규정된 선형 연산자 방정식에 대한 Conjugate gradient 방법에 관하여

金 道 鉉

ON THE CONJUGATE GRADIENT METHOD FOR CONSTRAINED SINGULAR LINEAR OPERATOR EQUATIONS

Do-Hyun Kim

1. Introuduction

In this paper we introduce the weighted generalized inverse of a linear operator in a general Hilbert space, and we establish the convergence of the conjugate gradient method to the least squares solution of minimal norm.

Throughout this paper, we shall let X, Y, and Z be (real or complex) Hibert spaces, and let A be a bounded linear operator on X into Y. The linear equation

(1) $Ax = y$ for $y \in Y$

may or may not have a solution depending on whether or not y is in $R(A)$, the range of A, and even if $y \in R(A)$ the solution of (1) need not be unique. In either case, one can seek a least squares solution i. e., a solution which minimizes the quadratic functional $f(x) = \| Ax-y \|^2$. Such a solution exists for all $y \in R(A) \oplus R(A)^{\perp}$. We shall also be interested in the least squares solution of minimial norm.

We consider the conjugate gradient method that minimizes $f(x)$ at each step. That is, choose an initial vector $x_0 \in X$, then compute $r_0 = P_0 = A^*(Ax_0-y)$, where A^* is the adjoint of A. If $p_0 \neq 0$, compute $x_1 = x_0 - \alpha_0 p_0$ where $\alpha_0 = \| r_0 \|^2 / \| Ap_0 \|^2$. For i = 1, 2, ⋯, compute

(2) $r_i = A^*(Ax_i-y) = r_{i-1} - \alpha_{i-1} A^* Ap_{i-1}$,

where

(3) $\alpha_{i-1} = \dfrac{\langle r_{n-1}, p_{n-1} \rangle}{\| Ap_{i-1} \|^2}$,

and if $r_i \neq 0$, then compute

(4) $p_i = r_i + \beta_{i-1} p_{i-1}$, where $\beta_{i-1} = - \dfrac{\langle r_i, A^* Ap_{i-1} \rangle}{\| Ap_{i-1} \|^2}$,

and set

(5) $x_{i+1} = x_i - \alpha_i p_i$

2. Least squares solutions and weighted generalized inverse

For any subspace S, we denote the orthogonal complement of S by $S^{\perp}$ and the closure of S by $\bar{S}$. Let D(A), R(A), and N(A) denote, respectively, the domain, the range and the null space of a linear operator A. The restriction of A to a set K is denoted by L|K. It is well Known [1]

$$X = N(A) \oplus N(A)^{\perp}$$
$$Y = N(A^{*}) \oplus N(A^{*})^{\perp}$$
$$\{R(T)\}^{\perp} = N(A^{*}),\ R(A^{*}) = N(A)^{\perp},\ R(A) = N(A^{*})^{\perp}.$$

For a given $y \epsilon Y$, an element $u \epsilon X$ is called a least squares solution of the linear opeator equation $Ax=y$ if $\| Au-y \| \leqq \| Ax-y \|$ for all $x \epsilon X$. Among least squares solutions an element v of minimal norm is called a best approximate solution of (1). For each $y \epsilon R(A) \oplus R(A)^{\perp}$, the set of all least squares solutions of (1) is a nonempty closed convex subset of X and hence has a unique element v of minimal norm. The generalized inverse A^{+} of A is the operator whose domatin is $D(A^{+}) = R(A) \oplus R(A)^{\perp}$ and $A^{+}y = v$, where v is the unique best approximate solution of the equation (1). If R(A) is not closed, then A^{+} is only densely defined and unbounded. If u is a least squares solution of (1), then $u = A^{+}y + (I - A^{+}A)x_0$ for some $x_0 \epsilon X$.

Let L be a bounded linear operator from X into Z. We assume that the range R(L) of L is closed in Z, but the range R(A) of A is not necessarily closed in Y. For a given y in $D(A^{+})$, let

(6) $S_y = \{u \epsilon X :\ \| Au-y \|_Y = \inf \| Ax-y \|_Y,\ x \epsilon X\}$.

Then the problem is to find $w \epsilon S_y$ such that

(7) $\| Lw \|_Z = \inf\{ \| Lu \|_Z : u \epsilon S_y\}$.

The problem (6)-(7) has a solution for every $y \epsilon D(A^{+})$ if and only if $N(A)+N(L)$ is closed, and the solution is unique if and only if $N(A) \cap N(L) = \{0\}$. Throughout this paper, we assume that $N(A) \cap N(L) = \{0\}$ and $N(A)+N(L)$ is closed.

We define a new inner product in X :

$$[u, v] = \langle Au, Av \rangle_Y + \langle Lu, Lv \rangle_Z \text{ for } u, v \epsilon X.$$

We denote the space X with the inner product $[\cdot, \cdot]$ by X_L.

