Seventieth anniversary of the conjugate gradient method and what do old papers reveal about our presence

Zdeněk Strakoš
Charles University, Prague
Jindřich Nečas Center for Mathematical Modelling

ILAS, Galway, June 2022

Many thanks

I am greatly indebted to many collaborators and friends. I am more and more thankful to those, whom I have tried to learn from without a possibility of meeting them in person.

Old work can be surprisingly revealing

Cornelius Lanczos, *Why Mathematics*, Dublin, 1966

The naive optimist, who believes in progress and is convinced that today is better than yesterday and in ten years time the world will be infinitely better off than today, will come to the conclusion that mathematics (and more generally all the exact sciences) started only about twenty years ago, while all the predecessors must have walked in a kind of limbo of half-digested and improperly conceived ideas.

Cornelius Lanczos (1950, 1953)

A computer, have to follow the notation!

Journal of Research of the National Bureau of Standards Vol. 49, No. 1, July 1952 Research Paper 2341

Solution of Systems of Linear Equations by Minimized Iterations[1]

Cornelius Lanczos

A simple algorithm is described which is well adapted to the effective solution of large systems of linear algebraic equations by a succession of well-convergent approximations.

1. Introduction

In an earlier publication [14][2] a method was described which generated the eigenvalues and eigenvectors of a matrix by a successive algorithm based on minimizations by least squares.[3] The advantage of this method consists in the fact that the successive iterations are constantly employed with maximum efficiency which guarantees fastest convergence for a given number of iterations. Moreover, with the proper care the accumulation of rounding errors can be avoided. The resulting high precision is of great advantage if the separation of closely bunched eigenvalues and eigenvectors is demanded [16].

It was pointed out in [14, p. 256] that the inversion of a matrix, and thus the solution of simultaneous systems of linear equations, is contained in the general procedure as a special case. However, in view of the great importance associated with the solution of large systems of linear equations, this problem deserved more than passing attention. It is the purpose of the present discussion to adopt the general principles of the previous investigation to the specific demands that arise if we are not interested in the complete analysis of a matrix but only in the more special problem of obtaining the solution of a given set of linear equations

$$Ay=b_0 \tag{1}$$

with a given matrix A and a given right side b_0. This is actually equivalent to the evaluation of one eigenvector only, of a symmetric, positive definite matrix. It is clear that this will require considerably less detailed analysis than the problem of constructing the *entire* set of eigenvectors and eigenvalues associated with an arbitrary matrix.

2. The Double Set of Vectors Associated With the Method of Minimized Iterations

The previous investigation [14] started out with an algorithm (see p. 261) which generated a double set of polynomials, later on denoted by $p_i(x)$ and $q_i(x)$ (see p. 274). Then a second algorithm was introduced, called "minimized iterations", which avoided the numerical difficulties of the first algorithm (see p. 287) and had, in addition, theoretically valuable properties for the solution of differential and integral equations (p. 272).

In this second algorithm, however, only *one-half* of the previous polynomials were represented, corresponding to the $p_i(x)$ polynomials whose coefficients appeared in the *full* columns of the original algorithm [14, (60)]. The polynomials $q_i(x)$, associated with the *half columns* of [14, (60)] did not come into evidence in the later procedure.

The vectors b_i, generated by minimized iterations, correspond to the polynomials $p_i(x)$ in the sense

$$b_k=p_k(A)b_0. \tag{2}$$

We should expect that the vectors generated by $q_k(A)b_0$ might also have some significance. We will see that this is actually the case. It is of considerable advantage to translate the *entire* scheme [14, (60)] into the language of minimized iterations, without omitting the half columns. We thus get a *double* set of vectors, instead of the single set considered before.

The additional work thus involved is not superfluous because the second set of polynomials can be put to good use. Moreover, the two sets of polynomials belong logically together and complement each other in a natural fashion. From the practical standpoint of adapting the resultant algorithm to the demands of large scale electronic computers, we gain in the simplicity of coding. The recurrence relations which exist between the polynomials $p_i(x)$; $q_i(x)$ are simpler in structure than the recurrence relation obtained by eliminating the second set of polynomials.

We want to simplify and systematize our notations. The vector obtained by letting the polynomial $p_k(A)$ operate on the original vector b_0, shall be called p_k:

$$p_k=p_k(A)b_0. \tag{3}$$

We thus distinguish between p_k as a *vector* and $p_k(A)$ as a *polynomial operator*. Hence the notation p_k will take the place of the previous b_k. Correspondingly we denote the associated second set of vectors by q_k:

$$q_k=q_k(A)b_0. \tag{4}$$

[1] The preparation of this paper was sponsored (in part) by the Office of Naval Research.

[2] Figures in brackets indicate the literature references at the end of this paper.

