SIAM J. NUMERICAL ANALYSIS
Vol. 16, No. 3, June 1979

0036-1429/79/1603-0002 $01.00/0

SOME CONVERGENCE PROPERTIES OF THE CONJUGATE GRADIENT METHOD IN HILBERT SPACE*

ZENON FORTUNA†

Abstract. A rate of convergence of the conjugate gradient method for minimizing the convex quadratic functionals in Hilbert space is investigated. Daniel's approach via spectral analysis is developed and the theory of moments is applied. Main results: the investigation of conjugate gradient method properties is replaced by the investigation of some approximation problems; a rate of convergence is characterized by a Jacobi matrix closely related to the Hessian of the functional. Examples have been constructed when the convergence of the conjugate gradient method is only linear.

1. Introduction. The problem of unconstrained minimization of a uniformly convex quadratic functional $Q(\cdot)$, $Q\colon U \to R^1$ where U is a real Hilbert space, has received a great deal of attention in recent years, e.g. see [1], [2], [3]. Usually it is assumed that the gradient $g(u)$ is available at each point $u \in U$ and thus gradient methods are applied. Among all gradient methods the conjugate method (CGM) seems to be most convenient from the numerical point of view. This method spares the memory of the computer, is not very time consuming and exhibits good convergence properties when U is a finite dimensional space.

The well known results are the following:

P1. When $U = R^n$, CGM exhibits finite termination (after at most n steps the optimal point is reached); see [2], p. 53.

P2. In Hilbert space at each iteration the reduction of the value of the functional by CGM is at least as good as by the steepest descent method (SDM); see [1], p. 117. This means that when $Q_0, Q_1, Q_2, \cdots, Q_i$ is the sequence of values of the functional Q generated by the CGM and $Q_1', Q_2', \cdots, Q_i'$ is the sequence of corrections generated by the SDM from points $u_0, u_1, \cdots, u_{i-1}$ obtained by the CGM then (by the assumption that $Q_i \to \hat{Q} = 0$)

$$\frac{Q_{i+1}}{Q_i} \leqq \frac{Q'_{i+1}}{Q_i}, \qquad i = 0, 1, 2, \cdots.$$

Let us assume for simplicity that the optimal point is $\hat{u} = 0$ and the optimal value is $\hat{Q} = 0$. The Hessian $H \in L(U, U)$ of the functional $Q(u)$ is assumed to be bounded and let $\mathrm{sp}(H) \subset [m, 1]$ for some $0 < m \leqq 1$. Then from P2 and SDM properties it follows that q-rate of convergence[1] for the CGM can be evaluated by

$$\frac{Q_{i+1}}{Q_i} \leqq \left(\frac{1-m}{1+m}\right)^2, \qquad i = 0, 1, 2, \cdots.$$

Daniel's [4] approach via spectral analysis proved that the factor $((1-m)/(1+m))^2$ could be decreased by using the r-rate of convergence: the asymptotic r-rate of convergence for the CGM can be bounded from above by $((1-\sqrt{m})/(1+\sqrt{m}))^2$.

Daniel proved also that, when the Hessian $H = \lambda I + C$, where $\lambda > 0$ and C is a nonnegative symmetric compact operator, then the convergence is even q-superlinear in the sense that $Q_{i+1}/Q_i \to 0$. This result was generalized in [5] in the form:

* Received by the editors September 26, 1974, and in revised form August 9, 1978.

† Institut of Automatic Control, Warsaw Technical University, Warsaw, Poland.

[1] For the definition of rates of convergence see [11].

P3. When $H = H_1 + H_2$ where H_1 is an algebraic operator (that means it has a characteristic polynomial of degree N) and H_2 is compact, then $Q_{i+N}/Q_i \to 0$, which is sufficient for r-superlinear convergence in the sense that $\sqrt[i]{Q_i} \to 0$.

The case when $H = H_1 + H_2$ appears in practice very often, e.g. when $Q(u, x)$ is an integral quadratic functional and linear state equations $P(x, u) = 0$ have a resolvent operator of Volterra type. For such a type of functional the numerical results confirm the superlinear convergence. Because the majority of known examples of infinite dimensional quadratic functionals are linear-quadratic control problems, the hypothesis was suggested that, in Hilbert space, the CGM exhibits superlinear convergence. In the next section the spectral approach of Daniel will be developed and the analysis of convergence of the CGM will be replaced by equivalent mean-square approximation problem. In § 3 the speed of approximation process will be characterized by some Jacobi matrix closely related to the Hessian H. In § 4 the further properties of such an approach will be given. In the last section two examples will be presented with only linear convergence of the CGM.