THEOREM 1. An element $w \epsilon X$ is a solution to the problem (6)-(7) if and only if $A^{*}Aw = A^{*}y$ and $L^{*}Lw \epsilon N(A)$.

Proof. Refer to Nashed[2].

The solution w is the least squares solution of X_L-minimal norm of the equation (1). Let A^{+}_{L} denote the map induced by $y \rightarrow w$ and call it the weighted generalized inverse of A.

3. Regularization and some observations on the conjugate gradient method

When the range of A is closed, the problem (6)-(7) is well-posed. Hence our interest is in the case that the range of A is not closed and hence the problem is ill-posed. Instead of solving this ill-posed problem directly, we will regularize it by a family of stable minimization problems.

Let W be the product space of Y and Z with the usual inner product:

$$W = Y \times Z,$$

$$\langle (y_1, z_1), (y_2, z_2) \rangle_W = \langle y_1, y_2 \rangle_Y + \langle z_1, z_2 \rangle_Z$$

for $y_1, y_2 \in Y$ and $z_1, z_2 \in Z$.

For $\alpha > 0$, let C_α be a linear operator from X into W defined by $C_\alpha x = (Ax, \sqrt{\alpha} Lx)$ for $x \in X$. We denote by U_α the unique best approximate solution of the equation $C_\alpha x = \bar{y}$ for each $\alpha > 0$ where $\bar{y} = (y, 0)$ in W. That is, $U_\alpha = C_\alpha^+ \bar{y}$. Let us write $J_\alpha(x) = \| Ax - y \|^2 + \alpha \| Lx \|^2$.

THEOREM 2. Let $\alpha > 0$. An element X_α in X minimizes the quadratic functional $J_\alpha(x)$ if and only if $C_\alpha^* C_\alpha x = C_\alpha^* y$.

Proof. It is easy and omitted.

$\| X \|$ and $\| X \|_L$ are equivalent if AN(L) is closed. Throughout this paper we assume that AN(L) is closed.

THEOREM 3. For $\alpha > 0$, let U_α be the unique solution of the operator equation (8). Then $\lim_{\alpha \to 0} U_\alpha$ exists and $\lim_{\alpha \to 0} U_\alpha = A^+_L y$.

Proof. Refer to Song [3].

We now examine some properties of the conjugate gradient algorithm described in the introduction. Let P denote the orthogonal projection of X onto $\overline{R(C_\alpha^*)}$ and let Q denote the orthogonal projection of W onto $\overline{R(C_\alpha)}$. If $\bar{y} \in D(C_\alpha^+)$, then $Q\bar{y} = \hat{y} \in R(C_\alpha)$ and $v = C_\alpha^+ \bar{y} = C_\alpha^+ \hat{y}$ and $\hat{y} = Q\bar{y} = C_\alpha v = C_\alpha C_\alpha^+ \bar{y}$. Since Q is an orthogonal projection, the functional $J_\alpha(x)$ can be written as $J_\alpha(x) = \| C_\alpha x - y \|^2 = \| C_\alpha x - \hat{y} \|^2 + \| \hat{y} - y \|^2$. Thus, minimizing $J_\alpha(x)$ is equivalent to minimizing the functional $\| C_\alpha x - \hat{y} \|^2$ which we shall denote by $K_\alpha(X)$.

Setting $u = v + (I-P)x_0 = C_\alpha^+ y + (I-P)x_0$. One can define the error vector $e = x - u$ and the vector $r = C_\alpha^*(C_\alpha x - \hat{y}) = C_\alpha^*(C_\alpha x - \hat{y})$

Then $(C_\alpha^* C_\alpha)e = r$ and $[r, e] = \| C_\alpha x - \hat{y} \|^2 = K_\alpha(x)$.

The sequence of iterates $\{x_i\}$ generated by the conjugate gradient method (2)-(5) is contained in $x_0 + R(C_\alpha^*)$ with both r_i and p_i, for $i = 0, 1, 2, \cdots$, in $R(C_\alpha^*)$. Unless explicitly mentioned otherwise, we shall assume that the conjuate gradient method does not terminate in a finite number of steps, that is $r_i \neq 0$ for $i = 0, 1, 2 \cdots\cdots$.

We shall make use of the following lemmas.

Lemma 4. (a) For k=0, 1, 2, ⋯, i $K\alpha(x_i)=[r_i, e_k]=[e_i, r_k]$

(b) For i=0, 1, 2, ⋯⋯, $[p_i, e_i]$ $\|r_i\|^2=K\alpha(x_i)\|p_i\|^2$.