[3] The present paper is a natural sequel to the previous publication and depends on the previous findings. The reader's familiarity with the earlier development is assumed throughout this paper; the symbolism of the present paper is in harmony with that used before, in particular the notation pq, if applied to vectors, shall mean the scalar product of these two vectors.

1953

CHEBYSHEV POLYNOMIALS IN THE SOLUTION OF LARGE-SCALE LINEAR SYSTEMS*

(Toronto Symp. on Computing Techniques 1952, pp. 124–133)

By

Cornelius Lanczos
National Bureau of Standards, Los Angeles

An easily coded iterative routine is described which approximates the solution of a wide class of simultaneous linear algebraic equations by a succession of simple matrix multiplications.

The problem of finding the solution of the matrix equation

(1) Ay = b

demands the construction of the inverse matrix A^{-1} which in the case of very large matrices poses great numerical difficulties. On the other hand, the large scale machinery available in our days performs the mere substitution Ab with a given matrix A and a given vector b as a simple routine operation. Since this operation can be repeated any number of times, we have an arbitrary polynomial operation

(2) $G_m(A) \cdot b$

*The preparation of this paper was sponsored (in part) by the Office of Naval Research.

*The Editors wish to express their appreciation to the publisher for permission to reproduce this paper. A fuller statement of appreciation and permission appears in the Acknowledgements.

From Lanczos' Collected Works, Vol. VI (Paper + 2 commentaries)

Magnus R. Hestenes and Eduard Stiefel (1952)

Journal of Research of the National Bureau of Standards Vol. 49, No. 6, December 1952 Research Paper 2379

Methods of Conjugate Gradients for Solving Linear Systems[1]

Magnus R. Hestenes[2] and Eduard Stiefel[3]

An iterative algorithm is given for solving a system $Ax=k$ of n linear equations in n unknowns. The solution is given in n steps. It is shown that this method is a special case of a very general method which also includes Gaussian elimination. These general algorithms are essentially algorithms for finding an n dimensional ellipsoid. Connections are made with the theory of orthogonal polynomials and continued fractions.

1. Introduction

One of the major problems in machine computations is to find an effective method of solving a system of n simultaneous equations in n unknowns, particularly if n is large. There is, of course, no best method for all problems because the goodness of a method depends to some extent upon the particular system to be solved. In judging the goodness of a method for machine computations, one should bear in mind that criteria for a good machine method may be different from those for a hand method. By a hand method, we shall mean one in which a desk calculator may be used. By a machine method, we shall mean one in which sequence-controlled machines are used.

A machine method should have the following properties:

(1) The method should be simple, composed of a repetition of elementary routines requiring a minimum of storage space.

(2) The method should insure rapid convergence if the number of steps required for the solution is infinite. A method which—if no rounding-off errors occur—will yield the solution in a finite number of steps is to be preferred.

(3) The procedure should be stable with respect to rounding-off errors. If needed, a subroutine should be available to insure this stability. It should be possible to diminish rounding-off errors by a repetition of the same routine, starting with the previous result as the new estimate of the solution.

(4) Each step should give information about the solution and should yield a new and better estimate than the previous one.

(5) As many of the original data as possible should be used during each step of the routine. Special properties of the given linear system—such as having many vanishing coefficients—should be preserved. (For example, in the Gauss elimination special properties of this type may be destroyed.)

In our opinion there are two methods that best fit these criteria, namely, (a) the Gauss elimination method; (b) the conjugate gradient method presented in the present monograph.

There are many variations of the elimination method, just as there are many variations of the conjugate gradient method here presented. In the present paper it will be shown that both methods are special cases of a method that we call the method of conjugate directions. This enables one to compare the two methods from a theoretical point of view.

In our opinion, the conjugate gradient method is superior to the elimination method as a machine method. Our reasons can be stated as follows:

(a) Like the Gauss elimination method, the method of conjugate gradients gives the solution in n steps if no rounding-off error occurs.

(b) The conjugate gradient method is simpler to code and requires less storage space.

(c) The given matrix is unaltered during the process, so that a maximum of the original data is used. The advantage of having many zeros in the matrix is preserved. The method is, therefore, especially suited to handle linear systems arising from difference equations approximating boundary value problems.

(d) At each step an estimate of the solution is given, which is an improvement over the one given in the preceding step.

(e) At any step one can start anew by a very simple device, keeping the estimate last obtained as the initial estimate.

In the present paper, the conjugate gradient routines are developed for the symmetric and nonsymmetric cases. The principal results are described in section 3. For most of the theoretical considerations, we restrict ourselves to the positive definite symmetric case. No generality is lost thereby. We deal only with real matrices. The extension to complex matrices is simple.

