ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
VIII, 4 (1966)

K. FLOREK (Wrocław)

SOME REMARKS ON SOLVING SYSTEMS OF LINEAR EQUATIONS BY A RELAXATION METHOD COMBINED WITH A GENERALIZED HOTELLING METHOD FOR INVERTING MATRICES

Let A, B, C, D_n, and H_n be square matrices of degree r, let I denotes the unit matrix, and $|A|$ the determinant of a matrix A. Let P, Q, X_k, and Y be column vectors of degree r. Let $\|A\|$ and $\|P\|$ denote a metric for the matrix A and for the vector P respectively, such that: for any real number a we have $\|Aa\| = |a|\cdot\|A\|$; $\|P\| > 0$ for $P \neq 0$; $\|I\| = 1$; and for any A, B, C and P, Q the inequalities

$$\|AB+C\| \leqslant \|A\|\cdot\|B\|+\|C\|$$

and

$$\|AP+Q\| \leqslant \|A\|\cdot\|P\|+\|Q\|$$

are fulfilled.

A generalized Hotelling method for inverting matrices ([1], [3]) allows for fixed $m = 2, 3, 4, \ldots$; for any matrix A satisfying the condition

(1) $$|A| \neq 0;$$

and for any given D_0, the determination of

(2) $$D_n = (I-H_0^{m^n})A^{-1}$$

using the algorithm

$$D_{n+1} = \Big(I+\sum_{s=1}^{m-1} H_n^s\Big) D_n, \qquad n = 0, 1, 2, \ldots,$$

where

$$H_n = I-D_nA.$$

Moreover, if

(3) $$\|H_0\| \leqslant \theta < 1,$$

then the following inequality holds:

(4) $$\|D_n-A^{-1}\| \leqslant \frac{\theta^{m^n}\|D_0\|}{1-\theta}.$$

For $m = 2$ the procedure (2) reduces to the well-known Hotelling method [2].

If μ_r is the time necessary to perform the multiplication of two square matrices of degree r, and σ the addition or subtraction time of two elements, then the necessary time for the calculation of D_n is equal to

$$\tau(m, n) = nm(\mu_r + r\sigma). \tag{5}$$

Having given the required precision of the result, which according to (4) means a fixed m^n, the time (5) reaches its minimum for $m = 3$. For this value of m the generalized Hotelling method is faster than the original one. But its disadvantage lies in the necessity (for $m > 2$) of storeing one matrix more in the course of calculation.

Consider now a system of linear equations

$$AX = Y. \tag{6}$$

This system may be solved by a combination of the generalized Hotelling method with the relaxation method

$$X_k = DY + (I - DA)X_{k-1}, \quad k = 1, 2, 3, \ldots \tag{7}$$

The following theorem gives an estimate for the precision of the proposed combination of methods:

THEOREM 1. *If, for fixed* $m \geqslant 2$, *the relations* (1) *and* (3) *are fulfilled, and if* $D = D_n$ *is given by* (2), *then the procedure* (7) *leads to an approximate solution of* (6) *and*

$$\|X_k - X\| \leqslant \theta^{km^n} \|X_0 - X\|, \tag{8}$$

where X_0 *is the initial solution.*

Proof. Putting (6) into (7) we have

$$X_k - X = (I - DA)(X_{k-1} - X).$$

Expressing $X_{k-1} - X$ by $X_{k-2} - X$ and so on, in k steps we come to

$$X_k - X = (I - DA)^k (X_0 - X).$$

For $D = D_n$ we have from (2)

$$X_k - X = H_0^{km^n}(X_0 - X).$$

Inequality (3) allows then the formulation of the theorem.

Optimal proportions for the combination of the two methods are stated in the following

THEOREM 2. *If the assumptions of theorem* 1 *are fulfilled, then for a given sufficiently large exponent* $\varrho > mr + 1$ *in the estimate*

$$\frac{\|X_k - X\|}{\|X_0 - X\|} \leqslant \theta^{km^n} \leqslant \theta^{\varrho} \tag{9}$$

the minimum calculation time of X_k is reached approximately for

(10) $$k \approx \frac{mr(\mu_r + r\sigma)}{\mu_r + r^2\sigma}$$

and

(11) $$n \approx \ln(\varrho/k)/\ln m,$$

where μ_r and σ have the same meaning as in (5).