Proof. Refer to Kammerer [4]

Lemma 5. The inequality $\|e_{i+i}\|^2 \leqq \|e_i\|^2-\alpha_i K\alpha(x_i)$ holds for i=0, 1, 2, ⋯.

Proof. Making use of Lemma 4(b), we get the following sequence of identities:

$$\|e_{i+1}\|^2=\|e_i\|^2-2\alpha_i[e_i,p_i]+\alpha_i^2\|p_i\|^2 \quad {}_i$$

$$=\|e_i\|^2-\alpha_i\{2K\alpha(x_i)-\alpha_i\|r_i\|^2\}\frac{\|p_i\|^2}{\|r_i\|^2}$$

$$=\|e_i\|^2-\alpha_i\{K\alpha(x_i)+K\alpha(x_{i+i})\}\frac{\|p_i\|^2}{\|r_i\|^2}$$

$$\leqq \|e_i\|^2-\alpha_i K\alpha(x_i).$$

Lemma 6. For any nonnegative integers i and j, both $[p_i,e_i]$ and $[e_i,e_j]$ are nonnegative.

Proof. Lemma 4 (b) shows that $[p_i,e_i]$ is nonnegative. To show that $[e_i,e_j]$ is nonnegative, we shall assume without loss of generality that $i \geqq j$. Then $e_j=e_i+\alpha_{i-1}p_{i-1}+\cdots+\alpha_j p_j$, and $[e_i,e_j]=[e_i,e_i]+\sum_{k=j}^{i-1}\alpha_k[e_i,p_k]$, which is nonnegative.

4. Convergence of the conjugate gradient method.

In this section, using the conjugate gradient method, we find an approximate solution U_α of the regularized operator equation $C_\alpha^* C_\alpha x=C_\alpha^* \bar{y}$.
We prove the convergence of the conjugate gradient method to a solution of $C_\alpha^* C_\alpha x=C_\alpha^* \bar{y}$.

THEOREM 7. In the assumptions of section 2-3, the conjugate gradient method (2)-(5) converges monotonically to the least squares solution $u=C_\alpha^+\bar{y}+(I-P)x_0$ of $C_\alpha x=\bar{y}$.

Proof. Refer to [5].

Lemma 8. If $Q\bar{y}\varepsilon R(C_\alpha C_\alpha^*)$, then for i=0, 1, 2, ⋯⋯,

(a) $\|z_{i+1}-\bar{z}\| \leqq \|z_i-\bar{z}\| \leqq \|z_0-\bar{z}\|$,

(b) $\|e_i\|^2 \leqq K_\alpha(x_i)\|z_0-\bar{z}\|^2$

and

(c) $\|e_{i+1}\|^2 \leqq (1-B\|e_i\|^2)\|e_i\|^2$

where $B=\|z_0-\bar{z}\|^2\|C_\alpha\|^{-2}$.

Proof. (a). (5) imply that $z_i-z_{i+1}=\alpha_i v^{-1}p_i$ for i=0, 1, 2, ⋯⋯,
where $U=C_\alpha^* \mid R(C_\alpha^*)$

$$\|z_{i+1}-\bar{z}\|^2=\|z_i-\bar{z}\|^2-2\alpha_i[U^{-1}p_i,z_i-\bar{z}]+\alpha^2{}_i\|U^{-1}p_i\|^2$$

$$=\|z_i-\bar{z}\|^2-\alpha_i\{2[U^{-1}p_i, z_i-\bar{z}]-[U^{-1}p_i, z_i-z_{i+1}]\}$$

$$=\|z_1-\bar{z}\|^2-\alpha_i[U^{-1}p_i, (z_i-\bar{z})+(z_{i+1}-\bar{z})]$$

$$=\|z_i-\bar{z}\|^2-\alpha_i[U^{-1}p_i,v^{-1}(e_i+e_{i+1})]$$

$= \| z_i - \bar{z} \|^2 - \alpha_i [v^{-1}u^{-1}p_i, e_i + e_{i+1}]$ where $V = C\alpha \mid \overline{R(C^*)}$

Therefore, we need only show that $[v^{-1}u^{-1}p_i, e_i + e_{i+1}]$ is nonnegative.

$[v^{-1}u^{-1}p_i, e_i + e_j] = \| r_i \|^2 \sum_{k=0}^{i} \frac{1}{\| r_k \|^2} [v^{-1}u^{-1}r_k, e_i + e_{i+1}] = \| r_i \|^2 \sum_{k=0}^{i} \frac{1}{\| r_k \|^2} [e_k, e_i + e_{i+1}]$,

which by Lemma 6 is nonnegative.

(b) Using the Cauch-Schwarz inequality and part (a), we obtain $\| e_i \|^2 = | [x_i - u, C^* z_i - u] |^2 = | [C_{\bar{a}}(x_i - u), z_i - \bar{z}] |^2 \leq k_a(x_i) \| z_0 - \bar{z} \|^2$.