The method of conjugate gradients was developed independently by E. Stiefel of the Institute of Applied Mathematics at Zurich and by M. R. Hestenes with the cooperation of J. B. Rosser, G. Forsythe, and L. Paige of the Institute for Numerical Analysis, National Bureau of Standards. The present account was prepared jointly by M. R. Hestenes and E. Stiefel during the latter's stay at the National Bureau of Standards. The first papers on this method were

[1] This work was performed on a National Bureau of Standards contract with the University of California at Los Angeles, and was sponsored (in part) by the Office of Naval Research, United States Navy.
[2] National Bureau of Standards and University of California at Los Angeles.
[3] University of California at Los Angeles, and Eidgenössische Technische Hochschule, Zurich, Switzerland.

given by E. Stiefel[4] and by M. R. Hestenes.[5] Reports on this method were given by E. Stiefel[6] and J. B. Rosser[7] at a Symposium[8] on August 23–25, 1951. Recently, C. Lanczos[9] developed a closely related routine based on his earlier paper on eigenvalue problem.[10] Examples and numerical tests of the method have been by R. Hayes, U. Hochstrasser, and M. Stein.

2. Notations and Terminology

Throughout the following pages we shall be concerned with the problem of solving a system of linear equations

$$
\begin{aligned}
a_{11}x_1+a_{12}x_2+\ \ldots\ +a_{1n}x_n&=k_1\\
a_{21}x_1+a_{22}x_2+\ \ldots\ +a_{2n}x_n&=k_2\\
\ldots\quad \ldots\quad &\ldots\\
a_{n1}x_1+a_{n2}x_2+\ \ldots\ +a_{nn}x_n&=k_n.
\end{aligned}
\tag{2:1}
$$

These equations will be written in the vector form $Ax=k$. Here A is the matrix of coefficients (a_{ij}), x and k are the vectors $(x_1,\ldots,x_n)$ and $(k_1,\ldots,k_n)$. It is assumed that A is nonsingular. Its *inverse* A^{-1} therefore exists. We denote the *transpose* of A by A^*.

Given two vectors $x=(x_1,\ldots,x_n)$ and $y=(y_1,\ldots,y_n)$, their sum $x+y$ is the vector $(x_1+y_1,\ldots,x_n+y_n)$, and ax is the vector $(ax_1,\ldots,ax_n)$, where a is a scalar. The sum

$$(x,y)=x_1y_1+x_2y_2+\ldots+x_ny_n$$

is their *scalar product*. The *length* of x will be denoted by

$$|x|=(x_1^2+\ldots+x_n^2)^{\frac{1}{2}}=(x,x)^{\frac{1}{2}}.$$

The *Cauchy-Schwarz inequality* states that for all x,y:

$$(x,y)^2\leq(x,x)(y,y)\quad\text{or}\quad|(x,y)|\leq|x|\,|y|. \tag{2:2}$$

The matrix A and its transpose A^* satisfy the relation

$$(x,Ay)=\sum_{i,j=1}^{n}a_{ij}x_iy_j=(A^*x,y).$$

If $a_{ij}=a_{ji}$, that is, if $A=A^*$, then A is said to be *symmetric*. A matrix A is said to be *positive definite* in case $(x,Ax)>0$ whenever $x\neq0$. If $(x,Ax)\geq0$ for all x, then A is said to be *nonnegative*. If A is symmetric, then two vectors x and y are said to be *conjugate* or *A-orthogonal* if the relation $(x,Ay)=(Ax,y)=0$ holds. This is an extension of the orthogonality relation $(x,y)=0$.

By an *eigenvalue* of a matrix A is meant a number λ such that $Ay=\lambda y$ has a solution $y\neq0$, and y is called a corresponding *eigenvector*.

Unless otherwise expressly stated the matrix A, with which we are concerned, will be *assumed to be symmetric and positive definite*. Clearly no loss of generality is caused thereby from a theoretical point of view, because the system $Ax=k$ is equivalent to the system $Bx=l$, where $B=A^*A$, $l=A^*k$. From a numerical point of view, the two systems are different, because of rounding-off errors that occur in joining the product A^*A. Our applications to the nonsymmetric case do not involve the computation of A^*A.

In the sequel we shall not have occasion to refer to a particular coordinate of a vector. Accordingly we may use subscripts to distinguish vectors instead of components. Thus x_0 will denote the vector $(x_{01},\ldots,x_{0n})$ and x_i the vector $(x_{i1},\ldots,x_{in})$. In case a symbol is to be interpreted as a component, we shall call attention to this fact unless the interpretation is evident from the context.

The solution of the system $Ax=k$ will be denoted by h; that is, $Ah=k$. If x is an estimate of h, the difference $r=k-Ax$ will be called the *residual* of x as an estimate of h. The quantity $|r|^2$ will be called the *squared residual*. The vector $h-x$ will be called the *error vector* of x, as an estimate of h.

