ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 21 Март 1981 Апрель № 2

УДК 519:614.4

ОБ АСИМПТОТИЧЕСКИХ СВОЙСТВАХ МЕТОДА НАИСКОРЕЙШЕГО СПУСКА В ЗАДАЧАХ НА СОБСТВЕННЫЕ ЗНАЧЕНИЯ

ЖУК П. Ф.

(Киев)

Исследуется асимптотическое поведение нормированных градиентов метода наискорейшего спуска при определении наименьшего собственного значения положительно-определенного и самосопряженного конечномерного оператора, и на основе полученных результатов обосновываются приемы ускорения его сходимости, а также предложен способ одновременного отыскания этим методом трех собственных значений оператора.

Введение

Изучение асимптотического поведения итерационных процессов часто позволяет более полно использовать их внутренние резервы. Так, например, в [1] на основе изучения асимптотического поведения метода наискорейшего спуска при решении систем линейных алгебраических уравнений разработаны эффективные и легко реализуемые на ЭВМ приемы ускорения его сходимости. В [2] на основе работ [3], [4] об асимптотическом поведении нормированных градиентов метода наискорейшего спуска при решении систем линейных алгебраических уравнений обосновывается возможность применения к нему известного приема А. А. Абрамова ускорения сходимости линейных итерационных процессов.

В настоящей работе рассматриваются вопросы сходимости и асимптотического поведения метода наискорейшего спуска для задач на собственные значения

$$(1) \qquad Av=\lambda v,$$

где A — положительно-определенный и самосопряженный линейный оператор, действующий в n-мерном действительном гильбертовом пространстве H со скалярным произведением, обозначенным через (v, u), и нормой $\|v\|^2=(v, v)$, а $\lambda_1\leqslant\ldots\leqslant\lambda_n$ и $\{e_i\}_1^n$ — собственные значения и соответствующая им ортонормированная система собственных векторов оператора A. Для простоты изложения рассмотрен случай не кратных собственных значений λ_i.

Вычислительная схема метода наискорейшего спуска для отыскания λ_1 представима формулами

$$(2) \qquad \tilde v_0\neq 0, \qquad v_k=\tilde v_k/\|\tilde v_k\|, \qquad \mu_k=(Av_k, v_k),$$

$$(3) \qquad w_k=Av_k-\mu_k v_k, \qquad \tilde v_{k+1}=v_k-\tau_k w_k,$$

где множитель τ_k выбирается из условия минимума μ_{k+1}.

§ 1. Условие стабилизации метода наискорейшего спуска

Стабилизация метода наискорейшего спуска означает существование такого конечного $k=0,1,\ldots$, при котором $\mu_k=\lambda_1$. Основной задачей данного параграфа является нахождение необходимого и достаточного условия стабилизации. Для простоты изложения здесь и в дальнейшем предполагаем, что $\lambda_1<\ldots<\lambda_n$.

Укажем на некоторые известные свойства метода наискорейшего спуска [5], [6], необходимые для дальнейших исследований.

1. Для каждого $k=0, 1,\ldots$, при котором $w_k\neq 0$, имеют место соотношения

$$(1.1)\qquad (w_k, v_k)=0,$$

$$(1.2)\qquad \tau_k=2\{q_k-\mu_k+[(q_k-\mu_k)^2+4\|w_k\|^2]^{1/2}\}^{-1},$$

$$(1.3)\qquad \tau_k=(q_k-\mu_{k+1})^{-1},$$

$$(1.4)\qquad \mu_k-\mu_{k+1}=\tau_k\|w_k\|^2,$$

$$(1.5)\qquad (w_k, w_{k+1})=0,$$

где $q_k=(Aw_k, w_k)/(w_k, w_k)$.

2. Если $\lambda_1<\lambda_2$ и начальное приближение таково, что $\lambda_1<\mu_0<\lambda_2$, то последовательность μ_k сходится к λ_1 со скоростью геометрической прогрессии и справедлива оценка

$$\mu_{k+1}-\lambda_1\leqslant\rho^2(\mu_k-\lambda_1),\qquad \rho=\frac{1-\xi}{1+\xi},\quad \xi=\frac{\lambda_2-\mu_0}{\lambda_n-\lambda_1}.$$

