AN ITERATIVE APPROACH TO A CONSTRAINED LEAST SQUARES PROBLEM

SIMEON REICH AND HONG-KUN XU

Received 3 December 2001

A constrained least squares problem in a Hilbert space $H$ is considered. The standard Tikhonov regularization method is used. In the case where the set of the constraints is the nonempty intersection of a finite collection of closed convex subsets of $H$, an iterative algorithm is designed. The resulting sequence is shown to converge strongly to the unique solution of the regularized problem. The net of the solutions to the regularized problems strongly converges to the minimum norm solution of the least squares problem if its solution set is nonempty.

1. Introduction

Consider the following constrained least squares problem in a Hilbert space $H$:

$$
\min_{x \in C} f(x) := \frac{1}{2} \|Ax - b\|^2,
$$

where $C$ is a nonempty closed convex subset of $H$, $A : H \to H$ is a bounded linear operator, and $b$ is an element of $H$.

Let $S$ denote the solution set of (1.1). It is known that $S$ is closed and convex, but it may well be empty.

Since problem (1.1) may be ill-posed, a regularization process is necessary in order to find a solution of (1.1) (if any). We will apply the standard Tikhonov regularization to (1.1); that is, we consider the following regularized problem ($\epsilon > 0$):

$$
\min_{x \in C} f_\epsilon(x) := \frac{1}{2} \|Ax - b\|^2 + \frac{1}{2} \epsilon \|x\|^2.
$$

(1.2)
Constrained least squares problem

(The reader is referred to [6] by Engl et al. for regularization methods for inverse
and ill-posed problems. See also Tikhonov and Arsenin [12] and Björck [3].)

Since \( f_\varepsilon \) is continuous, convex, and coercive (i.e., \( \|x\| \to \infty \Rightarrow f_\varepsilon(x) \to \infty \)),
(1.2) has a unique solution which we denote by \( x_\varepsilon \).

In this paper, we first discuss the asymptotic behavior of \( \{x_\varepsilon\} \) as \( \varepsilon \downarrow 0 \) and then
design an iterative algorithm which produces a sequence converging strongly to
the unique solution \( x_\varepsilon \) of (1.2) in the case where \( C \) is the nonempty intersection
of a finite collection of closed convex subsets of \( H \). Moreover, if \( S \) is nonempty,
then the net \( \{x_\varepsilon\} \) of the approximate solutions converges strongly to the mini-
mum norm solution of (1.1).

2. Behavior of \( \{x_\varepsilon\} \)

Let \( C, f, f_\varepsilon, \{x_\varepsilon\}, \) and \( S \) be given as in the introduction. It is the objective of this
section to study the behavior of the net \( \{x_\varepsilon\} \) of the solutions of the regularized
problem (1.2). The next result shows that as long as the solution set \( S \) of (1.1)
is nonempty, \( \{x_\varepsilon\} \) behaves well and approaches the minimum norm solution of
(1.1).

**Theorem 2.1.** The behavior of \( \{x_\varepsilon\} \) is as follows:

(i) \( S \neq \emptyset \) if and only if \( \{x_\varepsilon\} \) is bounded as \( \varepsilon \downarrow 0 \). In this case, \( \{x_\varepsilon\} \) converges
strongly as \( \varepsilon \downarrow 0 \) to the minimum norm solution of (1.1);
(ii) \( S = \emptyset \) if and only if \( \lim_{\varepsilon \downarrow 0} \|x_\varepsilon\| = \infty \).

**Proof.** To prove (i), first assume that \( S \neq \emptyset \). Since \( S \) is closed and convex, there
is a unique element in \( \overline{S} \), denoted by \( S^* \), such that

\[
\|S^*\| = \min \{\|v\| : v \in S\}. \tag{2.1}
\]

Since \( f_\varepsilon(x_\varepsilon) \leq f_\varepsilon(S^*) \), we see that

\[
\varepsilon(\|x_\varepsilon\|^2 - \|S^*\|^2) \leq \|AS^* - b\|^2 - \|Ax_\varepsilon - b\|^2. \tag{2.2}
\]

But \( f(S^*) \leq f(x_\varepsilon) \); that is, \( \|AS^* - b\|^2 \leq \|Ax_\varepsilon - b\|^2 \), it follows from (2.2) that

\[
\|x_\varepsilon\| \leq \|S^*\|. \tag{2.3}
\]

Hence, \( \{x_\varepsilon\} \) is bounded.