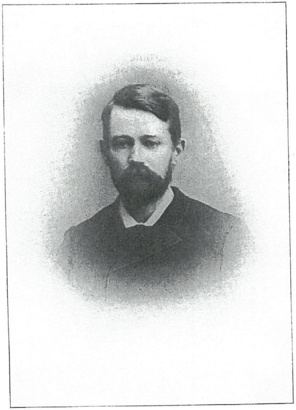
3. Method of Conjugate Gradients (cg-Method)

The present section will be devoted to a description of a method of solving a system of linear equations $Ax=k$. This method will be called the *conjugate gradient method* or, more briefly, the cg-method, for reasons which will unfold from the theory developed in later sections. For the moment, we shall limit ourselves to collecting in one place the basic formulas upon which the method is based and to describing briefly how these formulas are used.

The cg-method is an iterative method which terminates in at most n steps if no rounding-off errors are encountered. Starting with an initial estimate x_0 of the solution h, one determines successively new estimates $x_0, x_1, x_2, \ldots$ of h, the estimate x_i being closer to h than x_{i-1}. At each step the residual $r_i=k-Ax_i$ is computed. Normally this vector can be used as a measure of the "goodness" of the estimate x_i. However, this measure is not a reliable one because, as will be seen in section 18, it is possible to construct cases in which the *squared residual* $|r_i|^2$ increases at each step (except for the last) while the length of the error vector $|h-x_i|$ decreases monotonically. If no rounding-off error is encountered, one will reach an estimate $x_m (m\leq n)$ at which $r_m=0$. This estimate is the desired solution h. Normally, $m=n$. However, since rounding-

[4] E. Stiefel, Uebersicht über Methoden der Relaxationsrechnung, Z. angew. Math. Physik (3) (1952).
[5] M. R. Hestenes, Iterative methods for solving linear equations, NAML Report 52-9, National Bureau of Standards (1951).
[6] E. Stiefel, Some special methods of relaxation techniques, to appear in the Proceedings of the symposium (see footnote 8).
[7] J. B. Rosser, Rapidly converging iterative methods for solving linear equations, to appear in the Proceedings of the symposium (see footnote 8).
[8] Symposium on simultaneous linear equations and the determination of eigenvalues, Institute for Numerical Analysis, National Bureau of Standards, held on the campus of the University of California at Los Angeles (August 23–25, 1951).
[9] C. Lanczos, Solution of systems of linear equations by minimized iterations, NAML Report 52-13, National Bureau of Standards (1951).
[10] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Research NBS 45, 255 (1950) RP2133; Proceedings of Second Symposium on Large-Scale Digital Calculating Machinery, Cambridge, Mass., 1951, pages 164-206.

Thomas Jan Stieltjes (1856 - 1894)



Thomas Jan Stieltjes

1856–1894

Investigations on Continued Fractions

T. J. Stieltjes

Ann. Fac. Sci. Toulouse **8** (1894) J.1–122; **9** (1895) A.1–47 (translation)

Introduction

The object of this work is the study of the continued fraction

$$\cfrac{1}{a_1 z + \cfrac{1}{a_2 + \cfrac{1}{a_3 z + \cdots + \cfrac{1}{a_{2n} + \cfrac{1}{a_{2n+1} z + \cdots,}}}}} \qquad \text{(I)}$$

in which the a_i are positive real numbers, while z is a variable which can take all real or complex values.

Denoting by $\dfrac{P_n(z)}{Q_n(z)}$ the nth convergent[1], which depends only on the first n coefficients a_i, we shall determine in which cases this convergent tends to a limit for $n \to \infty$ and we shall investigate more closely the nature of this limit regarded as a function of z.

We shall summarize the principal result of this study. There are two distinct cases.

First case. – The series $\sum_1^\infty a_n$ is convergent.

In this case we have for each finite value of z,

$$\lim P_{2n}(z) = p(z),$$
$$\lim Q_{2n}(z) = q(z),$$
$$\lim P_{2n+1}(z) = p_1(z),$$
$$\lim Q_{2n+1}(z) = q_1(z),$$

$p(z), q(z), p_1(z), q_1(z)$ being holomorphic functions in the whole plane which satisfy the relation

$$q(z)p_1(z) - q_1(z)p(z) = +1.$$

These functions are of genus zero and admit only simple zeros which are

W. Karush (1952) and R. M. Hayes (1954)

CONVERGENCE OF A METHOD OF SOLVING LINEAR PROBLEMS[1]

W. KARUSH

1. **Introduction.** We are concerned with the solution of two problems associated with a linear operator A. First, the characteristic value problem

$$Ay = \lambda y \tag{1}$$

for the determination of the characteristic values λ and the characteristic vectors y; second, the linear equation problem

$$(A - \lambda I)x = b, \qquad b \neq 0, \tag{2}$$

for the determination of x, given the number λ and the vector b (I is the identity operator). Lanczos **[3]**[2] has described an interesting iterative method for the solution of these problems which appears to be effective for numerical calculation. It is our purpose to consider the convergence and rate of convergence of the method, in the Hilbert space sense, for a bounded self-adjoint operator.

