ON THE CONVERGENCE OF THE CLASSICAL ITERATIVE METHOD OF SOLVING LINEAR SIMULTANEOUS EQUATIONS[1]

By Edgar Reich

Massachusetts Institute of Technology

The classical iterative method, or Seidel method, is a scheme for solving the system of linear algebraic equations

$$\sum_{j=1}^{n} A_{ij} x_j = b_i, \qquad (i = 1, 2, \cdots, n),$$

by successive approximation, as follows:
If $x^{(\nu)} = (x_1^{(\nu)}, x_2^{(\nu)}, \cdots, x_n^{(\nu)})$ is the νth approximation of the solution, the $(\nu + 1)$st approximation, $x^{(\nu+1)} = (x_1^{(\nu+1)}, x_2^{(\nu+1)}, \cdots, x_n^{(\nu+1)})$, is obtained from the relations

$$\begin{cases} A_{11}x_1^{(\nu+1)} + A_{12}x_2^{(\nu)} + A_{13}x_3^{(\nu)} + \cdots + A_{1n}x_n^{(\nu)} = b_1, \\ A_{21}x_1^{(\nu+1)} + A_{22}x_2^{(\nu+1)} + A_{23}x_3^{(\nu)} + \cdots + A_{2n}x_n^{(\nu)} = b_2, \\ A_{31}x_1^{(\nu+1)} + A_{32}x_2^{(\nu+1)} + A_{33}x_3^{(\nu+1)} + \cdots + A_{3n}x_n^{(\nu)} = b_3, \\ \cdots \\ A_{n1}x_1^{(\nu+1)} + A_{n2}x_2^{(\nu+1)} + A_{n3}x_3^{(\nu+1)} + \cdots + A_{nn}x_n^{(\nu+1)} = b_n, \end{cases}$$

$x_1^{(\nu+1)}$ being obtained from the first equation, then $x_2^{(\nu+1)}$ from the second, and so on.

The given system can be written in matrix notation as $Ax = b$ where A is a non-singular square matrix of order n, and x and b are column vectors of order n. Let us define square matrices A_1 and A_2 as follows:

$$(A_1)_{ij} = \begin{cases} A_{ij} \text{ if } i \geq j \\ 0 \text{ if } i < j \end{cases},$$

$$(A_2)_{ij} = \begin{cases} A_{ij} \text{ if } i < j \\ 0 \text{ if } i \geq j \end{cases},$$

(Note that $A_1 + A_2 = A$.).

With this notation the Seidel method can be written as the matric difference equation

$$A_1 x^{(\nu+1)} + A_2 x^{(\nu)} = b.$$

Now various writers, among them C. E. Berry in this journal, (See list of refer-

[1] Work done under Office of Naval Research Contract N5ori60.

ences at end of this paper.) have shown that a necessary and sufficient condition for convergence, i.e., a necessary and sufficient condition for

$$\lim_{\nu \to \infty} (x_i^{(\nu)} - x_i) = 0, \qquad (i = 1, 2, \cdots, n),$$

is that

(1) A_1 has an inverse; that is $A_{ii} \neq 0$ for any i.

(2) The characteristic roots of $(A_1^{-1}A_2)$ all have an absolute value smaller than unity.

It would be advantageous to rephrase the above condition, if possible, in terms of simpler requirements on A. As a step in this direction the following theorem is offered:

THEOREM. *If A is a real, symmetric nth-order matrix with all terms on its main diagonal positive, then a necessary and sufficient condition for all the n characteristic roots of $(A_1^{-1}A_2)$ to be smaller than unity in magnitude is that A is positive definite.*

PROOF. Let z_j be a characteristic vector of $(A_1^{-1} A_2)$ corresponding to the characteristic root μ_j. Then

$$(1) \qquad (A_1^{-1}A_2)\, z_j = \mu_j z_j .$$

Premultiplying by $\bar{z}_i'A_1$, where the apostrophe and bar denote transposition and conjugation respectively:

$$(2) \qquad \bar{z}_i'A_2z_j = \mu_j\bar{z}_i'A_1z_j .$$

Consider the bilinear form $\bar{z}_i'Az_j$.

We have

$$(3) \qquad \bar{z}_i'Az_j = \bar{z}_i'A_1z_j + \bar{z}_i'A_2z_j = (1 + \mu_j)\, \bar{z}_i'A_1z_j .$$

Interchanging i and j:

$$(4) \qquad \bar{z}_j'Az_i = (1 + \mu_i)\bar{z}_j'A_1z_i .$$

Taking the conjugate:

$$(5) \qquad z_j'A\bar{z}_i = \bar{z}_i'Az_j = (1 + \bar{\mu}_i)z_j'A_1\bar{z}_i = (1 + \bar{\mu}_i)\bar{z}_i'A_1'z_j .$$

Let D be the diagonal matrix with elements

$$(6) \qquad D_{ij} = A_{ij}\delta_{ij} .$$

This makes $A_1' = D + A_2$.