Next, assume that \( \{x_\varepsilon\} \) is bounded as \( \varepsilon \downarrow 0 \). Assume that \( \varepsilon_n \downarrow 0 \) and \( x_{\varepsilon_n} \rightharpoonup \bar{x} \)
weakly. Clearly, \( \bar{x} \in C \).
We claim that \( \bar{x} \in S \). To see this, we note that the weak lower semicontinuity of \( f \) implies that for any \( x \in C \),
\[
 f(\bar{x}) \leq \liminf_{n \to \infty} f(x_{\varepsilon_n}) \leq \limsup_{\varepsilon \downarrow 0} f(x_{\varepsilon}) = \limsup_{\varepsilon \downarrow 0} f_\varepsilon(x_{\varepsilon}) \\
\leq \limsup_{\varepsilon \downarrow 0} f_\varepsilon(x) = f(x). \tag{2.4}
\]
This shows that \( \bar{x} \in S \).

Next, we prove, assuming that \( S \neq \emptyset \), that \( \{x_{\varepsilon}\} \) converges strongly as \( \varepsilon \downarrow 0 \) to \( S^\circ \), the minimum norm solution of (1.1). Let \( v \) be any weak limit point of \( \{x_{\varepsilon}\} \) when \( \varepsilon \downarrow 0 \). Then, we have a sequence \( \varepsilon_n \downarrow 0 \) such that \( x_{\varepsilon_n} \to v \) weakly. Since the norm of \( H \) is weakly lower semicontinuous, it follows from (2.3) that
\[
\|v\| \leq \|S^\circ\|. \tag{2.5}
\]
Due to the uniqueness of the minimum norm element, this implies that \( v = S^\circ \), which in turn implies that \( S^\circ \) is the only weak limit point of \( \{x_{\varepsilon}\} \). Therefore, again, due to the weak lower semicontinuity of the norm and (2.3), we have
\[
w - \lim_{\varepsilon \downarrow 0} x_{\varepsilon} = S^\circ, \quad \lim_{\varepsilon \downarrow 0} \|x_{\varepsilon}\| = \|S^\circ\|. \tag{2.6}
\]
This implies that \( s - \lim_{\varepsilon \downarrow 0} x_{\varepsilon} = S^\circ \).

In order to prove (ii), we first note that if \( S \neq \emptyset \), then by (i) \( \{x_{\varepsilon}\} \) is bounded. So, it remains to show that \( S = \emptyset \) implies
\[
\lim_{\varepsilon \downarrow 0} \|x_{\varepsilon}\| = \infty. \tag{2.7}
\]
Suppose on the contrary that (2.7) does not hold. Then we can find a sequence \( \{\varepsilon_n\} \downarrow 0 \) and some \( \tilde{x} \in C \) such that
\[
x_{\varepsilon_n} \rightharpoonup \tilde{x} \quad \text{weakly.} \tag{2.8}
\]
Repeating the first half of the proof of part (i), we obtain that
\[
f(\tilde{x}) \leq f(x), \quad x \in C. \tag{2.9}
\]
That is, \( \tilde{x} \in S \) and \( S \) is nonempty. The contradiction we have reached completes the proof. \( \square \)

**Remark 2.2.** Let \( y \in H \) and let \( x_\varepsilon(y) \) denote the unique solution of the regularized optimization problem
\[
\min_{x \in C} f_\varepsilon(x) := \frac{1}{2}\|Ax - b\|^2 + \frac{1}{2}\varepsilon\|x - y\|^2. \tag{2.10}
\]
Let \( P_S \) be the nearest point projection from \( H \) onto \( S \). Then, we can prove the following results:

(i) \( S \neq \emptyset \) if and only if \( \{ x_\varepsilon(y) \} \) is bounded as \( \varepsilon \downarrow 0 \). In this case, \( \{ x_\varepsilon(y) \} \) converges strongly as \( \varepsilon \downarrow 0 \) to the solution of (1.1) closest to \( y \) (i.e., \( P_S y \));

(ii) \( S = \emptyset \) if and only if \( \lim_{\varepsilon \downarrow 0} \| x_\varepsilon(y) \| = \infty \).