The procedure for obtaining the solution may be described as follows. Let $b \neq 0$ be a given initial vector, arbitrary for problem (1), equal to the right side of (2) for problem (2). Let

$$\mathcal{K}_i = (b, Ab, \cdots, A^{i-1}b), \tag{3}$$

i.e., the linear subspace spanned by the indicated vectors. Let $\mathcal{K}$ be the invariant subspace which is the closure of the linear subspace spanned by all non-negative powers A^ib; symbolically

$$\mathcal{K} = (b, Ab, \cdots, A^ib, \cdots). \tag{4}$$

Let π_i be the projection operator onto $\mathcal{K}_i$. Then to solve (1) and (2) we replace the operator A by the operator π_iA on $\mathcal{K}_i$, solve the corresponding finite-dimensional problem, and allow i to approach ∞. That is, (1) and (2) are approximated respectively by

$$\pi_iAy = \lambda y \qquad \text{on } \mathcal{K}_i \tag{5}$$

and

$$(\pi_iA - \lambda I)x = b \qquad \text{on } \mathcal{K}_i. \tag{6}$$

Received by the editors February 13, 1952.

[1] The preparation of this paper was sponsored (in part) by the Office of Naval Research.

[2] Numbers in brackets refer to the list of references at the end of the paper.

3. Iterative Methods of Solving Linear Problems on Hilbert Space[1, 2]

R. M. Hayes

I. Preliminaries

1. Introduction

This paper is concerned with the study of some iterative procedures for solving linear equations for a general class of linear operators. In particular, we will consider operators expressible as the sum of a positive definite operator and a completely continuous operator. Such will be called *Legendre operators*. These will be considered as operators on a Hilbert space, and hence it will first be convenient to recall some of the fundamental properties of operators on a Hilbert space. The first few sections of the paper will do this.

Following that, the basic convergence theorems which are the concern of this paper will be proved. They include as a special case the convergence of the Rayleigh-Ritz method. The proof for the most general problem will be done in four steps. First the case where the operator is *positive definite* is proved. Then the convergence theorem for the *nonsingular Legendre operator* is reduced to that for the positive definite one. Third, the *nonnegative operator* is treated. And finally, the proof for the *general Legendre operator* is reduced to that for the nonnegative case. It is also possible to consider convergence questions for eigenvalue problems involving the same type of operators. This is done in the final section of this part of the paper. The results in this part of the paper are dependent upon the work of M. R. Hestenes [6][3].

The third major part of the paper is concerned with a description of three specific examples of iterative procedures for solving such problems. These three procedures are considered as special cases of a general iterative process which might be called *relaxation*. The first procedure considered is the *gradient method*. Convergence is proved, and an estimate for *rate of convergence* is given. In fact, it is shown that in the gradient-method convergence is geometric.

The second procedure is the so-called *conjugate-direction method*. Convergence is shown for this method, and a geometric interpretation of such processes is given. Finally, a specialization of the conjugate-direction method is considered—the *conjugate-gradient method*. Some properties of the functions generated by this method are mentioned, and estimates for *rates of convergence* are given. Here again, it is shown that the conju-

[1] The preparation of this paper was sponsored (in part) by the Office of Naval Research, USN.

[2] A thesis submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Mathematics to the Faculty of the University of California, Los Angeles.

[3] Figures in brackets indicate the literature references at the end of this paper.

Hierarchy of linear problems starting at infinite dimension

A problem with bounded invertible operator $\mathcal{G}$ on an infinite dimensional Hilbert space V

$$\mathcal{G}\, u \;=\; f$$

is approximated on a finite dimensional subspace $V_n \subset V$ by a problem with the finite dimensional operator

$$\mathcal{G}_n\, u_n \;=\; f_n\,,$$

represented, using an appropriate basis of V_n, by the matrix problem

$$\mathbf{A}\,\mathbf{x} \;=\; \mathbf{b}\,.$$

Hierarchy of linear problems starting at infinite dimension

A problem with bounded invertible operator $\mathcal{G}$ on an infinite dimensional Hilbert space V

$$\mathcal{G}\, u \;=\; f$$

is approximated on a finite dimensional subspace $V_n \subset V$ by a problem with the finite dimensional operator

$$\mathcal{G}_n\, u_n \;=\; f_n\,,$$

represented, using an appropriate basis of V_n, by the matrix problem

$$\mathbf{A}\,\mathbf{x} \;=\; \mathbf{b}\,.$$

There is a continuous operator equation posed in infinite-dimensional spaces that underlines the linear system of equations [...] awareness of this connection is key to devising efficient solution strategies for the linear systems. Hiptmair (2006)

