Atti Accademia Nazionale dei Lincei

Classe Scienze Fisiche Matematiche Naturali

Rendiconti

William J. Kammerer, M. Zuhair Nashed

A generalization of a matrix iterative method of G. Cimmino to best approximate solution of linear integral equations of the first kind

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Serie 8, Vol. **51** (1971), n.1-2, p. 20–25.

Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLINA_1971_8_51_1-2_20_0>

Analisi matematica. — *A generalization of a matrix iterative method of G. Cimmino to best approximate solution of linear integral equations of the first kind.* Nota (*) di WILLIAM J. KAMMERER e M. ZUHAIR NASHED, presentata dal Corrisp. G. FICHERA.

RIASSUNTO. — Un metodo iterativo di G. Cimmino, relativo alle equazioni lineari algebriche, viene generalizzato ad una equazione integrale di prima specie: $\mathcal{K}x = y$. Ammettendo che il codominio $R(\mathcal{K})$ di $\mathcal{K}$ abbia dimensione maggiore di uno, si prova che il metodo converge ad un unico vettore che rende minima la norma di $\mathcal{K}x - y$ al variare di y nell'insieme $R(\mathcal{K}) + R(\mathcal{K})^{\perp}$.

1. G. Cimmino [1] proposed an iterative method for a finite system of linear algebraic equations which converges even if the system of equations is inconsistent, provided that the rank of the matrix is greater than one. It is the purpose of this note to generalize this method to linear integral equations of the first kind, and to announce a theorem establishing under mild conditions the convergence of the method to the solution of the integral equation if the equation has a unique solution, or to the solution of minimal norm if the equation has an infinite number of solutions. If the integral equation does not have a solution in the traditional sense, the method converges to a function which minimizes the L_2-norm of $\mathcal{K}x - y$, where $\mathcal{K}$ is the integral operator, and is of least L_2-norm.

2. In order to show that our proposed iterative method for least squares solutions of linear integral equations of the first kind may be viewed as a natural generalization of Cimmino's method for matrices, and to put the method itself in proper perspective, we highlight briefly the algebraic and geometric features of Cimmino's method.

Let $A = (a_{ij})$ be an $n \times n$ matrix over the field of complex numbers (the case of an $m \times n$ matrix can be incorporated by augmenting the matrix by the appropriate number of zero rows or columns). Let (u , v) denote the usual inner product in E^n, and let $r^{(1)}, \cdots, r^{(n)}$ denote the rows of the matrix A. These rows determine n hyperplanes in E^n defined for $x \in E^n$ by

$$\mathcal{H}_i = \{x : (r^{(i)}, x) = y_i\}, \qquad i = 1, \cdots, n.$$

Let $x^{(0)}$ be any selected vector. We place a mass $m_j > 0$ at the point which is symmetric to $x^{(0)}$ with respect to the hyperplane $\mathcal{H}_j$. We do this

(*) Pervenuta all'Accademia il 27 luglio 1971.

for each of the hyperplanes and we take the centroid of this system of n masses as our next iterate and continue the iteration in this manner using the same respective masses m_j. Algebraically, it is easy to show that this algorithm may be written in the form

$$x^{(k)} = \left(\sum_{j=1}^{n} m_j\right)^{-1} \sum_{i=1}^{n} m_i \left\{ x^{(k-1)} + 2\,\frac{y_i - (r^{(i)}, x^{(k-1)})}{(r^{(i)}, r^{(i)})}\, r^{(i)} \right\} =$$

$$= x^{(k-1)} - \frac{2}{\mu} \sum_{i=1}^{n} \left\{ \frac{(r^{(i)}, x^{(k-1)}) - y_i}{\|r^{(i)}\|^2} \right\} r^{(i)},$$

where $\mu = \sum_{i=1}^{n} m_i$. Setting $B = \left(\frac{\delta_{ij}}{\|r^{(i)}\|^2}\right)_{n\times n}$, $W = (\delta_{ij}\, m_i)_{n\times n}$, where δ_{ij} is the Kronecker delta, we obtain

$$x^{(k)} = x^{(k-1)} - \frac{2}{\mu} A^* WB\,[Ax^{(k-1)} - y],$$

or

$$x^{(k)} = \left[I - \frac{2}{\mu} A^* WBA\right] x^{(k-1)} + \frac{2}{\mu} A^* WBy, \tag{1}$$

which A^* denotes as usual the conjugate transpose of A.

