Journal of Computational Mathematics, Vol.14, No.4, 1996, 301–310.

A COLUMN RECURRENCE ALGORITHM FOR SOLVING LINEAR LEAST SQUARES PROBLEM[*1)]

J.X. Zhao
(*Department of Mathematics, Nanjing University, Nanjing, China*)

Abstract

A new column recurrence algorithm based on the classical Greville method and modified Huang update is proposed for computing generalized inverse matrix and least squares solution. The numerical results have shown the high efficiency and stability of the algorithm.

1. Introduction

Numerical method of a generalized inverse matrix and corresponding with the linear least squares is a standard tool for solving such problems as control, state evaluation and identification. Let A be an $m \times n$ real matrix. A real $n \times m$ matrix G is called the M-P generalized inverse matrix of A if G satisfies the following conditions:

$$\begin{aligned} \text{(I)} \quad AGA &= A, \quad \text{(II)} \quad GAG = G, \\ \text{(III)} \quad (AG)^T &= AG, \quad \text{(IV)} \quad (GA)^T = GA. \end{aligned} \tag{1}$$

Usually, we write G

$$A^+ = G.$$

The linear least square problem is defined as the minimization of the norm of the residual vector

$$\min_x \|Ax - b\|_2^2, \tag{2}$$

where b is an m-vector and x is an n-vector. Thus, the least square solution of the minimum norm of problem (2) is

$$x = A^+ b.$$

One of the major stability indices for computing generalized inverse matrix or linear least squares problem is

$$\kappa(A) = \|A\|\|A^+\|.$$

[*] Received May 23, 1994.

[1)] On leave at the Department of Mathematics, University of Trento, 38050 POVO (Trento), Italy.

An ill-conditioned matrix A, i.e. matrix with large κ, is quite common in control and identification problem[5]. It is thus important to have computational procedures suitable for solving ill-conditioned problem. An excellent survey on linear least square has been given in Björck[1].

Since the Greville scheme[6] is relative simple and is called G-method, it is adopted frequently for computing generalized inverse matrix in some cases. Computational practice and theoretical analysis show, however, when A is an ill-conditioned matrix, that the solution computed by G-method may bear no resemblance to the true solution. On the other hand, modified Huang method, one of the ABS class, may be more stable than that of some classical matrix factorization method[2]. But this method is only fit for solving the problem where $m \leq n$.

Our aim of this paper is to describe a new modification of the classical Greville method, which retains the main advantages of the classical scheme but in many cases is more stable.

Throughout this paper, let $\|\cdot\|$ stand for the 2-norm of a matrix or a vector.

2. Greville Method and Its Modification

Let A be a matrix of order $m \times n$ and will be denoted by

$$A = [a_1, a_2, \cdots, a_n] \in R^{m\times n},$$

where $a_i \in R^n$ and $m \geq n$. By convention, we assume rank $(A) = n$.

Denoted by

$$A_1 = [a_1], \quad A_k = [A_{k-1}, a_k] \in R^{m\times k}, \quad k \leq n. \tag{3}$$

We have known that G-method is an electable method for computing generalized inverse matrix if A is not too ill-conditioned. The G-method proceeds as follows[6].

G-method:

Set

$$A_1^+ = a_1^T/(a_1^T a_1).$$

For k=2:n Compute

$$d_k = A_{k-1}^+ a_k, \quad c_k = a_k - A_{k-1} d_k;$$

Take

$$y_k^T = \begin{cases} c_k^+ & c_k \neq 0 \\ (1 + d_k^T d_k) d_k^T A_{k-1}^+ & c_k = 0\ . \end{cases}$$

Compute

$$A_k^+ = \begin{bmatrix} A_{k-1}^+ - d_k y_k^T \\ y_k^T \end{bmatrix}.$$

Set

$$A^+ = A_n^+.$$

This algorithmic scheme is known as the Greville recurrence procedure. It is easy to see that the scheme is very compact and the numerical result is acceptable to well-behavior matrix.

Denote by $R(A)$ and $R(A)^{\perp}$ the range of A and corresponding orthogonal complement, respectively. Since

$$\begin{aligned} c_k &= a_k - A_{k-1}d_k \\ &= (I_m - A_{k-1}A_{k-1}^{+})a_k, \end{aligned} \tag{4}$$

c_k is an orthogonal projection vector of a_k onto $R(A_{k-1})^{\perp}$. The following lemma thus can be obtained directly[6].

