A NOTE ON THE ITERATIVE METHOD FOR COMPUTING THE SOLUTION OF OVERDETERMIND SYSTEM

Hwai Nien Yang

INTRODUCTION

An overdetermined system of linear equations is the form

(1) $AX=b$,

where the coefficient matrix is an m by n real matrix, $m>n$. The solution of this system has been provided by three methods: the least squares [1, P. 146], the Chebyshev or minimax [2, P. 375] and the best approximation [1, P. 146, 3, PP 17-19, 4, PP 38-43].

The objective of this note is to suggest use of the iterative method for the computing the approximate solution of (1). The iterative method is based on the convergent solution sequence of the following consistent system with a unique solution

(2) $AX=y$,

where A is an n square real matrix having the rank of n and which can also be denoted by [5, P. 96]

(3) $X_n=B(I+E+E^2+\cdots\cdots+E^n)y$,

where B is an approximate inverse of the non-singular n square matrix A, $n=1, 2, 3, \ldots$, and $E=I-AB$.

PRELIMINARY

1. For any m by n real matrix A, there is a unique generalized inverse n by m real matrix, denoted by A^-, having the useful properties as follows:

(4) $AA^-=(AA^-)'$,

(5) $A^-A=(A^-A)'$,

(6) $AA^-A=A$,

(7) $A^-AA^-=A^-$.

2. Let A be an m by n real matrix having the rank of k, ($k<m$, $k<n$) then A can be written in the form, $A=FR'$, where F is an m by k matrix and R' is a k by n matrix, defined by A. S. Householder in [6, P. 9]

(8) $AA^-=P_{r(A)}=F(F'F)^{-1}F'$,

(9) $A^{-}A = P_{r(A')} = R(R'R)^{-1}R'$,

(10) $A^{-} = R(R'R)^{-1}(F'F)^{-1}F'$,

where $P_{r(A)}$ and $P_{r(A')}$ are obviously the idempotent matrices because of (6) and (7).

3. The system of linear equations, $AX=b$, is consistent if and only if [1, P. 103]

(11) $b = AA^{-}b$,

and the best approximate solution X' is the form [1, P. 146]

(12) $X' = A^{-}b$.

RESULTS

Theorem 1. Let X_o be the initial approximate solution of the consistent system,

(13) $AX=b$,

where A is an m by n real matrix having the rank of k, ($k<m$, $k<n$), then the nth approximate solution is the form

(14) $X_n = B(I+E+E^2+\cdots\cdots+E^n)b$,

where B is an n by m approximately generalized inverse of A, $B \doteq A^{-}$, $n=1, 2, 3, \ldots$, and E can be written in the following form

(15) $E = AA^{-} - AB = P_{r(A)} - AB$.

Proof. Let the initial approximate solution be X_o and

(16) $X_o = Bb$,

then the residual vector r_o can be written in the following form

(17) $r_o = b - AX_o$,

from $b=AA^{-}b$ (11), $X_o=Bb$ (16), and $E=AA^{-}-AB$ (15), or

(18) $r_o = AA^{-}b - ABb = Eb$.

Obviously, we have

(19) $r_{n-1} = b - AX_{n-1}$,

and

(20) $X_n = X_{n-1}, + Br_{n-1}$,

for $n=1$.

Thus, we can show that

(21) $r_n = (AA^{-} - AB)^{n+1}b = E^{n+1}b$.

Because $r_1 = b - AX_1$(19) and $X_1 = X_o + Br_o$(20), we have

(22) $r_1 = b - A(X_o + Br_o)$,

and substituting $b=AA^{-}b$(11) and $A=AA^{-}A$(6) into the right hand side of(22), we obtain

(23) $r_1 = AA^-b - AA^-AX^o - ABr_o$.

Because the idempotent $AA^- = AA^-AA^-(8)$, $r_o = Eb(18)$, and $X_o = Bb(16)$, we have

(24) $r_1 = AA^-AA^-b - AA^-ABb - ABEb$,

and simplifying, we obtain

(25) $r_1 = E b$.

Similarly, from the iterative chains, we obtain

(26) $r_{n-1} = E^n b$.

Premultiply both sides of $X_n = X_{n-1} + Br_{n-1}$ (20) by A and substract it from b, we have

(27) $r_n = b - AX_n = b - AX_{n-1} - ABr_{n-1}$.

From (24), (25) and (26), we obtain by induction

(28) $r_n = (AA^- - AB)^{n+1} = E^{n+1}b$.