G. F. Votruba [2] has studied recently the iteration (1) in the context of generalized inverses of matrices and has shown that if the rank of the matrix A is greater than one, and $m_i = \|r^{(i)}\|^2$, then the sequence $\{x^{(k)}\}$ converges to the vector $A^\dagger y + Px^{(0)}$, where $A^\dagger$ is the Moore-Penrose generalized inverse of A (see for instance [3], [4]) and P is the orthogonal projection of E^n on the null space of A.

3. One can extend Cimmino's method to Fredholm integral equations of the first kind

$$\mathcal{K}x = \int_a^b K(\cdot\,, t)\, x(t)\, dt = y\,, \qquad y \in L_2[a\,, b]\,, \tag{2}$$

where $K(s\,, t) \in \mathfrak{L}_2\{[a\,, b] \times [a\,, b]\}$ by defining the family of hyperplanes

$$\mathcal{H}_s = \left\{ x \in L_2[a\,, b] : \int_a^b K(s\,, t)\, x(t)\, dt = y(s) \right\}$$

for almost every $s \in [a\,, b]$. The operator $\mathcal{K}$ is completely continuous and maps $L_2[a\,, b]$ into $L_2[a\,, b]$. The orthogonal projection of a function $x_0 \in L_2[a\,, b]$ on the hyperplane $\mathcal{H}_s$ is given by the function

$$z_s = x_0 + \lambda(s)\, \overline{K(s\,, \cdot)}\,, \tag{3}$$

where

$$\lambda(s) = \frac{y(s) - \int_a^b \mathrm{K}(s, r)\, x_0(r)\, \mathrm{d}r}{\int_a^b |\mathrm{K}(s, r)|^2\, \mathrm{d}r}, \tag{4}$$

and the reflection of x_0 through $\mathcal{H}_s$ is given by $x_0 + 2\lambda(s)\, \overline{\mathrm{K}(s, \cdot)}$. We first show that $z_s \in \mathcal{H}_s$:

$$\int_a^b \mathrm{K}(s, t)\, z_s(t)\, \mathrm{d}t = \int_a^b \mathrm{K}(s, t)\, x_0(t)\, \mathrm{d}t + \lambda(s) \int_a^b \mathrm{K}(s, t)\, \overline{\mathrm{K}(s, t)}\, \mathrm{d}t =$$

$$= \int_a^b \mathrm{K}(s, t)\, x_0(t)\, \mathrm{d}t + \frac{y(s) - \int_a^b \mathrm{K}(s, t)\, x_0(t)\, \mathrm{d}t}{\int_a^b |\mathrm{K}(s, t)|^2\, \mathrm{d}t} \int_a^b |\mathrm{K}(s, t)|^2\, \mathrm{d}t =$$

$$= y(s).$$

Hence for almost all s in $[a, b]$, $z_s \in \mathcal{H}_s$. On the other hand, letting $\langle \cdot, \cdot \rangle$ denote the inner product in $\mathrm{L}_2[a, b]$, we have from (3) and (4),

$$\langle x_0 - z_s, z_s - z \rangle = \langle -\lambda(s)\, \overline{\mathrm{K}(s, \cdot)}, x_0 + \lambda(s)\, \overline{\mathrm{K}(s, \cdot)} - z \rangle =$$

$$= -\lambda(s) \{ \langle \overline{\mathrm{K}(s, \cdot)}, x_0 \rangle + \overline{\lambda(s)} \langle \overline{\mathrm{K}(s, \cdot)}, \overline{\mathrm{K}(s, \cdot)} \rangle - \langle \overline{\mathrm{K}(s, \cdot)}, z \rangle \} =$$

$$= -\lambda(s) \left\{ \langle \overline{\mathrm{K}(s, \cdot)}, x_0 \rangle + \frac{\bar{y}(s) - \int_a^b \overline{\mathrm{K}(s, t)}\, \overline{x_0(t)}\, \mathrm{d}t}{\int_a^b |\mathrm{K}(s, t)|^2\, \mathrm{d}t} \int_a^b |\mathrm{K}(s, t)|^2\, \mathrm{d}t - \right.$$

$$\left. - \int_a^b \overline{\mathrm{K}(s, t)}\, \overline{z(t)}\, \mathrm{d}t \right\} = 0 \qquad \text{for } z \in \mathcal{H}_s.$$