Lemma. *The orthogonal projection $c_k = 0$ in G-method iff a_k is a linear combination of the vector system $a_1, a_2, \cdots, a_{k-1}$, i.e.*

$$a_k \in R(A_{k-1}).$$

In [4] the relation between the G-method and classical Gram-schmidt process is analyzed and the G-method for computing M-P inverse matrix, when $\operatorname{rank}(A) = n \le m$, is entirely equivalent to Gram-schmidt process from a purely theoretical viewpoint. In other words, we can prove the following relations to G-method[4].

Theorem. *If $c_1, c_2, \cdots, c_n$ are obtained by G-method, then*

1. $c_k^T a_j = c_k^T c_j = 0, \quad k = 1, 2, \cdots, n; \quad j = 1, 2, \cdots, k-1;$
2. $c_i^T a_i = c_i^T c_i, \quad i = 1, 2, \cdots, n;$
3. $(I_m - \Sigma_{l=1}^{k-1} c_l c_l^{+})a_j = (I_m - \Sigma_{l=1}^{k-1} c_l c_l^{+})c_j = 0, \quad j = 1, 2, \cdots, k.$

Proof. By induction.

From the above discussion, we have known that G-method is equivalent to classical Gram-schmidt process. Therefore, its numerical stability is relatively weak for ill-conditioned problem. In general, the modified Gram-schmidt process is more stable than that of classical scheme. Based on the analogous idea, we can obtain the following modified Greville method.

Take

$$A_1^{+} = a_1^{+},$$

Compute

$$d_2^{(j)} = A_1^{+} a_j, \quad j = 2, 3, \cdots, n.$$

Let

$$d_2 = d_2^{(2)}.$$

Then using the relations

$$d_3^{(j)} = A_2^{+} a_j = \begin{bmatrix} A_1^{+} - d_2 y_2^T \\ y_2^T \end{bmatrix} a_j$$

$$= \begin{bmatrix} d_2^{(j)} - (y_2^T a_j) d_2 \\ y_2^T a_j \end{bmatrix},$$

we can compute $d_3^{(j)}, j = 3, 4, \cdots, n$.

Set

$$d_3 = d_3^{(3)} = A_2^+ a_3.$$

$$\cdots\cdots$$

At the beginning of k-th ($k \leq n$) step, assume that $d_{k-1}^{(j)}, j = k, k+1, \cdots, n$, have already been computed. Then compute

$$\begin{aligned} d_k^{(j)} &= A_{k-1}^+ a_j = \begin{bmatrix} d_{k-1}^{(j)} - (y_{k-1}^T a_j) d_{k-1} \\ y_{k-1}^T a_j \end{bmatrix}, \\ j &= k, k+1, \cdots, n. \end{aligned} \tag{5}$$

Set

$$d_k = d_k^{(k)} = A_{k-1}^+ a_k.$$

We can see that the vector system $c_1, c_2, \cdots, c_k, k \leq n$, obtained by G-method must be in $R(A_k)$. Therefore,

$$A_k A_k^+ c_i = c_i \quad i = 1, 2, \cdots, k. \tag{6}$$

This means that the eigenvectors of $A_k A_k^+$ associated with eigenvalue 1 are $c_i, i = 1, 2, \cdots, k$. It follows from the spectral decomposition that

$$A_k A_k^+ = c_1 c_1^+ + c_2 c_2^+ + \cdots + c_k c_k^+.$$

Let

$$\begin{aligned} c_k^{(j)} &= (I_m - A_k A_{k-1}^+) a_j \\ &= (I_m - \Sigma_{l=1}^{k-1} c_l c_l^+) a_j \\ &= c_{k-1}^{(j)} - c_{k-1} c_{k-1}^+ a_j \\ &= c_{k-1}^{(j)} - \frac{(c_{k-1}^T a_j) c_{k-1}}{c_{k-1}^T c_{k-1}} \quad j = k, k+1, \cdots, n. \end{aligned} \tag{7}$$

Take

$$c_k = c_k^{(k)}.$$

Based on the above discussion, we can lead the modified Greville method, called the MG-method, which is equivalent to modified Gram-schmidt process. To ensure the numerical stability of G-method, a pivoting strategy is often performed in the size of $\|c_k\|$.

