제한된 최소화 문제에 관한 Steepest Descent 방법

金 道 鉉*

The Steepest Descent Method for Constrained Minimization Problems

*Kim, Do-hyun**

Summary

In this paper, we investigate the solution of constrained minimization problem for some given conditions and we establish the convergence of the steepest descent method to the least squares solutions of minimal norm.

국문요약

제한된 최소화 문제에 관한 Steepest Descent 방법

본 논문에서는 제한된 최소화 문제의 해의 존재성과 유일성을 연구하고 Steepest descent method에 의해 만들어지는 수열은 최적근사치 해에 수렴한다는 것을 보이고 오차를 추정하였다.

Introduction

The operator equation $Ax=y$ where A is a mapping on some space into another has a solution if and only if y is in the range of A. This embodies the notion of a solution in the traditional sense; it is an ideal situation. On the other hand, one may look at the problem from a different angle.

In this paper, we investigate the solutions of constrained minimization problem for some given conditions, and we establish the convergence of the steepest descent method to the least squares solutions of minimal

* 사범대학 수학교육과

norm.

Let X, Y and Z be (real or complex) Hilbert spaces, and let A be a bounded linear operator on X into Y. Unless otherwise indicated, $\langle \cdot , \cdot \rangle$ and $\| \cdot \|$ refer to the inner product and the norm, respectively. 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 always exist for all $y \in Y$ if R(A) is closed. If R(A) is arbitrary, a least squares solution does not exist for all $y \in Y$; however it exists for all $y \in R(A) \oplus R(A)^{\perp}$

For any subspace S, we denote the orthogonal complement of S by $S^{\perp}$ and the closure of S by $\overline{S}$. Let D(A), R(A), and N(A) denote the domain, the range and the null space of a linear operator A, respectively. It is well known

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

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

2. Least squares solutions. Existence and Uniqueness of the regularized solution

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.

We consider the following minimization problem;

(2) For a given y in $D(A^{\dagger})$, let $S_y=\{u \in X : \|Au-y\|_Y=\inf\|Ax-y\|_Y, x \in X\}$. Then the problem is to find $w \in S_y$ such that $\|Lw\|_Z=\inf\{\|Lu\|_Z : u \in S_y\}$.

We state the conditions under which the solution of the problem exists and is unique.

Proposition 1. The constrained minimization problem (2) has a solution for every $y \in D(A^{\dagger})$ if and only if LN(A) is closed.

proof : Since for any $u \in S_y$, $u=A^{\dagger}y+v$ for some $v \in N(A)$, the problem (2) is equivalent to $\inf\{\|Lu\| : u \in LS_y\}=\inf\{\|L(A^{\dagger}y+v)\| : v \in N(A)\}=\inf\{\|u\| : u \in LS_y\}$. Note that LS_y is a translation of the subspace LN(A). Thus, we can easily check that the proposition holds.

Proposition 2. In Proposition 1, there exists a unique solution if and only if $N(A) \cap N(L)=\{0\}$.

proof : (⇐) Suppose that $N(A) \cap N(L) = \{0\}$. Then since $N(L_A) = \{0\}$, where L_A is the restriction of L onto N(A), there exists a unique $w_1 \in N(A)$ such that $\|Lw_1 + L(A^{\cdot}z)\| \leq \|Lx_1 + L(A^{\cdot}z)\|$ for all $x_1 \in N(A)$. It shows that there exists a unique $v = A^{\cdot}z + w_1$ such that $\|Lw\| \leq \|Lx\|$ for all $x \in S_y$.

(⇒) Suppose that $N(A) \cap N(L) \neq \{0\}$. Then there exists at least one $w_2 \in N(A) \cap N(L)$ which is not zero. Thus, $\|Lw\| = \|L(w+w_2)\| \leq \|Lx\|$ for all $x \in S_y$. Hence w is not unique.

Theorem 3. An element $w \in X$ is a solution to the constrained minimization problem (2) if and only if $A^*Aw = A^*y$ and $L^*Lw \in N(A)$.

Proof : Let $w \in Sy$ such that $\|Lw\|z \leq \|Lu\|z$ for all $u \in Sy$. Then $A^*Aw = A^*y$ is obvious.