2. The CGM as an optimal approximation process. Let $Q(u) = \frac{1}{2}\langle u, Hu\rangle$ where $\langle \cdot, \cdot \rangle$ denotes the inner product in U. The simplest form of the CGM algorithm for quadratic Q is the following:

0. Choose u_0, find $g_0 = g(u_0)$; $i := 0$. IF $\|g_0\| = 0$, STOP. Otherwise take $s_0 = -g_0$ and GO TO 1.
1. Find $u_{i+1} = u_i + \tau_i \cdot s_i$, where step coefficient τ_i is found in 1-dimensional minimization procedure so that $Q(u_i + \tau_i s_i) \leqq Q(u_i + \tau s_i)$ for all $\tau \geqq 0$. GO TO 2.
2. Count $g_{i+1} = g(u_{i+1})$. IF $\|g_{i+1}\| = 0$, STOP. Otherwise take $s_{i+1} = -g_{i+1} + \alpha_i \cdot s_i$, where $\alpha_i = \|g_{i+1}\|^2/\|g_i\|^2$. $i := i+1$, GO TO 1.

It was noticed that the sequence $Q_0, Q_1, Q_2 \cdots$ is an optimal one among all gradient methods in the sense that Q_{i+1} is the minimal value of Q on a flat $u_0 + \operatorname{lin}\{g_0, g_1, \cdots, g_i\}$; see [1, p. 119].

Because $\operatorname{lin}\{g_0, g_1, \cdots, g_i\} = \operatorname{lin}\{g_0, Hg_0, \cdots, H^i g_0\}$ and $Q(u) = \frac{1}{2}\langle g, H^{-1} g\rangle$ then

$$Q_i = \tfrac{1}{2} \min_{c_1, \cdots, c_i} \langle g_0 + c_1 H g_0 + \cdots + c_i H^i g_0, H^{-1}(g_0 + c_1 H g_0 + \cdots + c_i H^i g_0)\rangle.$$

Let P_t be the resolution of the identity for the operator H, and $f(t) = \langle g_0, P_t g_0\rangle$ be the spectral function. The function $f(t)$ is a nondecreasing function such that $f(t+0) = f(t)$, $f(m) = 0$, $f(1) = \langle g_0, g_0\rangle$ and for any monomial $p(t) = t^i$, $i = 0, \pm 1, \pm 2, \cdots$, $\langle g_0, p(H) g_0\rangle = \int_m^1 p(t) \cdot df(t)$; see e.g. [7, p. 895]. Thus

$$Q_i = \tfrac{1}{2} \min_{c_1, \cdots, c_i} \int_m^1 (1 + c_1 t + \cdots + c_i t^i)^2 \cdot \frac{1}{t} \cdot df(t)$$

and the question whether the CGM is superlinearly convergent is replaced by a problem of how quickly is a constant function 1 approximated in a mean-square sense by functions $t, t^2, \cdots$ on the interval $[m, 1]$ with the integral measure $(1/t)\, df(t)$. In the theory of integral measures (weights) $df(t)$ two cases are distinguished:

I. The function $f(t)$ has a finite number of points of increase, say M points.

DEFINITION. t_0 is a *point of increase* of the nondecreasing function $f(t)$ defined on $[m, 1]$ iff: 1. $t_0 = m$ and $f(t_0) < f(t)$ for $t(t_0, 1]$, or 2. $t_0 \in (m, 1)$ and $f(t_1) < f(t_2)$ for any $t_1 \in [m, t_0)$ and $t_2 \in (t_0, 1]$, or 3. $t_0 = 1$ and $f(t) < f(t_0)$ for $t \in [m, t_0)$.

This case applies when $U = R^n$ or, more generally, when g_0 belongs to some finite dimensional subspace of u spanned by a set of eigenvectors of H. In this case $Q_{M-1} \neq 0$ but $Q_M = 0$, and finite termination occurs.

II. The function $f(t)$ has an infinite number of points of increase.

In this case $Q_i \neq 0$ for all i and the convergence properties could be investigated. Below it is assumed that case II applies.

3. Rate of convergence of the approximation process. Let $q_i = Q_{i+1}/Q_i$. Because of § 2 this ratio can be expressed by

$$q_i = \frac{\min_{c_1,\cdots,c_{i+1}} \int_m^1 (1+c_i+\cdots+c_{i+1}t^{i+1})^2 \cdots (1/t)\cdot df(t)}{\min_{c_1,\cdots,c_i} \int_m^1 (1+c_1t+\cdots+c_it^i)^2\cdot(1/t)\,df(t)}$$

LEMMA 1.

$$q_i = \frac{\min_{c_0,\cdots,c_i} \int_m^1 (t^{i+1}+c_it^i+\cdots+c_0)^2\cdot(1/t)\cdot df(t)}{\min_{c_0,\cdots,c_{i-1}} \int_m^1 (t^i+c_{i-1}t^{i-1}+\cdots+c_0)^2\cdot t\cdot df(t)}.$$

This lemma can be easily proved by means of the theory of moments, applied to the mean-square approximation problem; see e.g. [9, p. 324] for additional information.

