Control
and Cybernetics
VOL. 3 (1974) No. 1/2

Superlinear convergence of conjugate gradient method

by

ZENON FORTUNA

Technical University of Warsaw
Institute of Automatic Control

For the case of minimizing a quadratic functional in a real Hilbert space with Hessian $A=A_1+A_2$, where A_1 is a an algebraic operator and A_2 is a compact operator, superlinear convergence of the Conjugate Gradient Method is obtained. An example of a functional of such a type is encountered in an optimal control problem for a linear dynamic system with delays.

The Conjugate Gradient Method (CGM) has been used since its introduction by Hestenes [3] in many algorithms of numerical optimization. It was soon noticed that CGM is considerably better than the Steepest Descent Method (SDM). The theoretical evaluation, however, of the rate of convergence ($q_n=(Q_{n+1}-\hat{Q})/(Q_n-\hat{Q})$ Q_n is suboptimal value of functional at the n-th step of iterations) for CGM remained no better than for SDM, i.e. when sp $A\subset[m, M]$, $m>0$, then

$$q_n \leqslant \left(\frac{M-m}{M+m}\right)^2, \qquad n=0, 1, \ldots .$$

In [1] can be found a better evaluation

$$q_n \leqslant \left(\frac{\sqrt{M}-\sqrt{m}}{\sqrt{M}+\sqrt{m}}\right)^2 + \varepsilon_n,\ \varepsilon_n \to 0, \qquad n=0, 1, \ldots$$

but this is still linear convergence.

In this paper we show that for a wide class of minimized functionals we have, in fact, superlinear convergence of CGM, i.e., $\sqrt[n]{Q_n-\hat{Q}} \to 0$.

Problem. Let us have the following process:

$$\begin{aligned} &\dot{x}(t)=Dx(t)+Ex(t-h)+Fu(t), \\ &x(t)=\varphi(t) \text{ for } t\in[t_0-h, t_0]. \end{aligned} \tag{1}$$

We want to find a control function $\hat{u}$ which minimizes the functional

$$Q(u)=\tfrac{1}{2}\int_{t_0}^{t_1}[(x(t),Nx(t))+(u(t),Ru(t))]\,dt \tag{2}$$

$N\geqslant 0$, $R>0$, D, E, F — constant matrices.

Under these assumptions $\hat{u}$ exists and is unique.

If we want to find the actual solution, we have to solve the two-point boundary-value problem. Usually we use iterative algorithms for this problem. It is easy to compute the gradient of functional (2) with constraints (1) so it is reasonable to choose the Conjugate Gradient Method. In the case of a problem with delays ($E\neq 0$) the numerical procedures typically converge more slowly than without delays. We want to show that these numerical difficulties decreases with the number of iterations, because in the general case of process (1), the CGM assures superlinear convergence.

In [2] we find the solution of equation (1)

$$x(t)=x_0(t)+\int_{t_0}^{t}K(t,\tau)\,Fu(\tau)\,d\tau,$$

where $x_0(t)$ is the trajectory generated by $u(t)=0$ and $K(t,\tau)$ is a continuous function matrix. Let us denote $x(t)$ by the shorter notation

$$x=x_0+Lu \tag{3}$$

where L is above-defined compact linear operator in space U of piece-wise continuous function $u(t)$ (Volterra's operator with continuous kernel). After substituting (3) into (2) we obtain

$$Q(u)=\tfrac{1}{2}[\langle u,(R+L^*NL)u\rangle+\langle x_0,NLu\rangle+\langle u,L^*Nx_0\rangle+\langle x_0,Nx_0\rangle]=$$
$$=\tfrac{1}{2}\langle u-\hat{u},A(u-\hat{u})\rangle+\hat{Q}$$

where $\langle\ldots\rangle$ denotes inner product in $L^2(t_1,t_2,R^n)$

$$A=R+L^*NL,\ u=-A^{-1}L^*Nx_0,\ \hat{Q}=Q(\hat{u}).$$

It is sufficient then to prove the following theorem:

We use CGM for minimization of the quadratic functional

$$Q(u)=\tfrac{1}{2}\langle u-\hat{u},A(u-\hat{u})\rangle, \tag{*}$$

$u\in H$ — real Hilbert space,
$A\in L(H)$ — set of linear bounded operators on H
sp $A\subset[m,1]$, $m>0$
$A=A_1+A_2$
$A_1>0$, selfadjoint, $I+a_1A_1+a_2A_1^2+\ldots+a_NA_1^N=\mathbf{0}$
$A_2\geqslant 0$, compact, selfadjoint

(such an operator A_1 is called the algebraic operator).