Пусть α_{ik} — коэффициенты разложения вектора v_k по системе собственных векторов $\{e_i\}_1^n$ оператора A:

$$v_k=\sum_{i=1}^n \alpha_{ik}e_i.$$

Имеет место

Теорема 1. *Если $\mu_0<\lambda_2$, то метод наискорейшего спуска стабилизируется тогда и только тогда, когда вектор v_0 имеет вид*

$$v_0=\alpha_{10}e_1+\alpha_{p0}e_p,\qquad p=1, 2,\ldots, n.$$

Обозначим через U множество векторов $u=(u_1,\ldots, u_n)$ с неотрицательными компонентами, удовлетворяющими условиям

$$\sum_{i=1}^n u_i=1,\qquad \sum_{i=1}^n \lambda_i u_i=\mu,$$

а через U_1 — множество векторов $u^{(i)}=(u_{1i}, 0,\ldots, u_i^1, 0,\ldots, 0)$, $i=2, 3,\ldots, n$,

$$u_{1i}=\frac{\lambda_i-\mu}{\lambda_i-\lambda_1},\qquad u_i^1=\frac{\mu-\lambda_1}{\lambda_i-\lambda_1},\qquad \lambda_1<\mu<\lambda_2.$$

Докажем предварительно одно вспомогательное предложение.

Лемма. *Пусть*

$$\Phi(u)=\sum_{i=1}^{n} d_i u_i,$$

где $\{d_i\}_1^n$ — *вещественные числа. Тогда*

$$\max_{u\in U}\Phi(u)=\max_{u\in U_1}\Phi(u), \qquad \min_{u\in U}\Phi(u)=\min_{u\in U_1}\Phi(u).$$

Доказательство леммы. Достаточно показать, что U_1 — множество всех вершин многогранника U. Ясно, что $U_1\subseteq U$ и U_1 — линейно-независимое множество векторов. Пусть $u\in U$. Покажем, что существуют такие неотрицательные числа $\{\theta_i\}_{i=2}^{n}$, что выполняются равенства

$$\sum_{i=2}^{n}\theta_i=1, \qquad u=\sum_{i=2}^{n}\theta_i u^{(i)}. \tag{1.6}$$

Запишем (1.6) в виде

$$u_1=\sum_{i=2}^{n}\theta_i u_{1i}, \qquad u_i=\theta_i u_i^1, \quad i=2,3,\ldots,n. \tag{1.7}$$

Из (1.7) можно определить числа θ_i:

$$\theta_i=u_i/u_i^1\geqslant 0. \tag{1.8}$$

Обозначим через $\bar{u}_1$ выражение

$$\bar{u}_1=\sum_{i=2}^{n}\theta_i u_{1i},$$

где θ_i определены по формуле (1.8). Нетрудно убедиться в справедливости соотношений

$$\bar{u}_1=\sum_{i=2}^{n}\theta_i-1+u_1, \qquad \lambda_1(\bar{u}_1-u_1)=\mu\left(\sum_{i=2}^{n}\theta_i-1\right). \tag{1.9}$$

Поскольку $\lambda_1<\mu<\lambda_2$, то из (1.9) следует

$$\sum_{i=2}^{n}\theta_i=1, \qquad u_1=\bar{u}_1.$$

Лемма доказана.

Доказательство теоремы 1. Достаточно показать, что если

$$v_0=\alpha_{10}e_1+\alpha_{i0}e_i+\alpha_{j0}e_j, \qquad \alpha_{i0}\neq 0, \quad \alpha_{j0}\neq 0, \quad 1<i<j,$$

то метод наискорейшего спуска не стабилизируется. Без ограничения общности можно считать $i=2$, $j=n$. В силу формул (2), (3) вычислительной схемы метода наискорейшего спуска, имеем

$$\alpha_{2,\,k+1}=(1-\tau_k(\lambda_2-\mu_k))\alpha_{2k}\|\tilde v_{k+1}\|^{-1},$$

$$\alpha_{n,\,k+1}=(1-\tau_k(\lambda_n-\mu_k))\alpha_{nk}\|\tilde v_{k+1}\|^{-1},$$

где $\|\tilde v_{k+1}\|^2=1+\tau_k^2\|w_k\|^2\neq 0$. Следовательно, для доказательства теоремы достаточно показать, что для любого $k=0,\ 1,\dots$ выполнены неравенства

$$(1.10)\qquad \tau_k^{-1}+\mu_k\neq\lambda_2,\qquad \tau_k^{-1}+\mu_k\neq\lambda_n,$$

$$(1.11)\qquad \tau_k\neq\infty.$$

Неравенство (1.11) будет выполнено, если только $w_k\neq 0$. Действительно, в силу (1.4),

$$\mu_k-\mu_{k+1}=\tau_k\|w_k\|^2\leqslant\mu_k-\lambda_1,$$

следовательно,

$$\tau_k\leqslant(\mu_k-\lambda_1)\|w_k\|^{-2}.$$

Справедливость неравенств (1.10) доказывается от противного.