The next iterate x_1 in the Cimmino iteration would be the centroid of the family of points with the appropriate weight function. Taking $m(s) = \int_a^b |\mathrm{K}(s, t)|^2\, \mathrm{d}t$ to be the mass density and $\beta = \int_a^b \int_a^b |\mathrm{K}(s, t)|^2\, \mathrm{d}s\, \mathrm{d}t$

to be the total mass, we obtain

$$x_1(t) = \frac{1}{\beta}\int_a^b m(s)\left[x_0(t) + 2\lambda(s)\,\overline{K(s,t)}\right] ds =$$

$$= x_0(t) - \frac{2}{\beta}\int_a^b \overline{K(s,t)}\int_a^b K(s,r)\,x_0(r)\,dr\,ds +$$

$$+ \frac{2}{\beta}\int_a^b \overline{K(s,t)}\,y(s)\,ds .$$

In operator notation, this becomes

$$x_1(t) = \left(I - \frac{2}{\beta}\mathcal{K}^*\mathcal{K}\right)x_0(t) + \frac{2}{\beta}(\mathcal{K}^* y)(t)$$

where I is the identity map on $L_2[a, b]$ and $\mathcal{K}^*$ is the adjoint operator of $\mathcal{K}$. In general

$$x_{n+1} = \left(I - \frac{2}{\beta}\mathcal{K}^*\mathcal{K}\right)x_n + \frac{2}{\beta}\mathcal{K}^* y . \tag{5}$$

We note in passing that (5) is a particular realization of the successive approximation method with an averaging parameter,

$$x_{n+1} = x_n - \alpha\mathcal{K}^*\mathcal{K}x_n + \alpha\mathcal{K}^* y \tag{6}$$

for the solution of the equation

$$\mathcal{K}^*\mathcal{K}x = \mathcal{K}^* y . \tag{7}$$

Since the closure of the image of the unit ball under the completely continuous operator $\mathcal{K}$ is compact, $\mathfrak{R}(\mathcal{K})$, the range of $\mathcal{K}$, will be a closed subspace of $L_2[a, b]$ if and only if $\mathfrak{R}(\mathcal{K})$ is finite dimensional, i.e., if and only if the kernel $K(s, t)$ is separable or degenerate. Thus in general one cannot apply to (5) and (7) standard convergence theorems for successive approximations (for instance [5], [6]), nor recent convergence results for singular linear operators with closed range given by H. B. Keller [7], W. V. Petryshyn [8], G. F. Votruba [2] and others.

4. Let $\mathfrak{N}(\mathcal{K})$ denote the null space of $\mathcal{K}$, i.e., $\mathfrak{N}(\mathcal{K}) = \{x \in L_2[a, b] : \mathcal{K}x = 0\}$. Since $\mathcal{K}$ is continuous, $\mathfrak{N}(\mathcal{K})$ is a closed subspace of $L_2[a, b]$ and we have an orthogonal decomposition of $L_2[a, b]$, namely

$$L_2[a, b] = \mathfrak{N}(\mathcal{K}) \oplus \mathfrak{N}(\mathcal{K})^{\perp},$$