The algorithm is the following

$$
\text{(a)}\begin{cases}
\text{0. Choose } u_0 \text{ — start point; } i=0.\\
\text{1. Compute } g_i \text{ — gradient of } Q \text{ in point } u_i.\\
\text{2. Find } s_i=g_i+\alpha_{i-1}s_{i-1};\ \alpha_{-1}=0, \text{ for } i>0\ \alpha_{i-1} \text{ is such, that } \langle s_i, As_{i-1}\rangle=0.\\
\text{3. Find } u_{i+1}=u_i+\eta_i s_i \text{ where } \eta_i \text{ minimizes the function } f(\eta)=Q(u_i+\eta s_i).\\
\text{4. } i:=i+1, \text{ go to 1.}
\end{cases}
$$

Let us denote $Q_i=Q(u_i)$.

THEOREM. $\sqrt[n]{Q_n}\to 0$ (superlinear convergence).

The theorem is proved with two lemmas.

LEMMA 1. $\lim\limits_{k\to\infty}\dfrac{Q_{k+N}}{Q_k}=0$.

Proof. It is a well known fact [1] that if we have directions $s_0, s_1, \ldots, s_{k-1}$, then $Q_k=\frac{1}{2}\min\limits_{z\in S_{k-1}}\langle g_0-z, A^{-1}(g_0-z)\rangle$ for $S_{k-1}=\text{lin}\,\{As_0, As_1, \ldots, As_{k-1}\}$. ($Q_k$ is a suboptimal value of Q on subspace S_{k-1}).

From algorithm (a) it is easily seen that

$$S_{k-1}=\text{lin}\,\{Ag_0, A^2 g_0, \ldots, A^k g_0\}$$

and

$$Q_k=\tfrac{1}{2}\min_{l_1\ldots l_k}\langle g_0+l_1 Ag_0+\ldots+l_k A^k g_0, A^{-1}(g_0+\ldots+l_k A^k g_0)\rangle=$$
$$=\langle g_0+c_1^{(k)}Ag_0+c_2^{(k)}A^2 g_0+\ldots+c_k^{(k)}A^k g_0, A^{-1}(\ldots)\rangle=\langle g_k, A^{-1}g_k\rangle.$$

When we change the coefficients $c_i^{(k)}$ is any way, the value of Q_k will increase, so

$$\frac{Q_{k+N}}{Q_k}=\frac{\langle g_0+c_1^{(k+N)}Ag_0+\ldots+c_{k+N}^{(k+N)}A^{k+N}g_0, A^{-1}(\ldots)\rangle}{\langle g_0+c_1^{(k)}Ag_0+\ldots+c_k^{(k)}A^k g_0, A^{-1}(\ldots)\rangle}\leqslant$$
$$\leqslant\frac{\langle (I+a_1 A+a_2 A^2+\ldots+a_N A^N)(g_0+c_1^{(k)}Ag_0+\ldots+c_k^{(k)}A^k g_0), A^{-1}(\ldots)\rangle}{\langle g_0+c_1^{(k)}Ag_0+\ldots+c_k^{(k)}A^k g_0, A^{-1}(\ldots)\rangle}=$$
$$=\frac{\langle Cg_k, A^{-1}Cg_k\rangle}{\langle g_k, A^{-1}g_k\rangle}$$

where

$$C=I+a_1 A+\ldots+a_N A^N=I+a_1(A_1+A_2)+a_2(A_1+A_2)^2+\ldots+a_N(A_1+A_2)^N=$$
$$=[I+a_1 A_1+\ldots+a_N A_1^N]+[a_1 A_2+a_2 A_1 A_2+a_2 A_2 A_1+a_2 A_2^2+\ldots].$$

The first bracket is equal **0** because of the assumption on operator A_1. The second bracket is a linear, selfadjoint and compact operator (because A_2 is compact). So C is compact operator.

Because $\operatorname{sp} A \subset [m, 1]$, then

$$\frac{1}{\langle g_k, A^{-1} g_k \rangle} \leqslant \frac{1}{\langle g_k, g_k \rangle}$$

and

$$\frac{Q_{k+N}}{Q_k} \leqslant \langle C\bar{g}_k, A^{-1} C\bar{g}_k \rangle, \ \bar{g}_k = \frac{g_k}{\|g_k\|}.$$

The sequence $(\bar{g}_k)_0^\infty$ is weakly convergent to $\mathbf{0}$ (because $\langle \bar{g}_i, \bar{g}_j \rangle = \delta_{ij}$).

Thus $C\bar{g}_k \to 0$ and $\dfrac{Q_{k+N}}{Q_k} \to 0$. Q.E.D.

