ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
20,4 (1990), pp. 587–589

H. SPÄTH (Oldenburg)

MODIFIED GRAM–SCHMIDT FOR SOLVING LINEAR LEAST SQUARES PROBLEMS IS EQUIVALENT TO GAUSSIAN ELIMINATION FOR THE NORMAL EQUATIONS

Abstract. Gaussian elimination for the normal equations written in a suitable way is algebraically and numerically equivalent to the modified Gram–Schmidt process simultaneously applied to both the given coefficient matrix and the right-hand side.

Let A be a real $(m \times n)$-matrix with $m \geqslant n$ and $b \in \boldsymbol{R}^m$. Solving the overdetermined linear system of equations

$$Ax = b$$

for $x \in \boldsymbol{R}^n$ in the least squares sense means to determine x such that

$$\|Ax - b\|_2 \to \min.$$

By assuming $\operatorname{rank}(A) = n$, x is unique and can be found by solving the so-called normal equations

$$A^T A x = A^T b. \tag{1}$$

In most situations it is recommended for a numerical solution of the problem not to use (1) but to apply an orthogonalization method for A (and b) based on Householder transformations or the modified Gram–Schmidt (MGS) method ([1], [2]). MGS and Householder programs have found to be of essentially equivalent accuracy [2]. In this note we show that the Gaussian elimination for (1) written in a suitable way is algebraically and numerically equivalent to MGS.

Denoting the columns of A by $a_j (j = 1, \ldots, n)$ we can write equation (1) as

$$Cx = d, \tag{2}$$

where

$$\begin{aligned} c_{ij} &= a_i^T a_j \quad (i, j = 1, \ldots, n), \\ d_i &= a_i^T b \quad (i = 1, \ldots, n). \end{aligned} \tag{3}$$

Starting the Gaussian elimination for (2), using (3) we get at first

$$x_1 = g_1 - \sum_{j=2}^{n} r_{1j} x_j, \tag{4}$$

where

$$f_1 = \|a_1\|_2^2, \quad g_1 = a_1^T b/f_1, \quad r_{1j} = a_1^T a_j/f_1 \quad (j = 2, \ldots, n).$$

Substituting x_1 into (2) and omitting the first equation (4) we get

$$C^{(1)} \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} = d^{(1)}, \tag{5}$$

where the $((n-1)\times(n-1)$-matrix $C^{(1)}$ and $d^{(1)} \in \boldsymbol{R}^{n-1}$ are given by

$$\begin{aligned} c_{i-1,j-1}^{(1)} &= a_i^T a_j - a_1^T a_i \frac{a_1^T a_j}{a_1^T a_1} \quad (i, j = 2, \ldots, n), \\ d_{i-1}^{(1)} &= a_i^T b - a_1^T a_i \frac{a_1^T b}{a_1^T a_1} \quad (i = 2, \ldots, n). \end{aligned} \tag{6}$$

The easily verified key observation is now that (6) can be written as

$$\begin{aligned} c_{i-1,j-1}^{(1)} &= (a_i - r_{1i} a_1)^T (a_j - r_{1j} a_1) \quad (i, j = 2, \ldots, n), \\ d_{i-1}^{(1)} &= (a_i - r_{1i} a_1)^T (b - g_1 a_1) \quad (i = 2, \ldots, n) \end{aligned}$$

or, putting

$$\begin{aligned} a_j^{(1)} &= a_j^0 - r_{1j} a_1^{(0)} \quad (j = 2, \ldots, n), \\ b^{(1)} &= b^{(0)} - g_1 a_1^{(0)}, \end{aligned}$$

where $a_j^{(0)} = a_j$ $(j = 1, \ldots, n)$ and $b^{(0)} = b$, finally in the form

$$\begin{aligned} c_{i-1,j-1}^{(1)} &= (a_i^{(1)})^T a_j^{(1)} \quad (i, j = 2, \ldots, n), \\ d_{i-1}^{(1)} &= (a_i^{(1)})^T b^{(1)} \quad (i = 2, \ldots, n). \end{aligned} \tag{7}$$

But (7) means that (5) has the same structure as (2). Thus evidently the k-th elimination step $(k = 1, \ldots, n)$ can now be described in the following way:

$$x_k = g_k - \sum_{j=k+1}^{n} r_{kj} x_j \tag{8}$$

with

$$\begin{aligned} f_k &= \|a_k^{(k-1)}\|_2^2, \quad g_k = (a_k^{(k-1)})^T b^{(k-1)}/f_k, \\ r_{kj} &= (a_k^{(k-1)})^T a_j^{(k-1)}/f_k \quad (j = k+1, \ldots, n), \\ a_j^{(k)} &= a_j^{(k-1)} - r_{kj} a_k^{(k-1)} \quad (j = k+1, \ldots, n), \\ b^{(k)} &= b^{(k-1)} - g_k a_k^{(k-1)}. \end{aligned} \tag{9}$$

For the remaining linear system of equations with $n-k$ unknowns $x_{k+1}, \ldots, x_n$ and $n-k$ equations we would have

$$C^{(k)} \begin{bmatrix} x_{k+1} \\ \vdots \\ x_n \end{bmatrix} = d^{(k)}$$

with

$$\begin{aligned} c^{(k)}_{i-k,j-k} &= (a^{(k)}_i)^T a^{(k)}_j \qquad (i, j = k+1, \ldots, n), \\ d^{(k)}_{i-k} &= (a^{(k)}_i)^T b^{(k)} \qquad (i = k+1, \ldots, n). \end{aligned} \tag{10}$$

But as in (8) only values of (9) are required it turns out that the coefficient matrix and the right-hand side neither in (3) nor in (10) have to be calculated explicitly for the Gaussian elimination.

On the other hand, inspecting (9), you see the well-known modified Gram–Schmidt orthogonalization (including the right-hand side as required) and (8) is the corresponding backsubstitution ([1]–[3]).

References

[1] G. H. Golub and C. F. van Loan, *Matrix Computations*, North Oxford Academy, Baltimore 1983.

[2] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, Prentice Hall, Englewood Cliffs 1974.

[3] J. W. Longley, *Least Squares Computations Using Orthogonalization Methods*, Marcel Dekker, New York 1984.

PROF. DR. HELMUTH SPÄTH
FACHBEREICH MATHEMATIK
UNIVERSITÄT OLDENBURG
POSTFACH 2503
D-2900 OLDENBURG, FRG

Received on 1989.01.31