DEMONSTRATIO MATHEMATICA
Vol. XLIII No 4 2010

Zbigniew Bartoszewski

REMARKS ON THE CONVERGENCE OF AN ITERATIVE METHOD OF SOLUTION OF GENERALIZED LEAST SQUARES PROBLEM

Abstract. In this paper we consider an iterative method of finding a regularized solution of a general linear system $Ax = b$. For a given scalar α and an initial vector g it produces a sequence that converges to the least squares solution of this system. The limiting point minimizes the distance between g and the set of all least squares solutions of the problem. An estimate of the rate of convergence is also provided.

1. Introduction

Let $\mathbb{R}$ denote the set of real numbers. We denote the vector space of all $m \times n$ real matrices by $\mathbb{R}^{m\times n}$ and the space of all real n-vectors by $\mathbb{R}^n$. Consider the system of linear algebraic equations

$$Ax = b, \tag{1.1}$$

where $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ — a given $m \times n$ real matrix, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ — a given real m-vector and $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ — an unknown real n-vector. It is very well known that the system (1.1) may have exactly one solution or infinitely many solutions or no solutions. So, instead of seeking the solutions of (1.1) we can consider the more general problem:

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2, \tag{1.2}$$

2000 *Mathematics Subject Classification*: 65F20, 65F22, 65F10.
Key words and phrases: iterative methods, generalized inverse matrices, least squares problem, regularization of ill-posed problems.

where, for $u \in \mathbb{R}^n$

$$\|u\|_2 = (u_1^2 + \ldots + u_n^2)^{1/2}.$$

In contrast to problem (1.1), problem (1.2) always has a solution. It has more than one solution if the matrix A does not have the full column rank (for these well known facts see for instance [3]). It is clear that the problem (1.2) is equivalent to the problem

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2 . \tag{1.3}$$

To deal with a unique solution to (1.2), usually the solution with minimum norm is considered. But it is possible to consider a solution to (1.2) which is at the minimal distance to a given vector g instead. Although by a translation of the coordinate system we can reduce the latter problem to finding the minimizer to (1.2) with minimum norm we will show a direct and stable method of finding the minimizer to (1.2) at the closest distance to a given vector.

We denote by $\mathcal{X}$ the set of all minimizers x to (1.3):

$$\mathcal{X} = \{x \in \mathbb{R}^n : \|Ax - b\|_2^2 = \min\}. \tag{1.4}$$

Let $g \in \mathbb{R}^n$.

Definition 1.1. We will call the vector $x^g \in \mathcal{X}$ a g-pseudo-solution to (1.2) if it satisfies the relation

$$\|x^g - g\|_2 = \min_{x \in \mathcal{X}} \|x - g\|_2 . \tag{1.5}$$

It is obvious that condition (1.5) is equivalent to

$$\|x^g - g\|_2^2 = \min_{x \in \mathcal{X}} \|x - g\|_2^2 . \tag{1.6}$$

It is easy to prove the lemma

Lemma 1.1. *There exists exactly one g-pseudo-solution to problem* (1.1).

Proof. As in the case of the minimum norm solution, the statement follows from the convexity of $\mathcal{X}$ and strict convexity of $\|\cdot\|_2^2$. ∎

Let a matrix $A \in \mathbb{R}^{m \times n}$ and a vector $g \in \mathbb{R}^n$ be fixed. From Lemma 1.1 it follows that we have defined an operator $A^+(g) : \mathbb{R}^m \to \mathbb{R}^n$ which to each $b \in \mathbb{R}^m$ assigns a unique vector $x^g \in \mathbb{R}^n$. We refer to it as *the pseudo-inverse operator* to the operator A defined by the matrix A. It is easy to see that $A^+(\theta)$ (θ - zero in $\mathbb{R}^n$) is defined by the *Moore-Penrose* pseudo-inverse matrix A^+ of A and is linear. It is also easy to notice that $A^+(g)$ has the following property: for each $b \in \mathbb{R}^m$, $A^+(g)b = A^+(\theta)b + A^+(g)\theta$, where $A^+(\cdot)b$ is linear.