Предположим сначала, что существует такой номер $\bar k$, для которого имеют место соотношения

$$(1.12)\qquad \tau_{\bar k}^{-1}+\mu_{\bar k}=\lambda_2,\qquad \alpha_{2\bar k}\neq 0,\qquad \alpha_{n\bar k}\neq 0.$$

Используя (1.3), (1.4), преобразуем (1.12) к виду

$$q_{\bar k}+\tau_{\bar k}\|w_{\bar k}\|^2=\lambda_2,$$

или, полагая $\tau_{\bar k}^{-1}=\lambda_2-\mu_{\bar k}$, имеем

$$(1.13)\qquad q_{\bar k}+\|w_{\bar k}\|^2(\lambda_2-\mu_{\bar k})^{-1}=\lambda_2.$$

Разлагая вектор $w_{\bar k}$ по системе собственных векторов $\{e_i\}_1^n$, запишем $\|w_{\bar k}\|^2$ в виде

$$\|w_{\bar k}\|^2=\sum_{i=1}^n(\lambda_i-\mu_{\bar k})^2\alpha_{i\bar k}^2.$$

Если в качестве функции $\Phi(u)$ леммы взять функцию

$$\Phi(u)=\sum_{i=1}^n(\lambda_i-\mu_{\bar k})^2u_i,$$

а в качестве μ положить $\mu_{\bar k}$, то на основании этой леммы заключаем, что

$$(1.14)\qquad \|w_{\bar k}\|^2>\min_{u\in U_1}\Phi(u)=(\mu_{\bar k}-\lambda_1)(\lambda_2-\mu_{\bar k}).$$

Выполнение строгого неравенства в (1.14) обеспечено тем, что $\alpha_{n\bar k}\neq 0$. Используя соотношения (1.13), (1.14), получаем

$$(1.15)\qquad q_{\bar k}+\mu_{\bar k}<\lambda_1+\lambda_2.$$

С другой стороны, полагая

$$\Phi(u)=\sum_{i=1}^{n}(\lambda_i-q_{\bar k})(\mu_{\bar k}-\lambda_i)^2 u_i,\qquad \mu=\mu_{\bar k},$$

в силу леммы имеем

$$\max_{i=2,3,\dots,n}[(\mu_{\bar k}-\lambda_1)(\lambda_i-\mu_{\bar k})(\mu_{\bar k}+q_{\bar k}-\lambda_1-\lambda_i)]\geqslant 0. \tag{1.16}$$

Так как $\mu_{\bar k}<\lambda_2$, то из (1.16) следует неравенство $\mu_{\bar k}+q_{\bar k}\geqslant\lambda_1+\lambda_2$, противоречащее (1.15).

Предположим теперь, что существует такой номер $\bar k$, для которого выполнены соотношения

$$\tau_{\bar k}^{-1}+\mu_{\bar k}=\lambda_n,\qquad \alpha_{2\bar k}\neq 0,\qquad \alpha_{n\bar k}\neq 0. \tag{1.17}$$

Рассмотрим в качестве функции $\Phi(u)$ леммы следующее выражение:

$$\Phi(u)=\sum_{i=1}^{n}(\lambda_i-\mu_{\bar k+1})(\lambda_i-\mu_{\bar k})[1-\tau_{\bar k}(\lambda_i-\mu_{\bar k})]u_i,$$

где $u\in U$, $\mu=\mu_{\bar k}$. Если положить $\bar u=(\alpha_{1\bar k}^2,\dots,\alpha_{n\bar k}^2)$, то $\Phi(\bar u)=0$. Так как $\alpha_{2\bar k}\neq 0$, то, в силу леммы, имеем

$$\Phi(\bar u)>\min_{u\in U}\Phi(u)=\tau_{\bar k}(\lambda_n-\mu_{\bar k})(\mu_{\bar k}-\lambda_1)(\tau_{\bar k}^{-1}+\mu_{\bar k}+\mu_{\bar k+1}-\\ -\lambda_1-\lambda_n). \tag{1.18}$$

Из (1.18) получаем $\tau_{\bar k}^{-1}+\mu_{\bar k}<\lambda_n-\mu_{\bar k+1}+\lambda_1<\lambda_n$, что противоречит условию (1.17). Теорема доказана.