MG-method:

Compute

$$A_1^+$$

For $k = 2 : n$,

1. Compute

$$d_k^{(j)} = \begin{bmatrix} d_{k-1}^{(j)} - (y_{k-1}^T a_j) d_{k-1} \\ y_{k-1}^T a_j \end{bmatrix}, j = k, k+1, \cdots, n.$$

2. Compute

$$c_k^{(j)} = c_{k-1}^{(j)} - \frac{(c_{k-1}^T a_j) c_{k-1}}{c_{k-1}^T c_{k-1}}, j = k, k+1, \cdots, n.$$

Pivoting strategies:

$$c_{max} = \|c_k^{(j_0)}\| = \max_{k \le j \le n} (\|c_k^{(j)}\|)$$

Interchanging:

$$a_{j_0} \leftrightarrow a_k, \quad c_k^{(j_0)} \leftrightarrow c_k^{(k)}, \quad d_k^{(j_0)} \leftrightarrow d_k^{(k)}$$

Let

$$c_k = c_k^{(k)}, \quad d_k = d_k^{(k)}.$$

3. Take

$$y_k^T = c_k^+.$$

4. Compute

$$A_k^+ = \begin{bmatrix} A_{k-1}^+ - d_k y_k^T \\ y_k^T \end{bmatrix}.$$

Set

$$A^+ = A_n^+.$$

Furthermore, the numerical experiments show that the updating precision of $c_k, k = 1, 2, \cdots, n$, makes great influence on the numerical stability in the G-method. On the other hand, the ABS methods for proceeding successively to orthogonal bases of $R(A)$ have been obtained for the case where $m \le n$ [2].

ABS-updating:

Let H_1 be an arbitrary $n \times n$ nonsingular matrix. For i=1,2,$\cdots$,n, compute

(i) $p_i = H_i z_i$,

where $z_i \in R^n$ is arbitrary, subject to $z_i^T p_i \neq 0$;

(ii) $H_{i+1} = H_i - H_i w_i a_i^T H_i / (a_i^T H_i w_i)$,

where $w_i \in R^n$ is arbitrary, subject to $a_i^T H_i w_i \neq 0$.

The vector system $p_1, p_2, \cdots, p_k$ obtained by this algorithm satisfy

$$p_k^T a_j = 0, j = 1, 2, \cdots, k-1.$$

Among the particular algorithms in the ABS class obtained by making specific choices of the parameters, the modified Huang method is of special interest:

$H_1 = I_n$.

For $i = 1, 2, \cdots, n$. do

$$\begin{aligned} z_i &= H_i a_i, w_i = z_i / z_i^T z_i, \\ p_i &= H_i^T z_i, \\ H_{i+1} &= H_i - H_i a_i w_i^T H_i. \end{aligned}$$

This algorithm is numerically more stable than that of the Huang method or its analogues for the determined or underdetermined problem[2].

Here we wish to extend the modified Huang updating to overdetermined problem. It is to say that the orthogonal projection vector system $c_1, c_2, \cdots, c_n$, can be obtained by the modified Huang updating in the MG-method. Based on this idea, we can comprise a numerical method, called MHG-method, for computing the M-P generalized inverse.

MHG-method:

1. Set $H_1 = I_m$. and compute A_1^+,
2. For $k = 2 : n$ compute

2.1

$$\begin{aligned} d_k^{(j)} &= \begin{bmatrix} d_{k-1}^{(j)} - (y_k^T a_j) d_{k-1} \\ y_k^T a_j \end{bmatrix}, j = k, k+1, \cdots, n, \\ d_k &= d_k^{(k)}. \end{aligned}$$

2.2

$$\begin{aligned} z_k &= H_k a_k, \quad w_k = z_k^T z_k, \\ c_k &= H_k^T z_k, \\ H_{k+1} &= H_k - H_k a_k w_k^T H_k. \end{aligned}$$

2.3

$$A_k^+ = \begin{bmatrix} A_{k-1}^+ - d_k c_k^+ \\ c_k^+ \end{bmatrix}.$$

3. Set

$$A^+ = A_n^+.$$

In many cases, above the algorithm is much stabler than that of the G-method or MG-method for ill-conditioned problems.

3. Solution of Linear Least Squares Problem

Applying the above given MHG-method for computing the generalized inverse, we here can directly extend to solving overdetermined linear least squares problem

$$\min_x \|Ax - b\|^2, \tag{8}$$

where A is an $m \times n$ matrix,$m \geq n$, and b is an m-vector.