2. The regularized problem

Consider Tikhonov's regularization method, i.e. the regularized problem

$$\min_{x\in\mathbb{R}^n} \Phi_g^{(\alpha)}(x), \tag{2.1}$$

where

$$\Phi_g^{(\alpha)}(x) = \|Ax-b\|_2^2 + \alpha\,\|x-g\|_2^2\,, \quad x\in\mathbb{R}^n, \quad \alpha\in\mathbb{R}, \quad \alpha>0. \tag{2.2}$$

Definition 2.1. A vector x_α^g is referred to as an (α, g)-approximate solution to (1.1) if

$$\Phi_g^{(\alpha)}(x_\alpha^g) = \min_{x\in\mathbb{R}^n} \Phi_g^{(\alpha)}(x). \tag{2.3}$$

We have the following lemma.

Lemma 2.1. *For each $A\in\mathbb{R}^{m\times n}$, $b\in\mathbb{R}^m$, $g\in\mathbb{R}^n$, $\alpha\in\mathbb{R}$, $\alpha>0$ an (α, g)-approximate solution to* (1.1) *exists, is unique and belongs to the closed sphere $\bar{S}_{A,b}^{(g)} = \bar{S}\left(g, \|g-x^g\|_2\right)$.*

Proof. From (2.2) and the definition of x^g it follows immediately the relation

$$\min_{x\in\bar{S}_{A,b}^{(g)}} \Phi_g^{(\alpha)}(x) = \min_{x\in\mathbb{R}^n} \Phi_g^{(\alpha)}(x).$$

As $\bar{S}_{A,b}^{(g)}$ is compact in $\mathbb{R}^n$ with $\|\cdot\|_2$ norm then there exists at least one minimizer of $\Phi_g^{(\alpha)}$, say x_α^g, which by the construction belongs to $\bar{S}_{A,b}^{(g)}$. The uniqueness of the minimizer x_α^g follows from strict convexity of $\Phi_g^{(\alpha)}$. ∎

Let I denote the $n\times n$ unit matrix. We have the following theorem

Theorem 2.2. *An (α, g)-approximate solution to* (1.1) *satisfies the equation*

$$(A^TA+\alpha I)x = A^Tb+\alpha g, \tag{2.4}$$

which we refer to as the normal equations and is uniquely defined by this equation.

Proof. Assume $x\in\mathbb{R}^n$, $y\in\mathbb{R}^n$ and $a\in\mathbb{R}$. Consider the equality

$$\begin{aligned}\|A(x+ay)-b\|_2^2 + \alpha\,\|g-(x+ay)\|_2^2 &= \|Ax-b\|_2^2 + \alpha\,\|g-x\|_2^2 \\ &+ 2ay^T\left[A^T(Ax-b)-\alpha(g-x)\right] + a^2\left(\|Ay\|_2^2+\alpha\,\|y\|_2^2\right).\end{aligned}$$

If x is the (α, g)-approximate solution to (1.1) then we must have

$$A^T(Ax-b)-\alpha(g-x) = 0. \tag{2.5}$$

Otherwise, if $y = -\left[A^T(Ax-b) - \alpha(g-x)\right]$ and $\alpha > 0$ is small enough, then we obtain the contradictory inequality

$$\|A(x+ay)-b\|_2^2 + \alpha\,\|g-(x+ay)\|_2^2 < \|Ax-b\|_2^2 + \alpha\,\|g-x\|_2^2 .$$

It is clear that (2.5) can be rewritten in the form (2.4).

As the matrix A^TA is symmetric and positive semidefinite and $\alpha > 0$ then $A^TA + \alpha I$ is nonsingular and the (α, g)-approximate solution to (1.1) is the unique solution to eq. (1.1) given by the formula

$$x_\alpha^g = (A^TA+\alpha I)^{-1}A^Tb + \alpha(A^TA+\alpha I)^{-1}g. \tag{2.6}$$

In this way we completed the proof of the theorem. ∎

We also have the following theorem about the convergence of the sequence of (α, g)-approximate solutions to the g-pseudo-solution to (1.1).