Since $X_o = Bb$ (18) and $X_n = X_{n-1} + Br_{n-1}$ (20) for $n=1$, therefore from the iterative chains, we obtain

$$X_n = B(I + E + E^2 + \ldots + E^n)b.$$

Theorem 2. The solution sequence $X_1, X_2, X_3, \ldots, X_n$(14) converges to the best approximate solution, $X' = A^-b$(12), of the consistent system, $AX = b$, as n tends to infinity.

Proof. The series $I + E + E^2 + \ldots + E^n$ on the right hand side of (14) converges to $(I - E)^{-1}$ as n tends to infinity under the condition $\| E \| < 1$ [4, P. 109]. Thus, we can show that

(29) $\lim_{n \to \infty} X_n = A^-b$.

Let the approximately generalized inverse B of A be $\frac{A'}{c}$ and let us choose a suitable value of c such that $\| AA^- - AB \| = \| E \| < 1$ [7, PP 452-455]. Thus, it is evident that

(30) $\lim_{n \to \infty} X_n = B(I - E)^{-1}b$.

Substituting $b = AA^-b$(11) into the right hand side of (30), we have

(31) $\lim_{n \to \infty} X_n = B(I - E)^{-1}AA^-b$.

or

(32) $\lim_{n \to \infty} X_n = B[(AA^-)^-(I - E)]^-b$.

Substituting $E = AA^- - AB$(15) into the right hand side of (32) and simplifying, we obtain

(33) $\lim_{n \to \infty} X_n = B(AA^-AB)^-b$,

or because $AA^{-}A=A(6)$, so

$$(34)\qquad \lim_{n\to\infty} X_n=B(AB)^{-}b,$$

or

$$(35)\qquad \lim_{n\to\infty} X_n=BB^{-}A^{-}b.$$

Substituting $B=\frac{A'}{c}$ and $B^{-}=c(A')^{-}$ into the right hand side of (35), we have

$$(36)\qquad \lim_{n\to\infty} X_n=A'(A')^{-}A^{-}b,$$

or because $A^{-}A=(A^{-}A)'(5)$ and $A^{-}AA^{-}=A^{-}(7)$, we obtain

$$(37)\qquad \lim_{n\to\infty} X_n=A^{-}b.$$

REMARKS

1. By use of the form (14) and $X^{o}=Bb(16)$, the following solution sequence can be obtained recursively:

$$X_1=B(I+E)b=X_o+BEb,$$
$$X_2=B(I+E+E^2)b=X_1+BE^2b,$$
$$X_3=B(I+E+E^2+E^3)b=X_2+BE^3b,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

and so on.

Thus, the computation can be conveniently written in the form of a formula

$$(38)\qquad X_n=X_{n-1}+BE^n b,$$

where $n=1, 2, 3,$.

Example 1.

$$2x_1-x_2+x_3=8,$$
$$x_1+2x_2-x_3=5,$$

where $A=\begin{pmatrix}2 & -1 & 1\\ 1 & 2 & -1\end{pmatrix}$ and $b=\begin{pmatrix}8\\5\end{pmatrix}$,

1. Let $A=FR'$, where $F=\begin{pmatrix}2 & -1\\ 1 & 2\end{pmatrix}$ and $R'=\frac{1}{5}\begin{pmatrix}5 & 0 & 1\\ 0 & 5 & -3\end{pmatrix}$, and then we shall find $AA^{-}=F(F'F)^{-1}F'=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$ and $A^{-}=R(R'R)^{-}1\,(F'F)^{-1}F'=\frac{1}{35}\begin{pmatrix}13 & 8\\ -4 & 11\\ 5 & -5\end{pmatrix}$. Hence, the best approximate solution $X'=\frac{1}{35}\begin{pmatrix}144\\ 23\\ 15\end{pmatrix}=\begin{pmatrix}4.11428571\\ 0.65714285\\ 0.42857142\end{pmatrix}$.

2. To find the approximate inverse of B by choosing $c=5$, we shall have $B=\frac{A'}{c}=\frac{1}{5}\begin{pmatrix}2 & 1\\ -1 & 2\\ 1 & -1\end{pmatrix}$, $E=AA^{-}-AB=\frac{1}{5}\begin{pmatrix}-1 & 1\\ 1 & -1\end{pmatrix}$, and the initial approximate sol-

ution $X_0=Bb=\frac{1}{5}\begin{pmatrix}21\\2\\3\end{pmatrix}$. The formula $X_n=X_{n-1}+BE^nb$ (38) can be used recursively to obtain

$X_1=\begin{pmatrix}4.08\\0.76\\0.36\end{pmatrix}$, $X_2=\begin{pmatrix}4.128\\0.616\\0.456\end{pmatrix}$, $X_3=\begin{pmatrix}4.1088\\0.6736\\0.4160\end{pmatrix}$, $X_4=\begin{pmatrix}4.11648\\0.65056\\0.43296\end{pmatrix}$, $X_5=\begin{pmatrix}4.113408\\0.659776\\0.426816\end{pmatrix}$,

$X_6=\begin{pmatrix}4.1146368\\0.6560896\\0.4292736\end{pmatrix}$, $X_7=\begin{pmatrix}4.11414528\\0.65756416\\0.42829056\end{pmatrix}$, ……and so on, which would obviously converge to the best approximate solution.