Substituting this in (5):

$$(7) \quad \bar{z}_i'Az_j = (1 + \bar{\mu}_i)\, (\bar{z}_i'Dz_j + \bar{z}_i'A_2z_j) = (1 + \bar{\mu}_i)\bar{z}_i'Dz_j + (1 + \bar{\mu}_i)\mu_j\bar{z}_i'A_1z_j .$$

Eliminating $\bar{z}_i'A_1z_j$ between relations (3) and (7) we obtain

$$(8) \qquad (1 - \bar{\mu}_i\mu_j)\bar{z}_i'Az_j = (1 + \bar{\mu}_i)\, (1 + \mu_j)\bar{z}_i'Dz_j .$$

To obtain the necessary condition we use the fact that we must have $|\mu_i| < 1$, and can therefore rewrite (8) as

$$\bar{z}_i' A z_j = \frac{(1 + \bar{\mu}_i)(1 + \mu_j)}{1 - \bar{\mu}_i \mu_j} \bar{z}_i' D z_j = \sum_{k=0}^{\infty} (1 + \bar{\mu}_i)\bar{\mu}_i^k (1 + \mu_j)\mu_j^k \bar{z}_i' D z_j. \tag{9}$$

If $x = \sum_{i=1}^{m} c_i z_i$ is any linear combination of the $m \leq n$ independent characteristic vectors of $(A_1^{-1} A_2)$ then

$$\begin{aligned} \bar{x}' A x &= \left(\sum_{i=1}^{m} \bar{c}_i \bar{z}_i'\right) A \left(\sum_{i=1}^{m} c_i z_i\right) = \sum_{i,j=1}^{m} \bar{c}_i c_j \bar{z}_i' A z_j \\ &= \sum_{i,j=1}^{m} \bar{c}_i c_j \sum_{k=0}^{\infty} (1 + \bar{\mu}_i)\bar{\mu}_i^k (1 + \mu_j)\mu_j^k \bar{z}_i' D z_j, \end{aligned} \tag{10}$$

or

$$\bar{x}' A x = \sum_{k=0}^{\infty} \bar{y}_k' D y_k$$

where

$$y_k = \sum_{i=1}^{m} c_i (1 + \mu_i)\mu_i^k z_i.$$

Since by hypothesis $A_{ii} > 0$, D is evidently positive definite, and therefore

$$\bar{x}' A x > 0. \tag{11}$$

In case the characteristic roots μ_i, $(i = 1, 2, \cdots n)$, are all distinct there will be n independent z_i assured, and in that case (11) implies that A is positive definite. Consider, on the other hand, the case where the μ_i are not all distinct. Note that (a) the definiteness properties of a matrix are not changed by sufficiently small alterations in the elements; (b) the μ's depend continuously on the elements of A; (c) the discriminant of (1) is a polynomial in the A_{ij} that does not vanish identically.[2] It follows that A must be positive definite even in the case of repeated roots because an arbitrarily small change in A will separate any multiple μ's, still keeping them smaller than unity in magnitude, and not changing the definiteness properties of A.

This completes the proof that the condition given in the statement of the theorem is necessary. Now to prove sufficiency:

Setting $i = j$ in relation (8) we obtain

$$(1 - |\mu_i|^2)\bar{z}_i' A z_i = |1 + \mu_i|^2 \bar{z}_i' D z_i \tag{12}$$

Since both A and D are positive definite

$$\bar{z}_i' A z_i > 0 \text{ and } \bar{z}_i' D z_i > 0. \tag{13}$$

[2] The fact that the discriminant is not identically zero follows from easily constructible counter-examples.

Moreover, we cannot have $\mu_i = -1$ because that would mean by (3) that

$$0 = \bar{z}_i' A_1 z_i + \bar{z}_i' A_2 z_i = \bar{z}_i' A z_i .$$

Relation (12) thus implies

$$(14) \qquad 1 - |\mu_i|^2 > 0$$

i.e. $|\mu_i| < 1$ as was to be proved.

The part of the theorem giving the sufficient condition was already obtained by L. Seidel [1] and G. Temple in a somewhat more indirect fashion.

REFERENCES

[1] L. SEIDEL, "Über ein Verfahren die Gleichungen, auf welche die Methode der kleinsten Quadrate führt, sowie lineare Gleichungen überhaupt, durch successive Annä herung aufzulösen," *Abhandlungen der Mathematisch-Physikalischen Classe der Königlich Bayerischen Akademie der Wissenschaften*, Vol. 11 (1874), pp. 81–108.

[2] C. E. BERRY, "A criterion of convergence for the classical iterative method of solving linear simultaneous equations," *Annals of Math. Stat.*, Vol. 16 (1945), pp. 398–400.

[3] L. CESARI, "Sulla risoluzione dei sistemi di equazioni lineari per approssimazioni successive," *Rassegna delle Poste, dei Telegrafi e dei Telefoni*, Anno 9 (1931).

[4] L. CESARI, "Sulla risoluzione dei sistemi di equazioni lineari per approssimazioni successive," *Reale Accademia Nazionale dei Lincei, Serie 6, Classe di Szienze fisiche, matematiche e naturali, Rendiconti*, Vol. 25 (1937), pp. 422–428.

[5] J. MORRIS, *The Escalator Method in Engineering Vibration Problems*, Chapman and Hall Ltd., London 1947, pp. 63–70.

[6] R. J. SCHMIDT, "On the numerical solution of linear simultaneous equations by an iterative method," *Phil. Mag.*, Ser. 7, Vol. 32 (1941), pp. 369–383.

SOME RECURRENCE FORMULAE IN THE INCOMPLETE BETA FUNCTION RATIO

BY T. A. BANCROFT

Alabama Polytechnic Institute

1. Introduction. It is well known that the incomplete beta function ratio, defined by

$$(1) \qquad I_x(p, q) = \frac{B_x(p, q)}{B(p, q)},$$

where

$$(2) \qquad B_x(p, q) = \int_0^x x^{p-1}(1 - x)^{q-1}\, dx,$$

and

$$(3) \qquad B(p, q) = B_1(p, q),$$