Theorem 2.3. *For any $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $g \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}$ satisfying $\alpha > 0$*

$$\lim_{\alpha\to 0} x_\alpha^g \longrightarrow x^g. \tag{2.7}$$

Proof. It follows from Lemma 2.1 that x_α^g belongs to the compact set $\bar{S}_{A,b}^{(g)}$. So, from each sequence x_α^g, $\alpha \to 0^+$ we can choose a convergent, say to $\bar{x}^g \in \bar{S}_{A,b}^{(g)}$, subsequence $x_{\alpha_\nu}^g$. We have the inequality

$$\Phi_g^{(\alpha_\nu)}(x_{\alpha_\nu}^g) \le \Phi_g^{(\alpha)}(x^g). \tag{2.8}$$

Taking the limit as $\alpha_\nu \to 0$ in both sides of inequality (2.8) we arrive at the equality

$$\|A\bar{x}^g - b\|_2^2 = \|Ax^g - b\|_2^2 \tag{2.9}$$

and from the uniqueness of g-pseudo-solution to (1.1) it follows that $\bar{x}^g = x^g$. In this way we proved that each sequence x_α^g is convergent to x^g as α approaches zero. ∎

3. An iterative method of computing of g-pseudo-solution

The problem of computing of g-pseudo-solution may cause many difficulties not only if the matrix A is rank deficient but also when A has a full rank but is nearly rank deficient. Many methods have been developed for solving the least squares problem and they are discussed in detail in literature (see, for example [2, 3] and references therein). Those methods can be also applied to determine the g-pseudo-solution. There is also a vast literature devoted to solving regularized problems where different methods of choosing the regularization parameter α are discussed (see, for instance, [5, 6] and references therein). If an (α, g)-approximate solution is to approximate the g-pseudo-solution to (1.1) the parameter α should be small and if A is rank deficient or nearly rank deficient then the problem of determining of x_α^g is ill-conditioned and we can expect troubles. In this section we propose an iterative method that produces a sequence convergent to the g-pseudo-solution to (1.1) for any $\alpha > 0$ and show how the rate of convergence of this sequence to x^g depends on the value of the parameter α. First, we use the singular value decomposition (SVD) to express the (α, g)-approximate solution and x^g-pseudo-solution in terms related to the matrices involved in SVD of A. It is well known (see, for example, [3]) that an arbitrary matrix $A \in \mathbb{R}^{m\times n}$ can be expressed in the form

$$A = U\Sigma V^T, \tag{3.1}$$

where $U = [u_1, \ldots, u_m] \in \mathbb{R}^{m\times m}$ and $V = [v_1, \ldots, v_n] \in \mathbb{R}^{n\times n}$ are orthogonal and $\Sigma = \operatorname{diag}[\sigma_1, \ldots, \sigma_p] \in \mathbb{R}^{m\times n}$, $p = \min\{m, n\}$.

Then $A^TA = V\Sigma^T\Sigma V^T$, $\left(A^TA + \alpha I\right)^{-1} = V\left(\Sigma^T\Sigma + \alpha I\right)^{-1} V^T$ and equality (2.6) can be rewritten in the form

$$x_\alpha^g = V\left(\Sigma^T\Sigma + \alpha I\right)^{-1} V^T V\Sigma^T U^T b + \alpha V\left(\Sigma^T\Sigma + \alpha I\right)^{-1} V^T g. \tag{3.2}$$

Assume, without loss of generality, that only the first r, $r \le p$, singular values σ_i are different from zero. Introducing the following notation $\beta = (\beta_1, \ldots, \beta_m)^T = U^Tb$, $w = (\beta_1\sigma_1, \ldots, \beta_r\sigma_r, 0, \ldots, 0)^T$, $w \in \mathbb{R}^n$, $z = ((v_1, g), \ldots, (v_n, g))^T$ the relation (3.2) can be written as

$$\begin{aligned} x_\alpha^g &= \\ &V\left(\Sigma^T\Sigma + \alpha I\right)^{-1} V^T(\beta_1\sigma_1 v_1 + \cdots + \beta_r\sigma_r v_r) + \alpha V\left(\Sigma^T\Sigma + \alpha I\right)^{-1} V^T g \end{aligned} \tag{3.3}$$

or in a shorter form

$$x_\alpha^g = V\left(\Sigma^T\Sigma + \alpha I\right)^{-1} w + \alpha V\left(\Sigma^T\Sigma + \alpha I\right)^{-1} z. \tag{3.4}$$

The matrix $\left(\Sigma^T\Sigma + \alpha I\right)^{-1}$ has the form

$$\left(\Sigma^T\Sigma + \alpha I\right)^{-1} = \operatorname{diag}\left[\frac{1}{\alpha + \sigma_1^2}, \ldots, \frac{1}{\alpha + \sigma_r^2}, \frac{1}{\alpha}, \ldots, \frac{1}{\alpha}\right]$$

and therefore the equality (3.4) can be rewritten as

$$x_\alpha^g = V\begin{bmatrix} \frac{\beta_1\sigma_1}{\alpha+\sigma_1^2} \\ \vdots \\ \frac{\beta_r\sigma_r}{\alpha+\sigma_r^2} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + V\begin{bmatrix} \frac{\alpha(v_1,g)}{\alpha+\sigma_1^2} \\ \vdots \\ \frac{\alpha(v_r,g)}{\alpha+\sigma_r^2} \\ (v_{r+1}, g) \\ \vdots \\ (v_n, g) \end{bmatrix}. \tag{3.5}$$