### 3. Iterative algorithm

In this section, we will introduce an iterative algorithm which produces a sequence that converges strongly to the unique solution of the regularized problem (1.2) in the case where the set \( C \) of constraints is the intersection of finitely many closed convex subsets of \( H \). We need some lemmas. Lemma 3.1 is elementary (see [14] for a proof). Lemma 3.2 is an easy exercise in an inner product space. Lemma 3.3 is well known in the fixed-point theory of nonexpansive mappings (see, e.g., Goebel and Kirk [7]). Lemma 3.4 may be known. A proof is, however, included.

**Lemma 3.1.** Let \( \{ s_n \} \) be a sequence of nonnegative numbers satisfying

\[
 s_{n+1} \leq (1 - \alpha_n) s_n + \alpha_n \beta_n, \quad n \geq 0, 
\]

(3.1)

where \( \{ \alpha_n \} \) and \( \{ \beta_n \} \) are sequences of real numbers such that

(i) \( \{ \alpha_n \} \subset [0, 1] \) and \( \sum_{n=0}^{\infty} \alpha_n = \infty \) or equivalently,

\[
 \prod_{n=0}^{\infty} (1 - \alpha_n) := \lim_{n \to \infty} \prod_{k=0}^{n} (1 - \alpha_k) = 0; 
\]

(ii) \( \limsup_{n \to \infty} \beta_n \leq 0 \), or \( \sum_n \alpha_n \beta_n \) is convergent.

Then \( \lim_{n \to \infty} s_n = 0 \).

**Lemma 3.2.** Let \( H \) be a Hilbert space. Then \( \| x + y \|^2 \leq \| x \|^2 + 2 \langle y, x + y \rangle \), for all \( x, y \in H \).

**Lemma 3.3.** Let \( D \) be a closed convex subset of \( H \) and \( T : D \to H \) a nonexpansive mapping with a fixed point. If \( \{ x_n \} \) is a sequence in \( D \) such that \( x_n \rightharpoonup x \) weakly and \( (I - T)x_n \to y \) strongly, then \( (I - T)x = y \). (Here \( I \) is the identity operator on \( H \).)

**Lemma 3.4** (the optimality condition for (1.2)). The point \( x_\varepsilon \in C \) solves (1.2) if and only if

\[
 \langle x - x_\varepsilon, A^* b - (A^* A + \varepsilon I) x_\varepsilon \rangle \leq 0, \quad x \in C, 
\]

(3.3)

where \( A^* \) is the adjoint of \( A \).
Proof. Assume that $x_\varepsilon \in C$ is the minimizer of $f_\varepsilon$ over $C$. We have, for any $x \in C$,

$$0 \leq \lim_{t \to 0} \frac{1}{t} \left[ f_\varepsilon(x_\varepsilon + t(x - x_\varepsilon)) - f_\varepsilon(x_\varepsilon) \right]$$

$$= \lim_{t \to 0} \left[ \langle Ax_\varepsilon - b, A(x - x_\varepsilon) \rangle + \varepsilon \langle x_\varepsilon, x - x_\varepsilon \rangle + \frac{t}{2} \left( ||A(x - x_\varepsilon)||^2 + \varepsilon ||x - x_\varepsilon||^2 \right) \right]$$

$$= \langle A^*(Ax_\varepsilon - b) + \varepsilon x_\varepsilon, x - x_\varepsilon \rangle$$

$$= \langle (A^*A + \varepsilon I)x_\varepsilon - A^*b, x - x_\varepsilon \rangle. \quad (3.4)$$

Assume next that (3.3) holds. Calculations yield the derivative

$$f'(x) = (A^*A + \varepsilon I)x - A^*b. \quad (3.5)$$

Hence, by the subdifferential inequality, we obtain, for all $x \in C$,

$$f(x) \geq f(x_\varepsilon) + \langle f'(x_\varepsilon), x - x_\varepsilon \rangle$$

$$= f(x_\varepsilon) + \langle (A^*A + \varepsilon I)x_\varepsilon - A^*b, x - x_\varepsilon \rangle \geq f(x_\varepsilon). \quad (3.6)$$