§ 2. О скорости сходимости

Будем говорить, что последовательность μ_k сходится к λ_1 почти линейно, если существуют такие постоянные $1>c_1\geqslant c_2>0$, что для любого $k=0,\ 1,\dots$ выполнено неравенство

$$c_1\geqslant\frac{\mu_{k+1}-\lambda_1}{\mu_k-\lambda_1}\geqslant c_2.$$

Теорема 2. *Если $\bar v_0$ — такое начальное приближение, что выполнены условия*

$$\lambda_1<\mu_0<\lambda_2,\qquad \alpha_{i0}\neq 0,\qquad \alpha_{j0}\neq 0,\qquad 1<i<j, \tag{2.1}$$

то последовательность μ_k сходится к λ_1 почти линейно.

Доказательство. В силу того что нулевые коэффициенты α_{p0} не оказывают влияния на характер сходимости метода наискорейшего спуска, без ограничения общности можно считать, что в условии (2.1) $i=2$, $j=n$. Образуем следующие последовательности векторов:

$$\bar v_0=v_0,\qquad \bar v_{k+1}=\tau_k^{-1}\bar v_k-\bar w_k,\qquad \bar w_k=A\bar v_k-\mu_k\bar v_k,\qquad k=0,\ 1,\dots .$$

Нетрудно видеть, что выполнены равенства

$$\bar{v}_k = v_k \|\bar{v}_k\|, \qquad \bar{w}_k = w_k \|\bar{v}_k\|,$$

поэтому, используя соотношения (1.1), (1.4), (1.5), имеем

$$(2.2) \qquad (\bar{w}_k, \bar{v}_k) = 0, \qquad (\bar{w}_k, \bar{w}_{k+1}) = 0, \qquad \mu_k - \mu_{k+1} = \tau_k \frac{\|\bar{w}_k\|^2}{\|\bar{v}_k\|^2}.$$

Для $k=0, 1, \dots$ положим

$$\beta_k = \|\bar{w}_{k+1}\| / \|\bar{w}_k\|.$$

Докажем, что последовательность β_k сходится и

$$\beta = \lim_{k\to\infty} \beta_k \neq 0.$$

Рассмотрим скалярное произведение $(\bar{w}_k, \bar{w}_{k+2})$. Используя (2.2), получаем

$$(2.3) \qquad (\bar{w}_k, \bar{w}_{k+2}) = (\bar{w}_k, A\bar{v}_{k+2} - \mu_{k+2}\bar{v}_{k+2}) = (\bar{w}_k, A\bar{v}_{k+2} - \mu_{k+1}\bar{v}_{k+2}) +$$
$$+ (\mu_{k+1} - \mu_{k+2})(\bar{w}_k, \bar{v}_{k+2}) = -(\bar{w}_k, A\bar{w}_{k+1}) - \|\bar{w}_{k+1}\|^2 \|\bar{w}_k\|^2 \|\bar{v}_{k+1}\|^{-2}.$$

Аналогично преобразуем (w_{k+1}, w_{k+1}):

$$(2.4) \qquad (\bar{w}_{k+1}, \bar{w}_{k+1}) = (\bar{v}_{k+1}, A\bar{w}_{k+1} - \mu_k \bar{w}_{k+1}) = -(\bar{w}_k, A\bar{w}_{k+1}).$$

Используя неравенство Коши — Буняковского

$$|(\bar{w}_k, \bar{w}_{k+2})| \leqslant \|\bar{w}_k\| \|\bar{w}_{k+2}\|$$

и соотношения (2.3), (2.4), имеем

$$(2.5) \qquad \beta_{k+1} \geqslant \beta_k (1 - \varepsilon_k), \qquad k = 0, 1, \dots,$$

где $\varepsilon_k = \|\bar{w}_k\|^2 \|\bar{v}_{k+1}\|^{-2}$.

В дальнейшем нужны оценки для $\|\bar{w}_k\|^2$, $\|\bar{v}_{k+1}\|^2$. Так как коэффициенты α_{ik} разложения вектора v_k по системе собственных векторов $\{e_i\}_1^n$ удовлетворяют условиям

$$\sum_{i=1}^{n} \alpha_{ik}^2 = 1, \qquad \sum_{i=1}^{n} \lambda_i \alpha_{ik}^2 = \mu_k,$$

то нетрудно видеть, что

$$(2.6) \qquad (\lambda_2 - \mu_k)^2 \frac{\mu_k - \lambda_1}{\lambda_n - \lambda_1} \|\bar{v}_k\|^2 \leqslant \|\bar{w}_k\|^2 \leqslant (\lambda_n - \mu_k)^2 \frac{\mu_k - \lambda_1}{\lambda_2 - \lambda_1} \|\bar{v}_k\|^2.$$