Taking in (3.5) the limit as $\alpha \to 0$, we obtain

$$x^g = \lim_{\alpha\to 0} x^g_\alpha = \frac{\beta_1}{\sigma_1}v_1 + \ldots + \frac{\beta_r}{\sigma_r}v_r + (v_{r+1}, g)v_{r+1} + \ldots + (v_n, g)v_n \tag{3.6}$$

Now, for $k = 0, 1, \ldots$, we define the following sequence

$$\begin{aligned} x_{k+1} &= (A^TA + \alpha I)^{-1}A^Tb + \alpha(A^TA + \alpha I)^{-1}x_k \\ x_0 &= g. \end{aligned} \tag{3.7}$$

We have the following theorem.

Theorem 3.1. *For any $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $g \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}$ satisfying $\alpha > 0$, the sequence (x_k) defined by* (3.7) *is convergent to the g-pseudo-solution of* (1.1), *i.e.*

$$\lim_{k\to\infty} x_k = x^g. \tag{3.8}$$

Proof. Let us introduce the following notation:

$$Q = (A^TA + \alpha I)^{-1}, \quad S = \alpha Q, \quad h = A^Tb.$$

Then, the first of equalities (3.7) can be written as

$$x_{k+1} = Sx_k + Qh, \quad k = 0, 1, \ldots$$

and implies the relation

$$x_{k+1} = S^{k+1}g + (S^k + \ldots + I)Qh, \quad k = 0, 1, \ldots. \tag{3.9}$$

Equality (3.9) can be written in the form

$$\begin{aligned} x_{k+1} = {} & V\mathrm{diag}\left[\left(\frac{\alpha}{\alpha+\sigma_1^2}\right)^{k+1}, \ldots, \left(\frac{\alpha}{\alpha+\sigma_r^2}\right)^{k+1}, 1, \ldots, 1\right]V^Tg \\ & + V\mathrm{diag}\left[\frac{\alpha+\sigma_1^2}{\sigma_1^2}\left(1 - \left(\frac{\alpha}{\alpha+\sigma_1^2}\right)^k\right), \ldots, \right. \\ & \left. \frac{\alpha+\sigma_r^2}{\sigma_r^2}\left(1 - \left(\frac{\alpha}{\alpha+\sigma_r^2}\right)^k\right), k+1, \ldots, k+1\right]V^TQh. \end{aligned} \tag{3.10}$$

Because of the equality

$$V^TQh = \left(\frac{\sigma_1\beta_1}{\alpha+\sigma_1^2}, \ldots, \frac{\sigma_r\beta_r}{\alpha+\sigma_r^2}, 0, \ldots, 0\right)^T \tag{3.11}$$

the relation (3.10) can be written in the form

$$\begin{aligned} x_{k+1} = {} & V\mathrm{diag}\left[\left(\frac{\alpha}{\alpha+\sigma_1^2}\right)^{k+1},\ldots,\left(\frac{\alpha}{\alpha+\sigma_r^2}\right)^{k+1},1,\ldots,1\right]V^Tg \\ & +V\left(\frac{\beta_1}{\sigma_1}\left(1-\left(\frac{\alpha}{\alpha+\sigma_1^2}\right)^k\right),\ldots,\right. \\ & \left.\frac{\beta_r}{\sigma_r}\left(1-\left(\frac{\alpha}{\alpha+\sigma_r^2}\right)^k\right),0,\ldots,0\right)^T. \end{aligned} \tag{3.12}$$