2. If $AX=b$(13) has a unique solution, then the n square matrix A should be non-singular and $AA=I$. It is thus evident that the form (3) is only special case of (14) for both of them are identical.

3. The formulas (16) and (38) also can be used to compute the approximate solution sequence of an overdetermined system of linear equations of the form $AX=b$(1).

Example 2.

$$x_1+2x_2=4,$$
$$2x_1-x_2=5,$$
$$x_1-2x_2=2,$$

where $A=\begin{pmatrix}1&2\\2&-1\\1&-2\end{pmatrix}$ and $b=\begin{pmatrix}4\\5\\2\end{pmatrix}$.

1. The best approximate solution $X'=A^-b=\begin{pmatrix}2.84\\0.52\end{pmatrix}$.

2. The least squares solution $X=(A'A)^{-1}A'b=\begin{pmatrix}2.84\\0.52\end{pmatrix}$.

3. The minimax solution $X=\begin{pmatrix}2.866667\\0.500000\end{pmatrix}$.

4. Let $A=FR'$, where $F=\begin{pmatrix}1&2\\2&-1\\1&-2\end{pmatrix}$ and $R'=\begin{pmatrix}1&0\\0&1\end{pmatrix}$ and $AA^-=F(F'F)^{-1}F'=\frac{1}{50}\begin{pmatrix}41&12&-15\\12&34&20\\-15&20&25\end{pmatrix}$. We shall find the approximate inverse $B=\frac{A'}{c}=\frac{1}{10}\begin{pmatrix}1&2&1\\2&-1&-2\end{pmatrix}$ by choosing $c=10$, $E=AA^--AB=\frac{1}{50}\begin{pmatrix}16&12&0\\12&9&0\\0&0&0\end{pmatrix}$, and the initial approximate solution $X_0=\begin{pmatrix}1.6\\-0.1\end{pmatrix}$. The formula $X_n=X_{n-1}+BE^nb$ (38) can be also used recursively to obtain

$X_1=\begin{pmatrix}2.22\\0.21\end{pmatrix}$, $X_2=\begin{pmatrix}2.530\\0.365\end{pmatrix}$, $X_3=\begin{pmatrix}2.6850\\0.4425\end{pmatrix}$, $X_4=\begin{pmatrix}2.76250\\0.48125\end{pmatrix}$, $X_5=\begin{pmatrix}2.801250\\0.500625\end{pmatrix}$, $X_6=\begin{pmatrix}2.8206250\\0.5103125\end{pmatrix}$, $X_7=\begin{pmatrix}2.83031250\\0.51515625\end{pmatrix}$, $X_8=\begin{pmatrix}2.835156250\\0.517578125\end{pmatrix}$, $X_9=\begin{pmatrix}2.8375781250\\0.5187890625\end{pmatrix}$, $X_{10}=\begin{pmatrix}2.83878906250\\0.51939453125\end{pmatrix}$, ………, and so on, which would converge to the same solutions as indicated above.

REFERENCES

1. Graybill, F. A., Introduction to matrices with applications in statistics, Wordsworth Publishing company, Inc., Belmont, Calif., 1969, 372P.

2. Scheid, F., Theory and problems of Numerical analysis, Schaum Publishing Co., N. Y. 1968, 422P.

3. Penrose, R., On best approximate solution of linear matrix equations, Proc., Cambridge Philos. Soci., 52, 1956, PP 17-19.

4. Greville, T. N. E., The pseudoinverse of a rectangular or singular matrix and its application to the solution of system of linear equations, SIAM Review Vol. 1, No. 1, 1959, PP38-43.

5. Fröberg, C. E., Introduction to numerical analysis, This volumn is an English translation of lärobok i Numerisk Analys by C. E. Fröberg, Published and sold by permission of Svenska Bokförlaget Bonniers, 2nd. Edi., 1969, 433P.

6. Householder, A. S., Theory of matrices in numerical analysis, Blaisdell, N. Y., 2nd. Edi., 1965, 257P.

7. Ben-Israel, Adi., An iterative method for computing the generalized inverse of an arbitrary matrix, Math. Comp., Vol. 10., 1965, PP452-455.