(Infinite dimensional) Krylov subspace methods implicitly construct at the step j the finite dimensional approximation $\mathcal{G}_j$ of $\mathcal{G}$ which determines the desired approximate solution $u_j \in u_0 + \mathcal{K}_j(\mathcal{G}, r), \quad r = f - \mathcal{G}u_0$

$$u_j \; := \; u_0 + p_{j-1}(\mathcal{G})\, r \;\; \approx \;\; u \; = \; \mathcal{G}^{-1} f \,.$$

Here $p_{j-1}(\lambda)$ is the associated polynomial of degree at most $j-1$ and $\mathcal{G}_j$ is obtained by restricting and projecting $\mathcal{G}$ onto the jth Krylov subspace

$$\mathcal{K}_j(\mathcal{G}, r) \; := \; \mathrm{span}\left\{ r, \mathcal{G}r, \ldots, \mathcal{G}^{j-1} r \right\}.$$

A.N. Krylov (1931), Gantmakher (1934), Hestenes and Stiefel (1952), Lanczos (1952-53); Karush (1952), Hayes (1954), Stesin (1954), Vorobyev (1958)

Algebraic CG for $Ax = b$ with A HPD (1952)

$r_0 = b - Ax_0,\ p_0 = r_0\,.$ For $n = 1, \dots, n_{\max}$:

$$\begin{aligned}
\alpha_{n-1} &= \frac{r_{n-1}^* r_{n-1}}{p_{n-1}^* A p_{n-1}} \\
x_n &= x_{n-1} + \alpha_{n-1} p_{n-1}\,, \quad \text{stop when the stopping criterion is satisfied} \\
r_n &= r_{n-1} - \alpha_{n-1} A p_{n-1} \\
\beta_n &= \frac{r_n^* r_n}{r_{n-1}^* r_{n-1}} \\
p_n &= r_n + \beta_n p_{n-1}
\end{aligned}$$

Here α_{n-1} ensures the minimization of the energy norm $\|x - x_n\|_A$ along the line

$$z(\alpha) = x_{n-1} + \alpha p_{n-1}\,.$$

Mathematical elegance of CG: Galerkin orthogonality gives optimality

Provided that

$$p_i \perp_A p_j , \quad i \neq j,$$

the one-dimensional line minimizations at the individual steps 1 to n result in the n-dimensional minimization over the whole shifted Krylov subspace

$$x_0 + \mathcal{K}_n(A, r_0) \ = \ x_0 + \mathrm{span}\{p_0, p_1, \ldots, p_{n-1}\} \, .$$

Indeed,

$$x - x_0 \ = \ \sum_{\ell=0}^{N-1} \alpha_\ell p_\ell \ = \ \sum_{\ell=0}^{n-1} \alpha_\ell p_\ell \ + \ x - x_n \, ,$$

where

$$x - x_n \ \perp_A \ K_n(A, r_0) \, , \quad \text{or, equivalently,} \quad r_n \ \perp \ K_n(A, r_0) \, .$$

Cornelius Lanczos, March 9, 1947

On (what are now called) the Lanczos and CG methods:

The reason why I am strongly drawn to such
approximation mathematics problems is ... the fact that
a very "economical" solution is possible only when it is very "adequate".

To obtain a solution in very few steps
means nearly always that one has found a way
that does justice to the inner nature of the problem.

Albert Einstein, March 18, 1947

Your remark on the importance of adapted approximation methods makes very good sense to me, and I am convinced that this is a fruitful mathematical aspect, and not just a utilitarian one.

Your remark on the importance of
adapted approximation methods makes very
good sense to me, and I am convinced
that this is a fruitful mathematical aspect,
and not just a utilitarian one.

Nonlinear and globally optimal adaptation of the iterations within Krylov subspaces of increasing dimensionality to the problem data.

Magnus R. Hestenes and Eduard Stiefel (1952)

- Abstract. *An iterative algorithm is given for solving a system $Ax = k$ of n linear equations in n unknowns. The solution is given in n steps. [...] Connections are made with the theory of orthogonal polynomials and continued fractions.*

- Section 3. *At each step the residual $r_i = k - Ax_i$ is computed. Normally this vector can be used as a measure of the "goodness" of the estimate x_i. However, this measure is not a reliable one because [...] it is possible to construct cases in which the squared residual $|r_i|^2$ increases at each step (except for the last) while the length of the error vector decreases monotonically.*

- Section 8. *Propagation of Rounding-Off Errors in the cg-Method.*

- Sections 14. - 17. Orthogonal polynomials, (Riemann)-Stieltjes integral, *mass distribution on the positive axis. [...] During the following investigations we use the Gauss mechanical quadrature as a basic tool ...*

  Gauss quadrature is equivalent to solving the simplified Stieltjes problem of moments.