The equality (3.12) implies

$$\begin{aligned} \lim_{k\to\infty} x_k &= V\begin{bmatrix} 0 \\ \vdots \\ 0 \\ (v_{r+1},g) \\ \vdots \\ (v_n,g) \end{bmatrix} + V\begin{bmatrix} \frac{\beta_1}{\sigma_1} \\ \vdots \\ \frac{\beta_r}{\sigma_r} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = V\begin{bmatrix} \frac{\beta_1}{\sigma_1} \\ \vdots \\ \frac{\beta_r}{\sigma_r} \\ (v_{r+1},g) \\ \vdots \\ (v_n,g) \end{bmatrix} \\ &= \frac{\beta_1}{\sigma_1}v_1+\ldots+\frac{\beta_r}{\sigma_r}v_r+(v_{r+1},g)v_{r+1}+\ldots+(v_n,g)v_n, \end{aligned} \tag{3.13}$$

i.e.

$$\lim_{k\to\infty} x_k = x^g, \tag{3.14}$$

which completes the proof of the theorem. ∎

Now, we use the relation (3.12) to prove the estimate of the rate of convergence of x_k to the g-pseudo-solution x^g. Namely, we have the following theorem.

Theorem 3.2. *For the sequence of iteration x_k, $k=0,1,\ldots$, the following estimate of rate of convergence*

$$\|x_{k+1}-x^g\|_2 \le \left(\frac{\alpha}{\alpha+\sigma_j^2}\right)^k\left(\frac{\alpha}{\alpha+\sigma_j^2}\|g\|_2+\sqrt{r}\max_{1\le i\le r}\left|\frac{\beta_i}{\sigma_i}\right|\right) \tag{3.15}$$

holds with the index j defined by the relation $\sigma_j = \min_{1\le i\le r}\sigma_i$.

Proof. It follows from (3.12) that

$$
\begin{aligned}
x_{k+1} - x^g = {} & V\operatorname{diag}\left[\left(\frac{\alpha}{\alpha+\sigma_1^2}\right)^{k+1}, \dots, \left(\frac{\alpha}{\alpha+\sigma_r^2}\right)^{k+1}, 1, \dots, 1\right] V^T g \\
& + V\left(\frac{\beta_1}{\sigma_1}\left(1-\left(\frac{\alpha}{\alpha+\sigma_1^2}\right)^k\right), \dots, \right. \\
& \left. \frac{\beta_r}{\sigma_r}\left(1-\left(\frac{\alpha}{\alpha+\sigma_r^2}\right)^k\right), 0, \dots, 0\right)^T \\
& - V\operatorname{diag}[0, \dots, 0, 1 \dots, 1] V^T g - V\left(\frac{\beta_1}{\sigma_1}, \dots, \frac{\beta_r}{\sigma_r}, 0, \dots, 0\right)^T \\
= {} & V\operatorname{diag}\left[\left(\frac{\alpha}{\alpha+\sigma_1^2}\right)^{k+1}, \dots, \left(\frac{\alpha}{\alpha+\sigma_r^2}\right)^{k+1}, 0, \dots, 0\right] V^T g \\
& - V\left(\frac{\beta_1}{\sigma_1}\left(\frac{\alpha}{\alpha+\sigma_1^2}\right)^k, \dots, \frac{\beta_r}{\sigma_r}\left(\frac{\alpha}{\alpha+\sigma_r^2}\right)^k, 0, \dots, 0\right)^T .
\end{aligned}
\tag{3.16}
$$

The relation (3.16) implies the estimate (3.15). ∎

References

[1] Z. Bartoszewski, *A correction and discretization of the general linear boundary value problem for nonlinear systems of ordinary differential equations in normal form*, Ph. D. Thesis, Lomonosov University, Moscow, 1982 (in Russian).

[2] Å. Björk, *Numerical Methods for Least Squers Problems*, SIAM, Philadelphia, 1996.

[3] G. H. Golub, C. F. Van Loan, *Matrix Computations*, The John Hopkins University Press, Baltimore, 1989.

[4] A. D. Gorbunov, *Methods of Solving of Nonlinear Problems*, Lomonosov University, Moscow, 1980 (in Russian).

[5] V. A. Morozov, *Regularization Methods for Ill-posed Problems*, CRC Press, Boca Raton, FL, 1993.

[6] V. A. Morozov, *On the rate of convergence for regularized solutions*, Numer. Funct. Anal. Optim. 19, 3-4 (1998), 345–352.

THE FACULTY OF APPLIED PHYSICS AND MATHEMATICS
GDAŃSK UNIVERSITY OF TECHNOLOGY
ul. G. Narutowicza 11/12
80-233 GDAŃSK, POLAND
E-mail: zbart@mif.pg.gda.pl

Received May 23, 2009; revised version September 20, 2009.