COROLLARY. *When*

$$r_i = \frac{\min_{c_0,\cdots,c_i} \int_m^1 (t^{i+1}+c_it^i+\cdots+c_0)^2\,df(t)}{\min_{c_0,\cdots,c_{i-1}} \int_m^1 (t^i+c_{i-1}t^{i-1}+\cdots+c_0)^2\,df(t)}$$

then

$$r_i \leqq q_i \leqq 1/m^2\cdot r_i, \qquad i=0,1,\cdots,$$

which follows from the inequalities $m \leqq t \leqq 1$ *and* $1 \leqq 1/t \leqq 1/m$.

Now a Jacobi matrix closely related to the Hessian H will be defined. The ratios r_i will be equal to the squared subdiagonal elements of this matrix.

Let us recall that a Jacobi matrix J is an infinite symmetric tridiagonal matrix with nonnegative elements a_i, $i=0,1,\cdots$, on the diagonal and positive elements b_i, $i=0,1,\cdots$, on the super- and subdiagonal.

Let $J^{(k)}$ denote the kth main matrix fraction of J. Thus $J^{(k)}$ has elements $a_0, a_1, \cdots, a_{k-1}$ and $b_0, b_1, \cdots, b_{k-2}$. The notation $mI \leqq J \leqq I$ means that sp $(J^{(k)}) \subset [m, 1]$ for any $k=1,2,\cdots$.

LEMMA 2. *If* $b_{i-1}+a_i+b_i \leqq 1$ *and* $-b_{i-1}+a_i-b_i \geqq m$ *for* $i=0,1,\cdots$; $b_{-1} \triangleq 0$, *then* $mI \leqq J \leqq I$.

This lemma follows easily from Gershgorin's theorem. Let $f(t)$ be the spectral function for the initial gradient g_0, and let $x_0, x_1, \cdots$ be the orthonormal sequence of positive polynominals obtained in Gram–Schmidt orthonormalization from the sequence $1, t, \cdots$ and with inner product $[\cdot,\cdot] = \int_m^1(\cdot)(t)(\cdot)(t)\,df(t)$.

LEMMA 3. *For any initial gradient* g_0 *there exists a unique Jacobi matrix denoted by* $J(H, g_0)$ *such that* $t\cdot x_i(t) = b_{i-1}\cdot x_{i-1}(t) + a_i\cdot x_i(t) + b_i\cdot x_{i+1}(t)$, $t \in [m, 1]$, $i = 0, 1, \cdots, b_{-1} = 0$.

This is a version of the three-term recurrence relation.

Remark. It can be shown that the matrix $J(H, g_0)$ is the matrix form of the Hessian H with the basis $\bar{g}_0, \bar{g}_1, \bar{g}_2, \cdots$ (the sequence of normed gradients obtained in the CGM algorithm).

LEMMA 4.

$$r_i = b_i^2, \qquad i=0,1,\cdots.$$

The proof follows from the three-term recurrence relation and the property that the polynomials in mean-square approximation problems are mutually orthogonal.

COROLLARY. *The sequence $Q_0, Q_1, \cdots$ converges q-superlinearly iff $b_i \to 0$ and r-superlinearly iff $\sqrt[i]{b_0 b_1 \cdot \cdots \cdot b_{i-1}} \to 0$.*

This lemma follows from Lemma 1 and its Corollary and Lemma 4.

4. Properties of matrices $J(H, g_0)$.

LEMMA 5. *When* sp $(H) \subset [m, 1]$ *then for any* $g_0 \in U$, $mI \leqq J(H, g_0) \leqq I$.

Proof. The proof of the lemma follows from the properties of the moments of "distribution of masses"—see [10, p. 91].

The construction of the matrix $J(H, g_0)$ can in some sense be inverted and for given $\tilde{J}$-matrix it is possible to define such an operator H and vector g_0 that $\tilde{J} = J(H, g_0)$: Let $e_0, e_1, \cdots$ be the orthonormal basis in the space U. Then the operator $\mathring{H}(J)$ defined by $\mathring{H}e_i = b_{i-1}e_{i-1} + a_i e_i + b_i e_{i+1}$, $i = 0, 1, \cdots, b_{-1} = 0$, is the linear operator with the property $\langle \mathring{H}e_i, e_j \rangle = \langle e_i, \mathring{H}e_j \rangle$ for any i, j. Thus also for any vectors v_1 and v_2 being finite linear combinations of $e_0, e_1, \cdots$ we have $\langle \mathring{H}v_1, v_2 \rangle = \langle v_1, \mathring{H}v_2 \rangle$. The set of vectors v is dense in U, so $\mathring{H}(J)$ is a symmetric operator.