This shows that $x_\varepsilon$ is indeed a minimizer of $f_\varepsilon$ over $C$. \hfill \Box

We now assume that $S \neq \emptyset$ and look for the minimum norm solution of (1.1) in the case where

$$C = \bigcap_{i=1}^{N} C_i, \quad (3.7)$$

where $N \geq 1$ is an integer and $\{C_1, \ldots, C_N\}$ is a finite collection of closed convex subsets of a Hilbert space $H$. We assume that for each $i$, $C_i$ is the fixed-point set of a nonexpansive mapping $T_i$ (i.e., $C_i = \{x \in H : T_i x = x\}$). Recall that a mapping $T : H \to H$ is nonexpansive if $||Tx - Ty|| \leq ||x - y||$ for all $x, y \in H$.

It is well known that the (nearest point) projection $P_D$ from $H$ onto a closed convex subset $D$ of $H$ is nonexpansive. Note that each $C_i$ is the fixed-point set of the projection $P_{C_i}$.

We now introduce the iterative algorithm. Let a sequence $\{\alpha_n\}$ of real numbers in the interval $[0, 1]$ be given. For $\varepsilon > 0$, we put

$$B \equiv B_\varepsilon := A^*A + \varepsilon I, \quad u := A^*b. \quad (3.8)$$
Starting with an arbitrary initial guess $x_0 \in H$, we define the sequence $\{x_n\}$ recursively by the following algorithm:

$$
x_1 = (I - \alpha_1 B) T_1 x_0 + \alpha_1 u,
$$
$$
x_2 = (I - \alpha_2 B) T_2 x_1 + \alpha_2 u,
$$
$$
\vdots
$$
$$
x_N = (I - \alpha_N B) T_N x_{N-1} + \alpha_N u,
$$
$$
x_{N+1} = (I - \alpha_{N+1} B) T_1 x_N + \alpha_{N+1} u,
$$
$$
\vdots
$$

We can rewrite the algorithm in a more compact form

$$
x_{n+1} = (I - \alpha_{n+1} B) T_{n+1} x_n + \alpha_{n+1} u, \quad n \geq 0,
$$

where $T_n = T_{n \text{mod} N}$ and the mod function takes values in $\{1, 2, \ldots, N\}$.

**Remark 3.5.** The algorithm is a modification of an iteration scheme proposed by Halpern [8] and developed by Lions [9], Wittmann [13], Reich [10, 11], Bauschke [1], Xu [14, 15], and others. The algorithm and its variants apply to the convex feasibility problem (CFP). (See [2, 4] for some of the recent developments regarding the CFP.)

We now prove the main result of the paper. Some related work can also be found in a paper by Deutsch and Yamada [5].

**Theorem 3.6.** Assume that

1. $\alpha_n \to 0$;
2. $\sum_n \alpha_n = \infty$;
3. $\alpha_n/\alpha_{n+1} \to 1$ or $\sum_n |\alpha_n - \alpha_{n+1}| < \infty$;
4. $C = \bigcap_{i=1}^N F(T_i) = F(T_1 T_2 \cdots T_N) = F(T_N T_1 \cdots T_{N-1}) = \cdots = F(T_2 T_3 \cdots T_N T_1)$.

Then, the sequence $(x_n)$ generated by algorithm (3.10) converges in norm to the unique solution $x_\varepsilon$ of (1.2).

**Proof.** We will divide the proof into several steps.

1. First, we show that the sequence $\{x_n\}$ is bounded. Indeed, since $B$ is self-adjoint, we have that for any $0 < \alpha < (\|A\|^2 + \varepsilon)^{-1}$, $I - \alpha B$ is positive. Hence,

$$
\|I - \alpha B\| = \sup_{\|x\|=1} \langle (I - \alpha B)x, x \rangle \leq 1 - \alpha \varepsilon.
$$
(Without loss of generality, we may assume throughout the proof that \( \alpha_n < (\|A\|^2 + \epsilon)^{-1} \) for all \( n \)). It follows from (3.11) that for \( p \in C \),
\[
\|x_{n+1} - p\| = \|(I - \alpha_{n+1}B)(T_{n+1}x_n - p) + \alpha_{n+1}(u - Bp)\| \\
\leq (1 - \epsilon\alpha_{n+1})\|x_n - p\| + \alpha_{n+1}\|u - Bp\|. \tag{3.12}
\]
Hence, by induction, we obtain
\[
\|x_n - p\| \leq \max\left\{\|x_0 - p\|, \frac{\|u - Bp\|}{\epsilon}\right\}, \quad n \geq 0. \tag{3.13}
\]