Assume that A_{k-1}^{+} has been computed at the beginning of k-th step. Then we can form

$$x_{k-1} = A_{k-1}^{+} b$$

and

$$A_k = [A_{k-1}, a_k].$$

Therefore, we have

$$\begin{aligned} x_k &= A_k^{+} b = \begin{bmatrix} A_{k-1}^{+} b - (y_k^T b) d_k \\ y_k^T b \end{bmatrix} \\ &= \begin{bmatrix} x_{k-1} - x_k^{'} d_k \\ x_k^{'} \end{bmatrix}, \end{aligned}$$

where

$$x_k^{'} = y_k^T b$$

and y_k, d_k are computed by the MHG-method.

To be more compact, the new algorithm takes the vector b in the right hand side of the above equality as (n+1)th column of A. Based on the above idea, the following algorithm has been given, called the MHGS-method, for solving the linear least squares problem.

MHGS-method:

1. Set $[A \vdots b] = [a_1, a_2, \cdots, , a_{n+1}] \in R^{m \times (n+1)}, H_1 = I_m,\ R = [r_{ij}] = 0 \in R^{n \times (n+1)}$, a working array of the upper triangular matrix;

2. For $i = 1, 2, \cdots, n$ do

$z_i = H_i a_i,$

$w_i = z_i / z_i^T z_i,$

$c_i = H_i^T z_i$

If i=1 then $r_{1j} = c_1^T a_j / c_1^T c_1, \quad j = 2, \cdots, n+1;$

else for $j = i+1, \cdots, n+1$ do

for $k = 1, 2, \cdots, i-1$ do

$r_{kj} = r_{kj} - \frac{c_i^T a_j}{c_i^T c_i} r_{ki},$

$r_{ij} = \frac{c_i^T a_i}{c_i^T c_i}.$

$H_{i+1} = H_i - H_i a_i w_i^T H_i.$

3. $x_i = r_{i.n+1}, i = 1, 2, \cdots, n.$

$x = (x_1, x_2, \cdots, x_n)^T$ is the least squares solution of the minimum norm of (8).

4. Numerical Experiments

The algorithm described in previous sections has been implemented and has been run in double precision on VAX computer of the University of Trento. The considered linear least squares problems,

$$\min_x \|Ax - b\|^2,$$

have the following data.

Problem 1. (ill-conditioned problem)

$$\begin{aligned} A &= [a_{ij}] = [1.0/(i+j-1.0)] \in R^{m\times n}, \\ b &= (b_1, b_2, \cdots, b_m)^T \in R^m, \\ b_i &= \Sigma_{j=1}^n a_{ij}, \quad i = 1, 2, \cdots, m. \end{aligned}$$

Problem 2. (mid-conditioned problem)

$$\begin{aligned} A &= [a_{ij}] = [\max(i,j)] \in R^{m\times n} \\ b &= (b_1, b_2, \cdots, b_m)^T \in R^m, \\ b_i &= \Sigma_{j=1}^n a_{ij}, \quad i = 1, 2, \cdots, m. \end{aligned}$$

Problem 3. (mid-conditioned problem)

$$\begin{aligned} A &= \begin{bmatrix} n & n-1 & n-2 & \dots & 2 & 1 \\ n-1 & n-1 & n-2 & \dots & 2 & 1 \\ n-2 & n-2 & n-2 & \dots & 2 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 2 & 2 & 2 & \dots & 2 & 1 \\ 1 & 1 & 1 & \dots & 1 & 1 \end{bmatrix} = [a_{ij}] \in R^{n\times n}, \\ b &= (b_1, b_2, \cdots, b_n)^T \in R^n, \\ b_i &= \Sigma_{j=1}^n a_{ij}, \quad i = 1, 2, \cdots, n. \end{aligned}$$

Algorithm:
G-method: Column Greville method,
CGLSP[2)]-method : Algorithm given in [5],

[2)] Modified conjugate gradient method for least squares problem

MHGS-method: Algorithm obtained in this paper.

The following tables present the computed precision with three different algorithms for the given problems. These results indicate that the new method (MHGS-method) obtained in this paper is fairly efficient, especially as an ill-conditioned least square as problem 1. The index of the relative accuracy is defined by

$$P = \frac{\|x - x^*\|_2}{\|x^*\|_2},$$

where x^* is the true solution and x is a computed solution.