Так как $\|\bar{v}_{k+1}\|^2 = \tau_k^{-2} \|\bar{v}_k\|^2 + \|\bar{w}_k\|^2$, то для $\|\bar{v}_{k+1}\|^2$ справедлива оценка

$$(2.7) \qquad \tau_k^{-2} \|\bar{v}_k\|^2 < \|\bar{v}_{k+1}\|^2 < (\tau_k^{-2} + \lambda_n^2) \|\bar{v}_k\|^2.$$

Кроме того, в силу теоремы 1, имеет место неравенство

$$(2.8) \qquad \lambda_2 - \mu_k < \tau_k^{-1} < \lambda_n - \mu_k.$$

Из (2.6)—(2.8) следует, что для $k=0, 1, \ldots$

$$(2.9) \qquad \beta_k^2 < 2\left(\frac{\lambda_n-\lambda_1}{\lambda_2-\mu_0}\right)^3 \lambda_n^2;$$

таким образом, последовательность β_k ограничена сверху.

Рассмотрим бесконечное произведение

$$P=\prod_{i=0}^{\infty}(1-\varepsilon_i),$$

где ε_i определены в (2.5).

Используя (2.6)—(2.8), нетрудно показать, что для $k=0, 1, \ldots$

$$(2.10) \qquad 0<\varepsilon_k<1, \qquad \varepsilon_k<\frac{(\lambda_n-\lambda_1)^2(\mu_k-\lambda_1)}{(\lambda_2-\lambda_1)(\lambda_2-\mu_0)^2}.$$

В силу свойства 2 метода наискорейшего спуска, имеем

$$\sum_{k=0}^{\infty}(\mu_k-\lambda_1)<\infty,$$

поэтому из неравенств (2.10) вытекает сходимость бесконечного произведения P.

Образуем новую последовательность β_k^* по правилу

$$\beta_k^*=\beta_k\prod_{i=0}^{k-1}(1-\varepsilon_i)^{-1}, \qquad k=1, 2, \ldots .$$

Используя неравенство (2.5), получаем

$$\beta_{k+1}^*=\beta_{k+1}\prod_{i=0}^{k}(1-\varepsilon_i)^{-1}\geqslant\beta_k\prod_{i=0}^{k-1}(1-\varepsilon_i)^{-1}=\beta_k^*;$$

таким образом, последовательность β_k^* неубывающая. Из сходимости бесконечного произведения P и ограниченности сверху последовательности β_k следует ограниченность сверху последовательности β_k^*. Пусть

$$\beta^*=\lim_{k\to\infty}\beta_k^*;$$

тогда последовательность β_k имеет предел и

$$(2.11) \qquad \beta=\lim_{k\to\infty}\beta_k=\beta^*P>0.$$

Пусть $0<\delta<\beta$ и $\tilde{k}$ — такой номер последовательности β_k, что для всех $k\geqslant\tilde{k}$ имеет место неравенство $\beta_k\geqslant\beta-\delta$. Используя (2.6)—(2.8), получаем, что для $k=0, 1, \ldots$

$$\beta_k^2<\frac{\mu_{k+1}-\lambda_1}{\mu_k-\lambda_1}\,\frac{(\lambda_n-\lambda_1)^3(\tau_k^{-2}+\lambda_n^2)}{(\lambda_2-\mu_0)^2(\lambda_2-\lambda_1)},$$

следовательно, для всех $k\geqslant\tilde{k}$ имеем

$$\frac{\mu_{k+1}-\lambda_1}{\mu_k-\lambda_1}>\frac{(\lambda_2-\mu_0)^2(\lambda_2-\lambda_1)(\beta-\delta)^2}{2(\lambda_n-\lambda_1)^3\lambda_n^2}=d,$$

Поэтому в качестве постоянной c_2 для последовательности μ_k можно взять

$$c_2=\min\left(d, \min_{i=0,1,\ldots,\tilde{k}} \frac{\mu_{i+1}-\lambda_1}{\mu_i-\lambda_1}\right).$$

В качестве c_1 можно взять ρ^2, где ρ определено в свойстве 2 метода наискорейшего спуска. Теорема доказана.

Таким образом, последовательность μ_k при естественных ограничениях сходится к λ_1 почти линейно. Этот факт усиливает результат работы [7] и утверждает, что метод наискорейшего спуска сходится не более чем со скоростью геометрической прогрессии.