(2) Next, we note that \( \|x_{n+1} - T_{n+1}x_n\| \to 0 \). This follows from assumption (1), step (1), and the fact that \( \|x_{n+1} - T_{n+1}x_n\| = \alpha_{n+1}\|u - BT_{n+1}x_n\| \).

(3) Now, we prove that \( \|x_{n+N} - x_n\| \to 0 \). Indeed, we have (note that \( T_{n+N} = T_n \))
\[
\|x_{n+N} - x_n\| = \|(I - \alpha_{n+N}B)(T_{n+N}x_{n+N-1} - T_nx_{n-1}) \\\n+ (\alpha_{n+N} - \alpha_n)(u - BT_nx_{n-1})\| \\
\leq (1 - \epsilon\alpha_{n+N})\|x_{n+N-1} - x_{n-1}\| + M|\alpha_{n+N} - \alpha_n| \\
= (1 - \epsilon\alpha_{n+N})\|x_{n+N-1} - x_{n-1}\| + \epsilon\alpha_{n+N}\beta_n, \tag{3.14}
\]
where \( M \) is a constant such that \( \|u - BT_nx_{n-1}\| \leq M \) for all \( n \) and \( \beta_n := M\epsilon^{-1}\alpha_{n+N}^{-1}|\alpha_{n+N} - \alpha_n| \). Applying Lemma 3.1 together with assumptions (2) and (3), we obtain that \( \|x_{n+N} - x_n\| \to 0 \).

(4) Next, we observe that \( x_n - T_{n+N} \cdots T_{n+1}x_n \xrightarrow{s} 0 \).

Indeed, noting that each \( T_n \) is nonexpansive and using step (2), we observe that
\[
\begin{align*}
    x_{n+N} - T_{n+N}x_{n+N-1} & \xrightarrow{s} 0, \\
    T_{n+N}x_{n+N-1} - T_{n+N}T_{n+N-1}x_{n+N-2} & \xrightarrow{s} 0, \\
    \vdots \\
    T_{n+N} \cdots T_{n+2}x_{n+1} - T_{n+N} \cdots T_{n+1}x_n & \xrightarrow{s} 0.
\end{align*} \tag{3.15}
\]
Adding up and using step (3), we see that \( x_n - T_{n+N} \cdots T_{n+1}x_n \xrightarrow{s} 0 \).

(5) Now, we show that \( \limsup_{n \to \infty} \langle u - Bx_\epsilon, x_n - x_\epsilon \rangle \leq 0 \), where \( x_\epsilon \) is the unique solution of (1.2).
Take a subsequence \( \{x_{n_j}\} \) of \( \{x_n\} \) such that
\[
\limsup_{n \to \infty} \langle u - Bx_\varepsilon, x_n - x_\varepsilon \rangle = \lim_{j \to \infty} \langle u - Bx_\varepsilon, x_{n_j} - x_\varepsilon \rangle. \tag{3.16}
\]
Since \( \{x_n\} \) is bounded, we may also assume that there exists some \( \tilde{x} \in H \) such that \( x_{n_j} \rightharpoonup \tilde{x} \). Since the pool of mappings is finite, we may further assume (passing to a further subsequence if necessary) that, for some \( i \in \{1, 2, \ldots, N\} \),
\[
T_{n_j} \equiv T_i, \quad \forall \ j \geq 1. \tag{3.17}
\]
It follows from step (4) that
\[
x_{n_j} - T_{i+N} \cdots T_{i+1}x_{n_j} \rightharpoonup 0. \tag{3.18}
\]
\textbf{Lemma 3.3} then ensures that the weak limit \( \tilde{x} \) of \( \{x_{n_j}\} \) is a fixed point of the mapping \( T_{i+N} \cdots T_{i+1} \). This together with assumption (4) implies that \( \tilde{x} \in F(T_{i+N} \cdots T_{i+1}) = C \). Therefore, we have, by (3.16) and \textbf{Lemma 3.4}, that \( \lim_{n \to \infty} \langle u - Bx_\varepsilon, x_n - x_\varepsilon \rangle = 0 \).