where $\mathfrak{N}(\mathcal{K})^{\perp}$ is the orthogonal complement of $\mathfrak{N}(\mathcal{K})$. The transformation $\mathcal{K}$ establishes a one to one correspondence between $\mathfrak{N}(\mathcal{K})^{\perp}$ and $\mathfrak{R}(\mathcal{K})$. The restriction $\mathcal{K} \mid \mathfrak{N}(\mathcal{K})^{\perp}$ of $\mathcal{K}$ to $\mathfrak{N}(\mathcal{K})^{\perp}$ possesses an inverse which need not be continuous unless $\mathfrak{R}(\mathcal{K})$ is closed. We define the *generalized inverse* $\mathcal{K}^{\dagger}$ of $\mathcal{K}$ to be the linear extension of $\{\mathcal{K} \mid \mathfrak{N}(\mathcal{K})^{\perp}\}^{-1}$ so that its domain is $\mathfrak{R}(\mathcal{K}) + \mathfrak{R}(\mathcal{K})^{\perp}$ and its null space is $\mathfrak{R}(\mathcal{K})^{\perp}$. It should be observed that $\mathcal{K}^{\dagger}$ is defined only on the dense linear manifold $\mathfrak{R}(\mathcal{K}) + \mathfrak{R}(\mathcal{K})^{\perp} = \mathfrak{R}(\mathcal{K}) + \mathfrak{N}(\mathcal{K}^*)$, which coincides with $L_2[a, b]$ if and only if $\mathfrak{R}(\mathcal{K})$ is closed. The range of $\mathcal{K}^{\dagger}$ is $\mathfrak{N}(\mathcal{K})^{\perp} = \overline{\mathfrak{R}(\mathcal{K}^*)}$. We note that this definition of the generalized inverse corresponds to the maximal generalized inverse with closed null space; see E. Arghiriade [9] and E. Arghiriade and A. Dragmir [10]. The Authors [11], [12] used this definition in establishing the convergence of the steepest descent and conjugate gradient methods to best approximate solutions of singular linear operator equations.

Let P be the orthogonal projection of $L_2[a, b]$ onto $\overline{\mathfrak{R}(\mathcal{K}^*)} = \mathfrak{N}(\mathcal{K})^{\perp}$ and Q the orthogonal projection of $L_2[a, b]$ onto $\mathfrak{N}(\mathcal{K}^*)$. Then it follows easily that $\mathcal{K}^{\dagger}\mathcal{K} = P$, $\mathcal{K}^{\dagger}\mathfrak{R}(\mathcal{K})^{\perp} = \{0\}$, and the projection Q is the unique continuous extension of $\mathcal{K}\mathcal{K}^{\dagger}$ to all of $L_2[a, b]$.

A function $u \in L_2[a, b]$ is called a *best approximate solution* of the integral equation (1) if u minimizes the L_2–norm of $\mathcal{K}x - y$. Clearly (1) need not have a best approximate solution for $\underline{\text{each } y} \in L_2[a, b]$, when the kernel is not separable. For if Qy is in $\overline{\mathfrak{R}(\mathcal{K})}$ but is not in $\mathfrak{R}(\mathcal{K})$, then $\inf\{\|\mathcal{K}x - y\| : x \in L_2\}$ is not attained for any $x \in L_2[a, b]$. However it is not hard to show (see [11] for a simple proof) that for each $y \in \mathfrak{R}(\mathcal{K}) + \mathfrak{R}(\mathcal{K})^{\perp}$, the domain of $\mathcal{K}^{\dagger}$, the equation (2) has a unique best approximate solution v of minimal norm given by $v = \mathcal{K}^{\dagger}y$, i.e.,

$$\|y - \mathcal{K}\mathcal{K}^{\dagger}y\| = \inf\{\|y - \mathcal{K}x\| : x \in L_2\}$$

and $\|\mathcal{K}^{\dagger}y\| < \|u\|$ for all $u \in L_2$ for which $\|y - \mathcal{K}\mathcal{K}^{\dagger}y\| = \|y - \mathcal{K}u\|$, $u \neq \mathcal{K}^{\dagger}y$. Throughout, $\|\cdot\|$ denotes the L_2–norm.

5. We are now in a position to state our Theorem on the convergence of the generalization of Cimmino's method proposed in Section 3, to the best approximate solution of minimal norm of Fredholm integral equation of the first kind. A similar result can be formulated for Volterra integral equations of the first kind. A proof in the context of general iterative methods (see also [13]) along with gradient methods for best approximate solutions of linear integral equations of the first and second kinds will appear elsewhere.