LEMMA 6. *When $mI \leqq J \leqq I$ then $m \cdot \|v\|^2 \leqq \langle \mathring{H}(J)v, v \rangle \leqq \|v\|^2$ for any $v = \sum_{i=0}^{n} c_i e_i$,* $n = 0, 1, \cdots$.

Proof. It is sufficient to notice that $\|v\| = \|c\|$ where $c = (c_0, c_1, \cdots c_n)^T$ and $\langle v, \mathring{H}(J)v \rangle = \langle c, J^{(n+1)}c \rangle$.

COROLLARY. Because $\mathring{H}(J)$ is a symmetric bounded linear operator, its closure $H(J)$ is a linear, bounded, selfadjoint operator; see e.g. [8].

LEMMA 7. $J(H(\tilde{J}), e_0) = \tilde{J}$.

Proof. The proof is straightforward. It is enough to evaluate the elements of the matrix $J(H(\tilde{J}), e_0)$, using the properties of $H(\tilde{J})$.

LEMMA 8. *If $mI \leqq J \leqq I$ then* sp $(H(J)) \subset [m, 1]$.

It follows from Lemma 6 and the fact that the set $e_0, e_1, \cdots$ forms a basis in U.

As a summary of §§ 2, 3 and 4 the following theorem is obtained.

THEOREM *The investigation of the superlinear convergence of the CGM for quadratic functionals with the Hessian H restricted by* sp $(H) \subset [m, 1]$, $0 < m \leqq 1$, *is equivalent to investigation of the properties of sequences $b_0, b_1, \cdots$, of such J-matrices that $mI \leqq J \leqq I$.*

5. The CGM could only be linearly convergent.

EXAMPLE 1. *Let J_0 be the Jacobi matrix with $a_i = (1+m)/2$, $b_i = (1-m)/4$, $i = 0, 1, \cdots, m \in (0, 1)$. When the CGM is applied to $Q(u) = \frac{1}{2}\langle u, H(J_0)u \rangle$ with $u_0 = H(J_0)^{-1} e_0$ (i.e. with $g_0 = e_0$) then $((1-m)/4)^2 \leqq Q_{i+1}/Q_i$ and the convergence $Q_i \to 0$ is only linear.*

Proof. The above example guarantees that $mI \leqq H(J_0) \leqq I$ by Lemmas 2 and 8. Thus $Q_{i+1}/Q_i \geqq b_i^2$ by Lemmas 1 and 4.

EXAMPLE 2. *Let $Q(u) = \int_\varepsilon^{2+\varepsilon} u(t) \cdot t \cdot u(t)\, dt$, $\varepsilon > 0$. When $u_0(t) = 1/t$, $t \in [\varepsilon, 2+\varepsilon]$ then the CGM generates sequence $Q_0, Q_1, \cdots$ converging to 0 and where $Q_i \geqq Q_0 \cdot (1/(4+2\varepsilon))^{2i}$, $i = 0, 1, 2, \cdots$.*

Proof. The proof of that is essentially similar to the proof of Lemma 1; it uses properties of the Legendre polynomials.

REFERENCES

[1] J. W. DANIEL, *The approximate minimization of functionals*, Prentice-Hall, Englewood Cliffs, NJ, 1971.

[2] E. POLAK, *Computational methods in optimization*, Academic Press, New York, 1971.

[3] M. J. POWELL, *Some convergence properties of the conjugate gradient method*, Math. Progr., 11 (1976), pp. 42–49.

[4] J. W. DANIEL, *The conjugate gradient method for linear and nonlinear operator equations*, this Journal, 4 (1967), pp. 10–26.

[5] Z. FORTUNA, *Superlinear convergence of conjugate gradient method*, Control and Cybernetics, 3 (1974), pp. 63–68.

[6] L. C. W. DIXON, *Conjugate gradient algorithms. Quadratic termination without linear searches*, J. Inst. Math. Appl., 15 (1975), pp. 9–18.

[7] N. DUNFORD AND J. SCHWARZ, *Linear Operators, Part II*, Interscience, New York, 1963.

[8] W. MLAK, *Introduction to Hilbert space theory*, Polish Scientific Publishers, Warsaw, 1972. (In Polish.)

[9] I. P. NATANSON, *Constructive theory of functions*, GITTL, Moscow, 1949. (In Russian.)

[10] M. G. KREJN AND A. A. NUDELMAN, *Markov's problem of moments and extremal problems*, Nauka, Moscow, 1973. (In Russian.)

[11] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.