Proof. According to (5) and (7) the necessary time for the calculation of X_k is given by

(12) $$\tau(m, n, k) = nm(\mu_r + r\sigma) + k(\mu_r + r^2\sigma) + \mu_r\left(1 + \frac{1}{\mu_r}\right) + r\sigma.$$

Using (3) and (9) we have for $\varrho > k$

(13) $$n \geqslant (\ln\varrho - \ln k)/\ln m.$$

If n was a real variable, the minimum calculation time would be obtained for the equality sign holding in (13). For integral n we get only the approximate equation (11). Let us put (11) into (12). Differentiation with regard to k (treated also as a real variable) gives

$$\frac{\partial}{\partial k}\tau(m, n, k) = -\frac{1}{k}m(\mu_r + r\sigma) + \frac{\mu_r}{r} + r\sigma.$$

This derivative is positive for

$$k > \gamma = \frac{mr(\mu_r + r\sigma)}{\mu_r + r^2\sigma},$$

and negative for $k < \gamma$; we have thus (10).

Remark. The time $\tau(m, n)$ of calculating D_n reaches its minimum for $m = 3$, but for $m = 2$ it is only about 5% greater. Thus searching for an optimal set of integers (m, n, k), the exponent ϱ given, it is necessary to consider at least two cases $m = 2$ and $m = 3$.

References

[1] M. Altman, *An optimum cubically convergent iterative method of inverting a linear bounded operator in Hilbert space*, Pacific Journal of Mathematics 10 (1960), pp. 1107-1113.

[2] H. Hotelling, *Some new methods in matrix calculation*, Annals of Mathematical Statistics 14 (1943), pp. 1-34.

[3] A. Kiełbasiński, *On the iterative procedures of best strategy for inverting a self-adjoint positive-definite bounded operator in Hilbert space*, Studia Mathematica 24 (1964), pp. 13-23.

INSTITUTE OF MATHEMATICS OF THE POLISH ACADEMY OF SCIENCES

Received on 22. 12. 1964.

K. FLOREK (Wrocław)

UWAGI O ROZWIĄZYWANIU RÓWNAŃ LINIOWYCH ZA POMOCĄ ZŁOŻENIA METODY RELAKSACJI I UOGÓLNIONEJ METODY HOTELLINGA ODWRACANIA MACIERZY

STRESZCZENIE

Rozważa się rozwiązywanie równania (5) za pomocą metody relaksacji (6), gdzie $D = D_n$ oblicza się według uogólnionej metody Hotellinga (2). Wzór (7) w tezie twierdzenia 1 podaje oszacowanie dokładności złożonej metody. Wzory (9) i (10) w tezie twierdzenia 2 podają przybliżone wartości parametrów k i n, które minimizują czas osiągnięcia rozwiązania X_k z daną dokładnością (8). We wzorze (9) μ_r oznacza czas mnożenia macierzy kwadratowych stopnia r, a σ czas dodawania dwu elementów macierzy.

К. ФЛЬОРЕК (Вроцлав)

ЗАМЕЧАНИЯ О РЕШЕНИИ ЛИНЕЙНЫХ УРАВНЕНИЙ С ПОМОЩЬЮ КОМБИНИРОВАНИЯ РЕЛАКСАЦИОННОГО МЕТОДА И ОБОБЩЁННОГО МЕТОДА ХОТЕЛЛИНГА ОБРАЩЕНИЯ МАТРИЦ

РЕЗЮМЕ

Рассматривается решение уравнения (5) методом релаксации (6), где $D = D_n$ вычисляется обобщённым методом Хотеллинга (2). Формула (7) в заключении теоремы 1 даёт оценку точности комбинированного метода. Формулы (9) и (10) в заключении теоремы 2 дают приближённые значения параметров k и n, которые минимизируют время получения решения X_k с данной точностью. В формуле (9) μ_r обозначает время умножения квадратных матриц r-го порядка, а σ — время сложения двух элементов матрицы.