LEMMA 2. If $\dfrac{Q_{k+N}}{Q_k} \to 0$, then $\sqrt[n]{Q_n} \to 0$.

Proof. Let $n = rN + s$, $0 \leqslant s < N$. Then

$$\sqrt[n]{Q_n} = \sqrt[n]{Q_0} \sqrt[n]{\frac{Q_n}{Q_0} \frac{Q_{2N}}{Q_N} \cdots \frac{Q_{rN}}{Q_{(r-1)N}}} \sqrt[n]{\frac{Q_n}{Q_{rN}}}.$$

For sufficiently large n we have $\sqrt[n]{Q_0} \leqslant 2$.

By assumption $\dfrac{Q_{rN}}{Q_{(r-1)N}} \xrightarrow[r\to\infty]{} 0$, so for any $\varepsilon \in (0, 1]$ there exists an integer r_0 such that for all $k \geqslant r_0$, $\dfrac{(Q)_{kN}}{Q_{(k-1)N}} \leqslant \varepsilon^N$.

Because

$$\frac{Q_n}{Q_{rN}} \leqslant 1, \ \frac{Q_{iN}}{Q_{(i-1)N}} \leqslant 1, \ i = 1, 2, \ldots, r_0,$$

then

$$\sqrt[n]{Q_n} \leqslant 2 \sqrt[n]{\frac{Q_{(r_0+1)N}}{Q_{r_0 N}} \frac{Q_{(r_0+2)N}}{Q_{(r_0+1)N}} \cdots \frac{Q_{rN}}{Q_{(r-1)N}}} \leqslant 2 \sqrt[n]{\varepsilon^{N(r-r_0)}} = \frac{2}{\sqrt[n]{\varepsilon^{Nr_0+s}}} \varepsilon.$$

For sufficiently large n

$$\sqrt[n]{\varepsilon^{Nr_0+s}} \geqslant \sqrt[n]{\varepsilon^{N(r_0+1)}} \geqslant \tfrac{1}{2}.$$

Finally, there exists such n_0 that for every $n \geqslant n_0$.

$$\sqrt[n]{Q_n} \leqslant 4\varepsilon.$$

This means $\sqrt[n]{Q_n} \to 0$. Q.E.D.

Convergence with property $\sqrt[n]{Q_n} \to 0$ is, surely, weaker than with $\dfrac{Q_{k+N}}{Q_k} \to 0$.

But the first property has a clearer interpretation. Namely, for any $\varepsilon > 0$ exists a n_0, such that for $n \geqslant 0$

$$\sqrt[n]{Q_n} < \varepsilon \sqrt[n]{Q_0} \quad \text{(because } \sqrt[n]{Q_0} \to 1\text{)}.$$

This means $Q_n < Q_0 \varepsilon^n$, i.e., starting from $n \geqslant n_0$ the sequence Q_n converges to 0 faster than a geometrical proression with ratio ε. So, we have superlinear convergence ($(Q_n)_0^\infty$ converges to 0 faster then any geometrical progression).

Remark 1. The property of superlinear convergence is not general. It is possible to show an operator A and gradient g_0 such that $(Q_n)_0^\infty$ has only linear convergence. An example

$(e_i)_0^\infty$ — orthonormal basis in H

$$Ae_k \stackrel{\Delta}{=} b_{k-1}\, e_{k-1} + a_k\, e_k + b_k\, e_{k+1},\ b_{-1}=0, \quad b_i = \frac{1-m}{4},$$

$$a_i = \frac{1+m}{2}, \quad i=0, 1, \ldots$$

$g_0 = e_0$.

It is possible to prove that the closure $\bar{A}$ of A is a positive, selfadjoint, linear operator such that when we minimize (*) we obtain $\frac{Q_{n+1}}{Q_n} \geqslant \left(\frac{1-m}{4}\right)^2$, $n=0, 1, \ldots$. For detailes see [5].

Remark 2. When $A=I+A_2$ where A_2 is compact, then $\frac{Q_{n+1}}{Q_n} \to 0$ (because $N=1$). If we can invert the operator A_1 (for problem (1) and (2) $A_1=R$ is a known matrix) we can obtain the above case by modifying the algoritm (a): we put

$$s_i = A_1^{-1} g_i + \alpha_{i-1}\, s_{i-1}.$$

In this way we shall have stronger superlinear convergence.