Замечание. Пусть x_0, x_1, ... — последовательные приближения при минимизации функционала

$$f(x)=0.5(Ax, x), \qquad A=A^*>0, \quad x\in H,$$

s-шаговым методом наискорейшего спуска. В [3] показано, что если $x_1\neq 0$, то последовательность $f(x_k)$ сходится к нулю почти линейно (для краткости изложения будем использовать здесь обозначения из [3]). При доказательстве почти линейной сходимости $f(x_k)$ к нулю в [3] существенно использовался тот факт, что для любого предельного вектора r последовательности нормированных градиентов $\{y_{2k}\}_0^\infty$ имеет место равенство

$$T^2r=r. \tag{2.12}$$

В § 3 данной работы будет показано аналогичное равенство для предельных векторов последовательности нормированных градиентов метода наискорейшего спуска в задачах на собственные значения, однако при этом существенную роль играет теорема 2.

Покажем, что для доказательства почти линейной сходимости s-шагового метода наискорейшего спуска не обязательно использовать равенство (2.12). Действительно, из [3] следует, что

$$\|w_{k+2}\|^2\|w_{k+1}\|^{-2}\geqslant\|w_{k+1}\|^2\|w_k\|^{-2}, \qquad k=0, 1, \ldots,$$

поэтому для $k=1, 2, \ldots$ справедливо неравенство

$$\|w_{k+1}\|^2\|w_k\|^{-2}\geqslant\|w_1\|^2\|w_0\|^{-2}>0.$$

Так как

$$\frac{f(x_{k+1})}{f(x_k)}=\frac{(A^{-1}z_{k+1}, z_{k+1})}{(A^{-1}z_k, z_k)}, \tag{2.13}$$

где

$$z_{k+1}=P_s(A, z_k)\gamma_s^{(k)}z_k, \qquad z_k=Ax_k, \tag{2.14}$$

$$P_s(A, z_k)=A^s+\frac{\gamma_{s-1}^{(k)}}{\gamma_s^{(k)}}A^{s-1}+\ldots+\frac{\gamma_1^{(k)}}{\gamma_s^{(k)}}A+\frac{1}{\gamma_s^{(k)}}E$$

есть матричный многочлен с коэффициентами $\gamma_i^{(k)}$, зависящими от z_k, имеющий различные вещественные корни в открытом интервале (λ_1, λ_n), а $w_0=z_0$, $w_{k+1}=P_s(A, z_k)w_k$, то из (2.14) вытекает, что для $k=0, 1, \ldots$

$$\|z_{k+1}\|^2/\|z_k\|^2>\|w_{k+1}\|^2/\lambda_n^{2s}\|w_k\|^2. \tag{2.15}$$

Используя (2.13), (2.15), получаем

$$\frac{f(x_{k+1})}{f(x_k)} > \frac{\lambda_1}{\lambda_n}\frac{\| z_{k+1}\|^2}{\| z_k\|^2} > \frac{\lambda_1}{\lambda_n^{2s+1}}\frac{\| w_1\|^2}{\| w_0\|^2},$$

и в качестве постоянной c_2 для последовательности $f(x_k)$ можно взять

$$c_2 = \frac{\lambda_1}{\lambda_n^{2s+1}}\frac{\| w_1\|^2}{\| w_0\|^2}. \tag{2.16}$$

Приведенное доказательство проще, чем доказательство того же факта в [3], и, кроме того, позволяет эффективно находить c_2 с помощью (2.16).

§ 3. Асимптотическое поведение нормированных градиентов

1. Пусть $\{z_k\}_0^\infty$ — последовательность нормированных градиентов метода наискорейшего спуска, $z_k=w_k/\|w_k\|$. Обозначим через V множество векторов $v\in H$ вида $v=\xi e_2+\eta e_n$, $\xi\neq 0$, $\eta\neq 0$. Имеет место

Теорема 3. *Если $\tilde{v}_0$ — такое начальное приближение, что $\mu_0<\lambda_2$, $\alpha_{20}\neq 0$, $\alpha_{n0}\neq 0$, то существуют пределы*

$$\lim_{k\to\infty} z_{2k}=z\in V, \qquad \lim_{k\to\infty} z_{2k+1}=\bar{z}\in V.$$

Доказательство. Обозначим через T оператор

$$Tv=\mu(v)v-Av,$$

где $\mu(v)=(Av, v)/(v, v)$.

Пусть z — предельный вектор последовательности $\{z_{2k}\}_0^\infty$, т. е. существует подпоследовательность $\{z_{2k_i}\}_{i=0}^\infty$ такая, что

$$z=\lim_{i\to\infty} z_{2k_i}.$$

Покажем, что имеет место равенство

$$z=T^2z/\|T^2z\|, \tag{3.1}$$

где $T^2z=T(Tz)$.