The condition numbers of the given matrices, cond(A), are computed by MATLAB.

Table 1. The relative accuracy P of Problem 1 ($m = n$).

m	Cond(A)	G-method	CGLSP-method	MHGS-method
5	4.7661E+05	1.1666822E-07	2.9090133E-05	2.1568097E-12
10	1.6025E+13	>1	3.7835974E-04	6.1374327E-09
15	3.3634E+17	>1	1.1718496E-03	7.3047523E-09
20	4.0020E+19	>>1	5.2199944E-04	2.4599253E-08
25	1.7138E+18	>>1	8.5164341E-04	1.0516242E-08
30	4.5397E+18	>>1	1.2079764E-03	2.2723464E-08
35	5.8601E+18	>>1	1.5771304E-03	2.0508478E-08
40	7.2654E+18	>>1	6.1066032E-04	5.0091549E-08

Table 2. The relative accuracy P of Problem 2 ($m = n$).

m	Cond(A)	G-method	CGLSP-method	MHGS-method
5	7.0766E+01	7.3474726E-14	4.7808928E-16	2.5225527E-16
10	2.8919E+02	1.7468340E-11	9.9344994E-16	3.2823535E-15
15	6.4639E+02	1.8298022E-07	1.4754814E-15	6.2574871E-15
20	1.1425E+03	3.3819112E-04	4.3725890E-15	1.5046502E-14
25	1.7775E+03	0.2837466E+00	5.6821201E-15	1.9495403E-14
30	2.5515E+03	>1	9.2010109E-15	2.2474395E-14
35	3.4644E+03	>1	1.1894571E-14	4.6867962E-14
40	4.5163E+03	>1	1.6454910E-14	5.3042908E-14

Table 3. The relative accuracy P of Problem 3 ($m = n$).

m	Cond(A)	G-method	CGLSP-method	MHGS-method
5	4.5455E+01	0.0000000E+00	1.4057041E-16	0.0000000E+00
10	1.7508E+02	3.6570011E-16	3.3419257E-16	0.0000000E+00
15	3.8582E+02	4.1579343E-15	6.8814798E-16	0.0000000E+00
20	6.7762E+02	1.4859172E-14	1.6301262E-15	0.0000000E+00
25	1.0505E+03	1.6141173E-14	3.4069208E-15	0.0000000E+00
30	1.5044E+03	3.4304422E-14	3.9338318E-15	0.0000000E+00
35	2.0394E+03	2.7719791E-14	8.5910155E-15	0.0000000E+00
40	2.6554E+03	4.8073519E-14	9.9144612E-15	0.0000000E+00

Table 4. The relative accuracy of the MHGS-method
for the ill-conditioned problem 1 ($m > n$).

m	n	Cond(A)	relative accuracy P
150	100	1.5173E+18	3.3504126E-08
150	110	2.1010E+18	4.0557843E-08
150	120	2.5797E+18	4.6187279E-08
150	130	4.0761E+18	5.2436966E-08
150	140	1.0964E+19	9.6172765E-08
150	150	1.9002E+20	2.0729776E-07
200	150	2.7419E+18	4.8961957E-08
500	10	9.9475E+09	1.6412854E-09
500	100	6.1819E+17	3.7023077E-08

References

[1] Å. Björck, Least squares methods, in Handbook of Numerical Analysis. Volume 1: Solution of Equations in R^n, P.G. Ciarlet and J.L. Lions, eds., Elsevier/North Holland, Amsterdam, (1990)

[2] J. Abaffy and E. Spedicato, ABS projection algorithms: Mathematical techniques for linear and nonlinear equations, Ellis Horwood (1989).

[3] J. Zhao, Huang's method for solution of consistent linear equations and its generalization, *J. Numer. Math. of Chinese Univer.*, 3(1981).

[4] J. Zhao, The study of recurrence methods for solving ill-conditioned linear system, Ph.D. thesis, Nanjing University, China, (1987).

[5] A.M. Vigodner and A.A. Pervozvanski, A modification of the conjugate gradient algorithm for the stable solution of an ill-conditioned problem by the method of least squares, *Comput. Math. Phys.*, 31(1991).

[6] A. Ben-Israel and T.N.E. Greville, Generalized inverse: Theory and Applications, New York, (1974)