(c) The boundedness of C_a show that

(9) $\alpha_i = \frac{\| r_i \|^2}{\| C_a p_i \|^2} \geq \frac{\| r_i \|^2}{\| C_a r_i \|^2} \geq \frac{1}{\| C_a \|^2}$

Part (C) is established by using of Lemma 5 and 8 (b) and (9) in the following manner:

$\| e_{i+1} \|^2 \leq \| e_i \|^2 - \alpha_i K\alpha(x_i) \leq \| e_i \|^2 - \frac{\alpha_i \| e_i \|^4}{\| z_0 - z \|} \leq (1 - \frac{\| e_i \|^2}{\| z_0 - z \|^2 \| C_a \|^2}) \| e_i \|^2$.

LEMMA 9. If the sequence $\{C_i\}$ of real numbers satisfies

$0 \leq C_{i+1} \leq (1 - BC_i)C_i$, $i = 0, 1, 2, \cdots\cdots$, with $B > 0$ and $0 < BC_0 \leq 1$ then

(10) $C_i \leq \frac{C_0}{1 + iBC_0}$ for $i = 0, 1, 2, \cdots\cdots$.

Proof. If $C_i = 0$, then $C_{i+n} = 0$ for $n = 0, 1, 2, \cdots\cdots$. Therefore, without loss of generality, we shall assume that $C_i > 0$ for all i.

The inequality $C_{k+1} < C_k$ for $k = 0, 1, 2, \cdots$, cna be established easily by induction. Then, summing the inequalities $\frac{1}{C_{k+1}} - \frac{1}{C_k} = \frac{C_k - C_{k+1}}{C_k C_{k+1}} \geq \frac{BC_k^2}{C_k C_{k+1}} > B$ from $k = 0$ to $k = i - 1$

yields $\frac{1}{C_i} - \frac{1}{C_0} > iB$.

Inequality (10) results when this inequality is solved for C_i.

THEOREM 10. If $Q\bar{y} \varepsilon R(C_a C_a^* C_a)$, then the conjugate gradient method (2)-(5), with initial value $x_0 \varepsilon R(C_a^* C_a)$, converges monotonically to the best approximate solution $u_a = C_a^+ \bar{y}$. In fact

(11) $\| x_i - u_a \|^2 \leq$

$$\frac{\| C_a \|^2 \| x_0 - C_a^+ \bar{y} \|^2 \| (C_a^*)^+ x_0 - (C_a C_a^*)^+ \bar{y} \|^2}{\| C_a \|^2 \| (C_a^*)^+ x_0 - (C_a C_a^* \bar{y})^+ \bar{y} \|^2 + i \| x_0 - C_a^* \bar{y} \|}$$

$i = 1, 2, \cdots\cdots$.

Proof. From Lemma 8(c), $\| e_{i+1} \|^2 \leq (1 - B \| e_i \|^2) \| e_i \|^2$,

where $B = \| z_0 - \bar{z} \|^{-2} \| C_a \|^{-2}$ holds for $i = 0, 1, 2, \cdots$ and

$B \| e_0 \|^2 B \| x_0 - u_a \|^2 = B \| C_a^* (z_0 - \bar{z}) \|^2 \leq \| C_a^* \|^2 \| C_a \|^{-2} = 1$. Lemma 9 can now be applied to this difference inequality showing that

$$\| e_i \|^2 \leq \frac{\| x_0 - u_\alpha \|^2}{1 + iB \| x_0 - u_\alpha \|^2}$$

Inequality (11) results when $x_0 - u_\alpha$ is replaced by $x_0 - C^+\bar{y}$ and relation $z = (C_\alpha C_\alpha^*)^+ y$ is utilized. This completes the proof of Theorm 10.

References

1. A. E. Taylor, Introduction to functional analysis, John-Wiley and Sons, New York, 1958.
2. M. Z. Nashed, Generalized inverses and applications, Academic Press, New York, 1976.
3. M. S. Song, Regularization-Projection methods and finite element approximations for ill-posed linear operator equations, ph. D. Dissrtation, University of Michign, 1978.
4. W. J. Kammerer and M. Z. Nashed, on the convergence of the conjugate gradient method for singular linear operator equations SIAM J. Nuer. Anal. vol. 9. 165-181, 1972.
5. D. H. Kim, The conjugate gradient method for least squares solutions of constrrained singular linear operator equations, Cheju National University J. 1985.

국 문 초 록

본 논문에서는 Weighted generalized inverse를 소개하고 Conjugate gradient 방법에 의해 형성되는 수열은 최적 근사치 해에 수렴한다는 것을 보이고 오차 범위를 결정하였다.