}|_F) |\hat{x}_n - \hat{x}|_E^2 \\ &\leq \; 2(\delta_n + \epsilon_n^2 |\hat{w}|_F) |\varphi(\hat{x}_n) - z|_F, \end{aligned}$$

and the desired result follows.

Remark 16. Using weak subsequential arguments one can show that any sequence $\{\hat{x}_n\}$ of solutions of (P_n) contains a subsequence which converges strongly to an x_0-minimum norm solution of (P), provided that $\epsilon_n \to 0$ and $\frac{\delta_n}{\epsilon_n} \to 0$. We also observe that under the assumptions of Theorem 14, the x_0-minimum norm solution $\hat{x}$ is unique in C.

References

[A] J.P. Aubin: L'Analyse Non Linéaire et ses Motivations Economiques, Massons, Paris, 1984.

[BM] V. Badeva and V. Morozov: Problems Incorrect Poses, Théorie et Applications, Masson, Paris, 1991.

[EKN] H.W. Engl, K. Kunisch and A. Neubauer: Convergence rates for Tikhonov regularization of non-linear ill-posed problems, Inverse Problems, 5 (1989), 523-540.

[G] C.W. Groetsch: The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind, Pitman, Boston, 1984.

[MZ] H. Maurer and J. Zowe: First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems, Mathematical Programming 16(1979), 98-110.

[N1] A. Neubauer: Finite-dimensional approximation of constrained Tikhonov-regularized solutions of ill-posed linear operator equations, Math. Comp. 48(1987), 565-583.

[N2] A. Neubauer: Tikhonov-regularization of ill-posed linear operator equations on closed convex sets, J. Approx. Theory 53(1988), 304-320.

[N3] A. Neubauer: Tikhonov-regularization of ill-posed linear operator equations on closed convex sets, Verlag der wissenschaftlichen Gesellschaften Österreichs, Vienna, 1986.

[N4] A. Neubauer: Tikhonov regularization for non-linear ill-posed problems: optimal convergence rate and finite dimensional approximation, Inverse Problems 5(1989), 541-558.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique

ISSN 0249 - 6399