- Section 18. Continued fractions.

Cornelius Lanczos (1952)

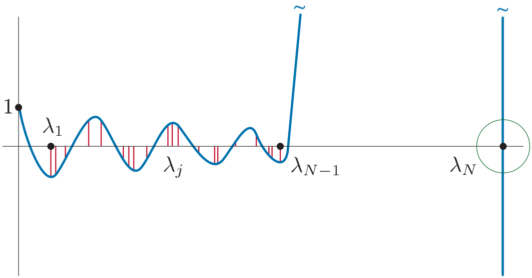
- Title. *Solution of Systems of Linear Equations by Minimized Iterations.*
- *In principle we have obtained a method for the solution of sets of linear equations which is simple and logical in structure. Yet from numerical standpoint we must not overlooked the danger of the possible accumulation of rounding errors.*
- Algorithm I: *purification* of the initial vector of the components in the direction of the eigenvectors corresponding to large eigenvalues using Chebyshev polynomials.
- Algorithm II: minimized iterations equivalent to CG.
- *The principle by which this process* [meant CG] *gives good attenuation is quite different from the previous one* [meant the purification using Chebyshev polynomials]. *The polynomials of this process have the peculiarity that they attenuate due to the nearness of their zeros to those λ-values which are present in A. The advantage of the process is its great economy.*
- *The price we have to pay is that the successive iterations of this process are more complicated than those of algorithm 1. Another difficulty arises from the inevitable accumulation of rounding errors.*

How does our presence match the visions in these papers?

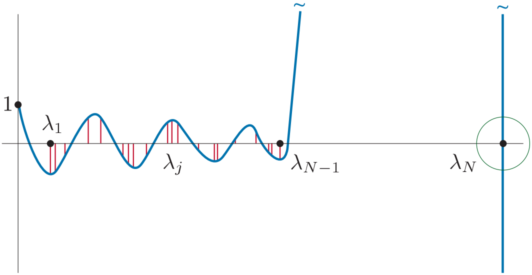
- The papers by Lanczos, Hestenes and Stiefel, Karush, Hayes, and the book by Vorobyev (1958R, 1965E, not discussed here), made many fundamental points.
- Some were painfully rediscovered (often through computational failures) decades later, other remain unnoticed in literature, including textbooks and monographs, until now.
- The common knowledge on CG has been reduced to an algorithmic description without broader context. Convergence rate is viewed through the two-D projection resulting in the linear Chebyshev-polynomial-based upper bound, which is sometimes combined with misguided or even plainly wrong arguments on clustering of eigenvalues.
- Rounding errors are typically excluded from theoretical considerations, while the derived results are subsequently applied to practical computations.
- Examples: Attenuation using the (composite) Chebyshev polynomials and attenuation given by the CG polynomials based on the Galerkin orthogonality, superlinear convergence.

Attenuation using Chebyshev polynomials and the CG polynomials

composite shifted Chebyshev: $$(1 - \frac{\lambda}{\lambda_N}) \min_{p \in \mathcal{P}_n(0)} \max_{\lambda \in [\lambda_1, \lambda_{N-1}]} |p(\lambda)|$$



composite shifted Chebyshev: $$(1 - \frac{\lambda}{\lambda_N}) \min_{p \in \mathcal{P}_n(0)} \max_{\lambda \in [\lambda_1, \lambda_{N-1}]} |p(\lambda)|$$



conjugate gradients: $$\|x - x_j\|^2 = \|r_0\|^2 \sum_{\ell=1}^{N} \omega_\ell \, \frac{(\varphi_j(\lambda_\ell))^2}{\lambda_\ell}$$

W. Karush (1952) and R. M. Hayes (1954), Superlinear convergence

Let $\mathcal{G} = \mathcal{I} + \mathcal{F}$ be self-adjoint, bounded and coercive,
with $\mathcal{F}$ being compact.

Let $\mathcal{G} = \mathcal{I} + \mathcal{F}$ be self-adjoint, bounded and coercive, with $\mathcal{F}$ being compact.

Then the rate of convergence of the conjugate gradient method for a linear problem with the operator $\mathcal{G}$ is accelerating and it is asymptotically faster than any geometric progression.