Обозначим через z^* правую часть соотношения (3.1). В силу (2.11), имеем

$$\beta_{k+1}-\beta_k=\beta_{k+1}\left[1-\frac{(z_k, z_{k+2})}{1-\varepsilon_k}\right]\to 0,$$

следовательно,

$$\lim_{k\to\infty} (z_k, z_{k+2})=1. \tag{3.2}$$

Рассмотрим последовательность $\{z_{2k_i+1}\}_{i=0}^\infty$. Легко убедиться в справедливости соотношений

$$z_{2k_i+1}=[(\tau_{2k_i}^{-1}+\mu_{2k_i})z_{2k_i}-Az_{2k_i}]\beta_{2k_i}^{-1}-(\mu_{2k_i}-\mu_{2k_i+1})\frac{\bar{v}_{2k_i+1}}{\|\bar{w}_{2k_i+1}\|}, \tag{3.3}$$

$$\lim_{i\to\infty}(\tau_{2k_i}^{-1}+\mu_{2k_i})=(Az, z). \tag{3.4}$$

В силу (2.6), имеем

$$(3.5)\qquad \frac{(\mu_{2k_i}-\mu_{2k_i+1})\,\|\bar{v}_{2k_i+1}\|}{\|\bar{w}_{2k_i+1}\|}\leqslant\frac{[(\lambda_n-\lambda_1)(\mu_{2k_i+1}-\lambda_1)]^{1/2}(\mu_{2k_i}-\lambda_1)}{(\lambda_2-\mu_0)(\mu_{2k_i+1}-\lambda_1)}.$$

Оценим правую часть этого неравенства. На основании теоремы 2, существует постоянная $c_2>0$ такая, что для $k=0, 1, \ldots$ выполнено неравенство $(\mu_{k+1}-\lambda_1)(\mu_k-\lambda_1)^{-1}\geqslant c_2$, следовательно, для $i=0, 1, \ldots$ справедливо

$$(3.6)\qquad (\mu_{2k_i+1}-\lambda_1)/(\mu_{2k_i}-\lambda_1)\geqslant c_2.$$

Используя (3.5), (3.6), убеждаемся, что

$$\lim_{i\to\infty}\frac{\|\bar{v}_{2k_i+1}\|}{\|\bar{w}_{2k_i+1}\|}(\mu_{2k_i}-\mu_{2k_i+1})=0.$$

Тогда из (3.3), (3.4) следует

$$\lim_{i\to\infty} z_{2k_i+1}=\beta^{-1}Tz.$$

Применяя предыдущие рассуждения к последовательности $\{z_{2k_i+1}\}_{i=0}^{\infty}$, убеждаемся в справедливости равенства

$$\lim_{i\to\infty} z_{2k_i+2}=\beta^{-2}T^2z.$$

Следовательно,

$$z^*=\lim_{i\to\infty} z_{2k_i+2}.$$

Пользуясь (3.2), имеем

$$\lim_{i\to\infty}(z_{2k_i}, z_{2k_i+2})=(z, z^*)=1,$$

откуда вытекает требуемое равенство $z=z^*$.

В [3] доказано, что если вектор $v\in H$ удовлетворяет равенству $v=T^2v/\|T^2v\|$, то он представляется в виде

$$v=\xi e_i+\eta e_j,\qquad i<j,\quad \xi\neq 0,\qquad \eta\neq 0,$$

следовательно,

$$(3.7)\qquad z=\xi e_i+\eta e_j.$$

Докажем, что $i=2$, $j=n$. Предположим, что $j<n$, и обозначим через $\bar{z}$ вектор $\bar{z}=Tz/\|Tz\|$. В силу (3.7), имеем $q=\mu(z)<\lambda_j$, $\bar{q}=\mu(\bar{z})<\lambda_j$, следовательно,

$$(3.8)\qquad \rho_j=(\bar{q}-\lambda_j)(q-\lambda_j)<(\bar{q}-\lambda_n)(q-\lambda_n)=\rho_n.$$

Пусть

$$z_k=\sum_{i=1}^{n}\theta_{ik}e_i.$$

Используя (3.3), (3.4), нетрудно показать, что

$$\theta_{j,2k_i+2}=c_i(\rho_j+\varepsilon_i^j)\theta_{j\,2k_i},$$

(3.9)

$$\theta_{n,2k_i+2}=c_i(\rho_n+\varepsilon_i^n)\theta_{n\,2k_i},$$

где $\varepsilon_i^j\to 0$, $\varepsilon_i^n\to 0$ при $i\to\infty$, а c_i — некоторое положительное число. Тогда из (3.8) следует

$$\lim_{i\to\infty}\theta_{j\,2k_i}\theta_{n\,2k_i}^{-1}=0,$$

что противоречит предположению. Аналогично доказывается, что $i=2$. В силу (3.7) имеем

$$\beta^2=\lim_{i\to\infty}\beta_{2k_i}^2=\|Tz\|^2=(\lambda_n-\lambda_2)^2\xi^2\eta^2. \tag{3.10}$$

Так как $\xi^2+\eta^2=1$, то уравнение (3.10) при заданном β может иметь не более восьми решений, поэтому множество предельных векторов последовательности $\{z_{2k}\}_0^\infty$ содержит не более восьми векторов. Однако в силу (3.2) это множество должно быть замкнутым и связным [3], следовательно, оно состоит из одного вектора $z\in V$.