References

1. Daniel J. W., The conjugate gradient method for linear and nonlinear operator equations. *SIAM J. Numer. Anal.* 4 (1957).
2. Габасов Р., Кириллова Ф., Качественная теория оптимальных процессов. Москва 1971.
3. Hestenes M. R., The conjugate gradient method for solving linear systems. Proc. Symp. Appl. Math. 1956 vol. 6 p. 83—102.
4. Polak E., Computational methods in optimization. New York 1971.
5. Fortuna Z., The characterization of convergence of conjugate gradient method for quadratic functionals in Hilbert space (to appear).

Superliniowa zbieżność metody gradientu sprzężonego

Dla przypadku minimalizacji funkcjonału kwadratowego w rzeczywistej przestrzeni Hilberta, w którym Hessian ma postać $A=A_1+A_2$, gdzie A_1 jest operatorem algebraicznym, a A_2 jest operatorem zwartym otrzymano superliniową zbieżność metody gradientu sprzężonego. Przykładem takiego funkcjonału jest układ liniowy z opóźnieniem.

Суперлинейная сходимость метода сопряженного градиента

Для случая минимизации квадратического функционала в действительном гильбертовом пространстве, в котором определитель Гессе имеет вид $A=A_1+A_2$, где A_1 является оператором алгебраическим, а A_2 является компактным оператором, получена суперлинейная сходимость метода сопряженного градиента. Примером такого функционала является линейная система с запаздыванием.

Wskazówki dla Autorów

W wydawnictwie "Control and Cybernetics" drukuje się prace oryginalne nie publikowane w innych czasopismach. Zalecane jest nadsyłanie artykułów w języku angielskim. W przypadku nadesłania artykułu w języku polskim, Redakcja może zalecić przetłumaczenie na język angielski. Objętość artykułu nie powinna przekraczać 1 arkusza wydawniczego, czyli ok. 20 stron maszynopisu formatu A4 z zachowaniem interlinii i marginesu szerokości 5 cm z lewej strony. Prace należy składać w 2 egzemplarzach. Układ pracy i forma powinny być dostosowane do niżej podanych wskazówek.

1. W nagłówku należy podać tytuł pracy, następnie imię (imiona) i nazwisko (nazwiska) autora (autorów) w porządku alfabetycznym oraz nazwę reprezentowanej instytucji i nazwę miasta. Po tytule należy umieścić krótkie streszczenie pracy (do 15 wierszy maszynopisu).

2. Materiał ilustracyjny powinien być dołączony na oddzielnych stronach. Podpisy pod rysunki należy podać oddzielnie.

3. Wzory i symbole powinny być wpisane na maszynie bardzo starannie.

Szczególną uwagę należy zwrócić na wyraźne zróżnicowanie małych i dużych liter. Litery greckie powinny być objaśnione na marginesie. Szczególnie dokładnie powinny być pisane indeksy (wskaźniki) i oznaczenia potęgowe. Należy stosować nawiasy okrągłe.

4. Spis literatury powinien być podany na końcu artykułu. Numery pozycji literatury w tekście zaopatruje się w nawiasy kwadratowe. Pozycje literatury powinny zawierać nazwisko autora (autorów) i pierwsze litery imion oraz dokładny tytuł pracy (w języku oryginału), a ponadto:

a) przy wydawnictwach zwartych (książki) — miejsce i rok wydania oraz wydawcę;

b) przy artykułach z czasopism: nazwę czasopisma, numer tomu, rok wydania i numer bieżący.

Pozycje literatury radzieckiej należy pisać alfabetem oryginalnym, czyli tzw. grażdanką.

Recomendations for the Authors

Control and Cybernetics publishes original papers which have not previously appeared in other journals. The publications of the papers in English is recommended. No paper should exceed in length 20 type written pages (210 × 297 mm) with lines spaced and a 50 mm margin on the lefthand side. Papers should be submitted in duplicate. The plan and form of the paper should be as follows:

1. The heading should include the title, the full names and surnames of the authors in alphabetic order, the name of the institution he represents and the name of the city or town. This heading should be followed by a brief summary (about 15 typewritten lines).

2. Figures, photographs tables, diagrams should be enclosed to the manuscript. The texts related to the figures should be typed on a separate page.

3. Of possible all mathematical expressions should be typewritten. Particular attention should be paid to differentiation between capital and small letters. Greek letters should as a rule be defined. Indices and exponents should be written with particular care. Round brackets should not be replaced by an inclined fraction line.

4. References should be put on the separate page. Numbers in the text identified by references should be enclosed in brackets. This should contain the surname and the initials of Christian names, of the author (or authors), the complete title of the work (in the original language) and, in addition:

a) for boks — the place and the year of publication and the publisher's name;

b) for journals — the name of the journal, the number of the volume, the year of the publication, and the ordinal number.