For any $u \in Sy$, $u = A^{\cdot}y + v$ for some $v \in N(A)$. Since $\|Lw\| = \inf\{\|Lu\| : u \in Sy\}$, $\|L(A^{\cdot}y + w_1)\| \leq \|L(A^{\cdot}y + u_1)\|$ for all $u_1 \in N(A)$, i.e., $\|L(A^{\cdot}y) + L(w_1)\| \leq \|L(A^{\cdot}y) + L(u_1)\|$ for all $u_1 \in N(A)$ — (*).

Now consider the restriction of L onto N(A), denoted by L_A. Then (*) induces that $\|L_A(w_1) + L(A^{\cdot}y)\| \leq \|L_A(u_1) + L(A^{\cdot}y)\|$ for all $u_1 \in N(A)$. Since L has a closed range, L_A has also a closed range. It shows that $w_1 = L_A^+(-L(A_y^+))$. Consequently, $L_A(w_1) + L(A^{\cdot}y) = L(w_1 + A^{\cdot}y) \in R(L_A)^{\perp}$.

Thus for all $u_1 \in N(A)$, $L_A(u_1) \in R(L_A)$ and

$$
\begin{aligned}
(L_A(u_1), L_A(w_1) + L(A^{\cdot}y)) &= (L_A(u_1),\ L(A^{\cdot}y + w_1)) \\
&= (L(u_1),\ L(w_1 + A^{\cdot}y)) \\
&= (u_1, L^*L(w_1 + A^{\cdot}y)) \\
&= (u_1, L^*L(w)) = 0.
\end{aligned}
$$

Namely, $L^*L(w) \in N(A)^{\perp}$.

We define a new inner product in L :

$$[u, v] = \langle Au, Av\rangle_Y + \langle Lu, Lv\rangle_z \quad \text{for } u, v \in X$$

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

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 $y \to w$ and call it the weighted generalized inverse of A.

The operator equation (1) is said to be well-posed (relative to the spaces X and Y) if for each $y \in Y$, (1) has a unique least squares solution of minimal norm which depend continuously on y. Otherwise the equation is said to be ill-posed. When the range of A is closed, the minimization problem 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 :

$$W = Y \times Z$$

$$\langle (y_1 z_1), (y_2, z_2)\rangle_W = \langle y_1, y_2\rangle_Y + \langle z_1, z_2\rangle_Z \quad \text{for } y_1, y_2 \in Y \text{ and } z_1, z_2 \in Z.$$

For $\alpha > 0$, let C_α be a linear operator from X into W defined by $C_\alpha x = (A_x, \sqrt{\alpha}\, Lx)$ 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 4. Let $\alpha > 0$. An element x_α in X minimizes the quadratic functional $J_\alpha(x)$ if and only if $C_\alpha^* C_\alpha x = C_\alpha^* \bar{y}$

Proof : Refer to Song

$\|x\|$ and $\|x\|_L$ are equivalent if AN(L) is closed. In addition to assuming the existence and uniqueness of the solution, we assume that AN(L) is closed throughout the paper.

Theorem 5. For $\alpha>0$, let U_α be the unique solution of the operator equation (2). Then $\lim_{\alpha\to 0} U_\alpha = A_L^+ y$

Proof : Refer to Song.

3. Convergence of the steepest descent method

In this section, using the steepest descent method, we find approximate solution U_α of the regularized operator equation $C_\alpha^* C_\alpha x = C_\alpha^* \bar{y}$.

We prove the convergence of the steepest descent method to a solution of $C_\alpha^* C_\alpha x = C_\alpha^* \bar{y}$.

Let $x_0 \in X$ be an initial approximation to a least squares solution U_α of the equation (3) $C_\alpha x = \bar{y}$, $\bar{y} = (y, 0) \in W$.

We show that the method converges to the unique solution of $C_\alpha^* C_\alpha x = \bar{y}$ with minimal norm if and only if x_0 is in the range of C_α^*.