THEOREM. *Let* $K(s, t) \in \mathfrak{L}_2\{[a, b] \times [a, b]\}$, *and let* $\mathcal{K}$ *denote the completely continuous linear operator* $\mathcal{K}x = \int_a^b K(\cdot, t)\, x(t)\, dt$ *which maps* $L_2[a, b]$ *into* $L_2[a, b]$. *If the dimension of the range of* $\mathcal{K}$ *is greater than one, then the generalization of Cimmino's method converges monotonically to the best approxi-*

mate solution of minimal norm of the integral equation of the first kind, starting from the initial approximation $x_0 = 0$, *for any* $y \in \mathfrak{R}(\mathcal{K}) + \mathfrak{R}(\mathcal{K})^{\perp}$ *and*

$$\|x_n - \mathcal{K}^{\dagger} y\|^2 \leq \frac{\|\mathcal{K}^{\dagger} y\|^2 \, \|(\mathcal{K}\mathcal{K}^*)^{\dagger} y\|^2}{\|(\mathcal{K}\mathcal{K}^*)^{\dagger} y\|^2 + \frac{4n}{\beta}(\beta - \|\mathcal{K}\|^2)\,\|\mathcal{K}^{\dagger} y\|^2},$$

where $\beta = \int_a^b \int_a^b |K(s,t)|^2 \, ds \, dt.$

Acknowledgement. This research was sponsored in part by the Mathematics Research Center, University of Wisconsin, under Contract No: DA-31-124-ARO-D-462 with the United States Army, while the second author was on leave from Georgia the Institute of Technology.

REFERENCES

[1] G. CIMMINO, *Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari*, « La Ricerca Scientifica », Roma, Serie II, 326–333 (1938).
[2] G. F. VOTRUBA, *Generalized Inverses and Singular Equations in Functional Analysis*, Ph. D. Dissertation, University of Michigan, Ann Arbor, 1963.
[3] R. PENROSE, *A generalized inverse for matrices*, « Proc. Camb. Phil. Soc. », *51*, 406–413 (1955).
[4] R. PENROSE, *On best approximate solutions of linear matrix equations*, « Proc. Camb. Phil. Soc. », *52*, 17–19 (1956).
[5] F. TRICOMI, *Integral Equations*, Interscience, New York, 1957.
[6] S. G. MIKHLIN and K. L. SMOLITSKIY, *Approximate Methods for Solution of Differential and Integral Equations*, American Elsevier, New York, 1967.
[7] H. B. KELLER, *The solution of singular and semidefinite linear systems by iteration*, « SIAM J. Numer. Anal. », *2*, 281–290 (1965).
[8] W. V. PETRYSHYN, *On generalized inverses and uniform convergence of* $(I - \beta K)^n$ *with applications to iterative methods*, « J. Math. Anal. Appl. », *18*, 417–439 (1967).
[9] E. ARGHIRIADE, *Sur l'inverse generalisée d'un operateur lineaire dans les espaces de Hilbert*, « Rend. Acc. dei Lincei », *45*, fasc. 6, 471–477 (1968).
[10] E. ARGHIRIADE and A. DRAGOMIR, *Remarques sur quelques theorems relatives à l'inverse generalisée d'un operateur lineaire dans les espaces de Hilbert*, « Rend. Acc. dei Lincei », *46*, fasc. 4, 333–338 (1969).
[11] W. J. KAMMERER and M. Z. NASHED, *Steepest descent for singular linear operators with nonclosed range*, « Applicable Analysis », *1* 143–159 (1971).
[12] W. J. KAMMERER and M. Z. NASHED, *On the convergence of the conjugate gradient method for singular linear operator equations*, « SIAM J. Numer. Analysis », to appear.
[13] D. W. SHOWALTER and A. BEN-ISRAEL, *Representation and computation of the generalized inverse of a bounded linear operator between two Hilbert spaces*, « Rend. Acc. dei Lincei », *48*, fasc. 2, 184–194 (1970).