Lanczos, Hestenes, Stiefel - phrases from the papers, Liesen, S (2013)

Numerical analysis

Convergence analysis Rounding error analysis Cost of computations Floating point computations

Iterative methods Polynomial preconditioning Stopping criteria Data uncertainty

Least squares solutions

Optimisation

Convex geometry

Minimising functionals

Approximation theory

Orthogonal polynomials

Chebyshev, Jacobi and Legendre polynomials

Green's function

Gibbs oscillation

Rayleigh quotients

Fourier series

Cornelius Lanczos

An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, 1950

Solution of systems of linear equations by minimized iterations, 1952

Chebyshev polynomials in the solution of large-scale linear systems, 1952

Magnus R. Hestenes & Eduard Stiefel

Methods of conjugate gradients for solving linear systems, 1952

Structure and sparsity

Gaussian elimination

Vandermonde determinant

Matrix theory

Linear algebra

General inner products

Cauchy-Schwarz inequality

Orthogonalisation

Projections

Functional analysis

Differential and integral operators

Liouville-Neumann expansion

Trigonometric interpolation Continued fractions Sturm sequences Fredholm problem

Gauss-Christoffel quadrature Riemann-Stieltjes integral Dirichlet and Fejér kernel

Real analysis

Hiptmair, CMA (2006):

Operator preconditioning is a very general recipe [...]. It is simple to apply, but may not be particularly efficient, because in case of the [condition number] *bound of Theorem [...] is too large, the operator preconditioning offers no hint how to improve the preconditioner. Hence, operator preconditioner may often achieve [...] the much-vaunted mesh independence of the preconditioner, but it may not perform satisfactorily on a given mesh.*

An example of warnings: Spectral equivalence and asymptotic behavior

Faber, Manteuffel and Parter, Adv. in Appl. Math. (1990):

For a fixed h, using a preconditioning strategy based on an equivalent operator may not be superior to classical methods [...] Equivalence alone is not sufficient for a good preconditioning strategy. One must also choose an equivalent operator for which the bound is small.

There is no flaw in the analysis, only a flaw in the conclusions drawn from the analysis [...] asymptotic estimates ignore the constant multiplier. Methods with similar asymptotic work estimates may behave quite differently in practice.

The state of the art opinions often do not follow the foundations

Referee report: *The only new items presented here have to do with analysis involving floating point operations [...] . These are likely to bear very little interest to the audience of* [this Journal] *...*

... the authors give a misguided argument. The main advantage of iterative methods over direct methods does not primarily lie in the fact that the iteration can be stopped early (whatever this means), but that their memory (mostly) and computational requirements are moderate.

It appears obvious to the authors that the A-norm is the quantity to measure to stop the iteration. In some case [...] it is the residual norm (yes) that matters. For example, in nonlinear iterations, it is important to monitor the decrease of the residual norm - because the nonlinear iteration looks at the non-linear residual to build globally convergent strategies. This is known to practitioners, yet it is vehemently rejected by the authors.

And they even refer to the nonexistent and principally wrong results

A quote from a very influential paper:

And they even refer to the nonexistent and principally wrong results

A quote from a very influential paper:

Soon after the introduction of $\kappa(A)$ for error analysis, Hestenes and Stiefel showed that this quantity also played a role in complexity analysis. More precisely, they showed that the number of iterations of the conjugate gradient method (assuming infinite precision) needed to ensure that the current approximation to the solution of a linear system attained a given accuracy is proportional to $\sqrt{\kappa(A)}$.

It seems to be time to humbly return to the founding authors

Cornelius Lanczos, *Why Mathematics*, Dublin, 1966

In a recent comment on mathematical preparation an educator wanted to characterize our backwardness by the following statement: "Is it not astonishing that a person graduating in mathematics today knows hardly more than what Euler knew already at the end of the eighteenth century?". On its face value this sounds a convincing argument. Yet it misses the point completely. Personally I would not hesitate not only to graduate with first class honors, but to give the Ph.D. (and with summa cum laude) without asking any further questions, to anybody who knew only one quarter of what Euler knew, provided that he knew it in the way in which Euler knew it.

Some references

- Liesen,S, *Krylov subspace methods: principles and analysis,* Oxford University Press (2013)
- Málek, S, *Preconditioning and the conjugate gradient method in the context of solving PDEs,* SIAM (2015)
- Carson, S, *On the cost of iterative computations*, Phil. Trans. R. Soc. A 378:20190050

Some references

- Liesen,S, *Krylov subspace methods: principles and analysis,* Oxford University Press (2013)
- Málek, S, *Preconditioning and the conjugate gradient method in the context of solving PDEs,* SIAM (2015)
- Carson, S, *On the cost of iterative computations*, Phil. Trans. R. Soc. A 378:20190050

Thank you very much for your kind attention.