Let

$$J_\alpha(x) = \|Ax - y\|^2 + \alpha \|Lx\|^2 \quad \text{for } \alpha>0.$$

The steepest descent method for minimizing $J_\alpha(x)$ is given by (4) $x_{n+1} = x_n - \alpha_n \operatorname{grad} J_\alpha(x_n)$ where α_n is chosen to minimize $J_\alpha(x_{n+1})$.

It is easy to show that $\operatorname{grad} J_\alpha(x) = (A^*A + \alpha L^* L)x - A^* y$, so that the algorithm (4) may be written in the form

$$x_{n+1} = x_n - \alpha_n r_n$$

where $r_n = (A^*A + \alpha L^* L)x_n - A^* y$ and

$$\alpha_n = \frac{\|r_n\|^2}{\|Ar_n\|^2 + \alpha\|Lr_n\|^2} = \frac{\|r_n\|^2}{\|C_\alpha r_n\|^2}$$

Note that if $\alpha_k = 0$ for some k, then x_k is a least squares solution of (3). Also if $C_\alpha r_k = 0$, then $r_k = 0$ since C_α is one-to-one on the range of C_α^*. Thus we shall assume that $r_n \neq 0$ for all n.

Theorem 6. The sequence generated by steepest descent method defined by (4) converges to an element $u \in S_\alpha = \{z : \inf\|C_\alpha x - \bar{y}\| = \|C_\alpha z - y\|\}$. $\{x_n\}$ converges to u_α if and only if $x_0 \in R(C_\alpha^*)$ for any initial approximation $x_0 \in X$.

Proof. We first show that $\{x_n\}$ is a minimizing sequence for any initial approximation $x_0 \in X$. Using (4),

$$\begin{aligned} J_\alpha(x_{n+1}) &= \|Ax_{n+1} - y\|^2 + \alpha\|Lx_{n+1}\|^2 \\ &= \|Ax_n - \alpha_n A r_n - y\|^2 + \alpha\|Lx_n - \alpha_n L r_n\|^2 \\ &= J_\alpha(x_n) - \frac{\|r_n\|^4}{\|Ar_n\|^2 + \alpha\|Lr_n\|^2} \end{aligned}$$

Thus $J_\alpha(x_{n+1}) \le J_\alpha(x_n)$ for all n, with the equality holding when $r_n = 0$.

We obtain recursively, $J_\alpha(x_{n+1}) = J_\alpha(x_0) - \sum_{i=1}^{n} \frac{\|r_i\|^4}{\|Ar_i\|^2 + \alpha\|Lr_i\|^2}$

Note that since J_α is bounded below by zero,

$$\sum_{i=0}^{\infty} \frac{\|r_i\|^4}{\|Ar_i\|^2 + \alpha\|Lr_i\|^2} = \sum_{i=0}^{\infty} \frac{\|r_i\|^4}{\|C_\alpha r_i\|^2} < \infty$$

Moreover,

$$\frac{1}{\|C_\alpha\|^2} \sum_{i=0}^{\infty} \|r_i\|^2 \le \sum_{i=0}^{\infty} \frac{\|r_i\|^4}{\|C_\alpha r_i\|^2}$$

Hence, $\sum_{i=0}^{\infty} \|r_i\|^2 < \infty$. Therefore, $r_n = C_\alpha^* C_\alpha x_n -$

$C_\alpha^* \bar{y} \to 0$ as $n \to \infty$.

Now we show strong convergence of $\{x_n\}$. From (4) we obtain recurrsively, $x_{n+1} = x_0 - \sum_{i=0}^{n} \alpha_i r_i$ and for $m > n$, $x_m - x_n = -\sum_{i=n}^{m-1} \alpha_i r_i$. Since $r_i = (A^*A + \alpha L^*L)x_i - A^*\bar{y} \in R(C_\alpha^*)$ for all i,

$x_m - x_n \in R(C_\alpha^*)$ for all m and n.