(6) Finally, we claim that \( x_n \rightharpoonup x_\varepsilon \). Indeed, we can write
\[
x_{n+1} - x_\varepsilon = (I - \alpha_{n+1}B)(T_{n+1}x_n - x_\varepsilon) + \alpha_{n+1}(u - Bx_\varepsilon). \tag{3.19}
\]
Apply \textbf{Lemma 3.2} to get
\[
\|x_{n+1} - x_\varepsilon\|^2 \leq \|(I - \alpha_{n+1}B)(T_{n+1}x_n - x_\varepsilon)\|^2 + 2\alpha_{n+1}\langle u - Bx_\varepsilon, x_{n+1} - x_\varepsilon \rangle
\leq (1 - \varepsilon\alpha_{n+1})\|x_n - x_\varepsilon\|^2 + 2\alpha_{n+1}\langle u - Bx_\varepsilon, x_{n+1} - x_\varepsilon \rangle. \tag{3.20}
\]
Using \textbf{Lemma 3.1} and step (5), we now conclude that \( \|x_n - x_\varepsilon\|^2 \to 0 \). □

\textbf{Remark 3.7.} Assumption (4) in \textbf{Theorem 3.6} is automatically satisfied if each \( T_i \) is \textit{firmly nonexpansive} (i.e., \( T_i \) satisfies the property: \( \langle x - y, T_i x - T_i y \rangle \geq \|T_i x - T_i y\|^2 \), for \( x, y \in H \)). Since a projection is firmly nonexpansive, we have the following consequence of \textbf{Theorem 3.6}.

\textbf{Corollary 3.8.} Assume that
\begin{itemize}
  \item[(1)] \( \alpha_n \to 0 \);
  \item[(2)] \( \sum_n \alpha_n = \infty \);
  \item[(3)] \( \alpha_n / \alpha_{n+1} \to 1 \) or \( \sum_n |\alpha_n - \alpha_{n+1}| < \infty \).
\end{itemize}
Let \( x_0 \in H \) be chosen arbitrarily and let \( \{x_n\} \) be generated by the iterative algorithm
\[
x_{n+1} = (I - \alpha_{n+1}B)P_{n+1}x_n + \alpha_{n+1}u, \quad n \geq 0, \tag{3.21}
\]
where \( P_k = P_{C_k} (1 \leq k \leq N) \). Then, \( \{x_n\} \) converges in norm to the unique solution \( x_\varepsilon \) of (1.2).
Remark 3.9. We have used two steps to find a solution of problem (1.1); that is, we first find the unique solution $x_\varepsilon$ of the regularized problem (1.2) and then approximate $x_\varepsilon$ via the iterative algorithm (3.10). It is not clear if one can combine the two steps into one. In other words, if one defines a sequence $\{x_n\}$ by

$$x_{n+1} = (I - \alpha_{n+1}B_{\varepsilon_n}) T_{n+1} x_n + \alpha_{n+1} u, \quad n \geq 0,$$

(3.22)

where $\{\varepsilon_n\} \subset (0, 1)$ and $\varepsilon_n \to 0$ as $n \to \infty$, does the sequence $\{x_n\}$ defined by (3.22) converge to the minimum norm solution of problem (1.1) (if the solution set $S$ of (1.1) is nonempty)?

Acknowledgment

The first author was partially supported by the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities (Grant 592/00), by the Fund for the Promotion of Research at the Technion, and by the Technion VPR Fund.

References


Constrained least squares problem


Simeon Reich: Department of Mathematics, The Technion - Israel Institute of Technology, 32000 Haifa, Israel
E-mail address: sreich@techunix.technion.ac.il

Hong-Kun Xu: Department of Mathematics, University of Durban-Westville, Private Bag X54001, Durban 4000, South Africa
E-mail address: hlxu@pixie.udw.ac.za