There exists a positive number r such that

$$r^2 \| x_m - x_n \|^2 \le \langle C_\alpha^* C_\alpha (x_m - x_n), x_m - x_n \rangle = \| C_\alpha (x_m - x_n) \|^2$$

On the other hand,

$$\begin{aligned}&\langle C_\alpha^* C_\alpha (x_m - x_n), x_m - x_n \rangle \le \\ &|\langle C_\alpha^* C_\alpha x_m - C_\alpha^* \bar{y}, x_m - x_n \rangle| + |\langle C_\alpha^* C_\alpha x_n - C_\alpha^* \bar{y}, x_m - x_n \rangle| \le \\ &\frac{1}{r} (\| C_\alpha^* C_\alpha x_m - C_\alpha^* \bar{y} \| + \| C_\alpha^* C_\alpha x_n - C_\alpha^* \bar{y} \|) \| C_\alpha (x_m - x_n) \|\end{aligned}$$

But $\{\| C_\alpha (x_m - x_n) \|\}$ is bounded, say by M, since $r_i \to 0$ and C_α^* has a bounded inverse on $R(C_\alpha)$. Thus, $r^2 \| x_m - x_n \|^2 \le \frac{M}{r} (\| C_\alpha^* C_\alpha x_m - C_\alpha^* \bar{y} \| + \| C_\alpha^* C_\alpha x_n - C_\alpha^* \bar{y} \|)$. But the right-hand side of the above inequality goes to zero as m, n $\to \infty$, which shows that $\{x_n\}$ is a Cauchy sequence, and hence converges to an element $u \in X$: $x_n \to u_\alpha$ and $\lim_{n \to \infty} J_\alpha(x_n) = J_\alpha(u) = \inf \{J_\alpha(x) : x \in X\}$. Note that u is any least squares solution of the equation (3). Finally we show that $\{x_n\}$ converges to the least squares solution u_α of minimal norm if and only if $x_0 \in R(C_\alpha^*)$.

If $x_0 \in R(C_\alpha^*)$, then $x_n = x_0 - \sum_{i=0}^{n-1} \alpha_i r_i \in R(C_\alpha^*)$ since $r_i \in R(C_\alpha^*)$ for all i.

Therefore since $R(C_\alpha^*)$ is closed, $\{x_n\}$ converges to a least squares solution $u_\alpha \in R(C_\alpha^*)$ and $C_\alpha^+ \bar{y}$ is the unique least squares solution in $R(C_\alpha^*)$.

Thus $\{x_n\}$ converges to $C_\alpha^+ \bar{y}$.

Conversely, if $x_0 \notin R(C_\alpha^*)$, then $x_0 = x_0' + x_0''$ where $x_0' \in R(C_\alpha^*)$ and $x_0'' \in R(C_\alpha^*)^\perp = N(C_\alpha)$.

Hence $X_{n+1} = x_0' - \sum_{i=0}^{n} \alpha_i r_i + P_{N(C\alpha)} x_0$, where $P_{N(C\alpha)}$ denotes the orthogonal projection on $N(C_\alpha)$.

But $x_0 - \sum_{i=0}^{n} \alpha_i r_i$ converges to $C_\alpha^+ \bar{y}$.

Thus $\{x_n\}$ converges to $C_\alpha \bar{y} + P_{N(C\alpha)} x_0$.

This completes the proof of the theorem.

Remark. In view of the above theorem, and Kantorovich's error estimates for the steepest descent method for bounded linear operator (3), it follows that for any $x_0 \in R(C_\alpha^*)$, $\| x_n - C_\alpha^* \bar{y} \| \le \beta \left(\frac{M-m}{M+m} \right)^n$, where $m \| x \|^2 \le \langle C_\alpha x, C_\alpha x \rangle \le M \| x \|^2$, $x \in R(C_\alpha^*)$ and β is constant.

REFERENCES

Groetsch, C. W., 1977, Generalized inverses of linear operators, dekker, New York, pp.1 -89.

Groetsch, C. W., 1984, The theory of Tikhnov regularization of Fredholm equation of the first kind, Research Notes in Mathematics, Boston

Kantorvich, L. V., 1982, Functional Analysis, Pergamon press, New York pp.443-473.

Nashed M. Z., and Grace Wahba, 1974, Generalized Inverses in reproducing kernel spaces, SIAM J. Math. Anal. vol 5, NO.6, pp.974-987.

Nashed M. Z., 1970, Steepest Descent for singular Linear operator equations, SIAM J. Numer. Anal. 7, pp.358-362.

Song, M. S., 1985, J. of Kor. Math. Soc., 22, 35.