Вторая часть теоремы следует из того, что

$$\lim_{k\to\infty}z_{2k+1}=Tz/\|Tz\|=\bar{z}\in V.$$

Теорема доказана.

2. Пусть $\mu_0<\lambda_2$, $\alpha_{20}\neq 0$, $\alpha_{n0}\neq 0$, а z, $\bar{z}$, q, $\bar{q}$ — пределы соответствующих последовательностей z_{2k}, z_{2k+1}, q_{2k}, q_{2k+1} и $z=\theta_2e_2+\theta_ne_n$, $c=(Az,\bar{z})$. Относительно λ_2, λ_n, θ_2, θ_n получим следующую систему уравнений:

$$\theta_2^2+\theta_n^2=1,\quad \lambda_2\theta_2^2+\lambda_n\theta_n^2=q,\qquad \pm(\lambda_n-\lambda_2)\theta_2\theta_n=c,$$

$$\lambda_2\theta_n^2+\lambda_n\theta_2^2=\bar{q}.$$

Отсюда находим λ_2 и λ_n:

$$\lambda_2=\{q+\bar{q}-[(q-\bar{q})^2+4c^2]^{1/2}\}/2,\qquad \lambda_n=q+\bar{q}-\lambda_2.$$

Так как q, $\bar{q}$, c нам не известны, то строим последовательные приближения λ_2^k к λ_2, λ_n^k к λ_n.

Пусть для $k=0,1,\ldots$

$$c_k=(Az_k,z_{k+1}),$$

$$\lambda_2^k=\{q_k+q_{k+1}-[(q_k-q_{k+1})^2+4c_k^2]^{1/2}\}/2,$$

$$\lambda_n^k=q_k+q_{k+1}-\lambda_2^k.$$

Тогда

$$\lambda_2=\lim_{k\to\infty}\lambda_2^k,\qquad \lambda_n=\lim_{k\to\infty}\lambda_n^k,$$

причем скорость сходимости λ_2^k к λ_2, λ_n^k к λ_n определяется скоростью сходимости z_{2k} к z, z_{2k+1} к $\bar{z}$.

Преимуществом предложенного приема одновременного определения λ_1, λ_2, λ_n с помощью метода наискорейшего спуска является то, что не требуется дополнительной вычислительной работы на промежуточных операциях.

Для практического применения данного приема существенна характеристика скорости сходимости z_{2k} к z, z_{2k+1} к $\bar{z}$. Оценим скорость сходимости z_{2k} к z (для z_{2k+1} оценка аналогична).

Пусть $z_{2k}=\theta_{2,\,2k}e_2+\theta_{n,\,2k}e_n+y_{2k}$, $y_{2k}\perp V$. В качестве меры близости z_{2k} и z возьмем $\|y_{2k}\|$. Обозначим через I множество таких индексов i, $i\neq 1$, что α_{i0} — невырождающийся коэффициент (т. е. $\alpha_{ik}\neq 0$, $k=0,1,\ldots$), а

$$\sigma=\min_{i\in I}\left|\lambda_i-\frac{\lambda_2+\lambda_n}{2}\right|.$$

Теорема 4. *Если* $\mu_0<\lambda_2$, $\alpha_{20}\neq 0$, $\alpha_{n0}\neq 0$,

$$|q-\bar{q}|\neq\left[\frac{(\lambda_n-\lambda_2)^2}{2}+2\sigma^2\right]^{1/2},$$

то существуют такие $\rho<1$, a_1, *зависящие от* v_0, *что для* $k=0,1,\ldots$ *выполнено неравенства* $\|y_{2k}\|\leqslant a_1\rho^k$.

Доказательство. Из (3.1) следует

$$\rho_2=\rho_n,\qquad \max_{i\in I,\ 2<i<n}|\rho_i|\leqslant\rho_2, \tag{3.11}$$

где $\rho_i=(\bar{q}-\lambda_i)(q-\lambda_i)$. Неравенство (3.11) эквивалентно неравенству

$$|q-\bar{q}|\leqslant\left[\frac{(\lambda_n-\lambda_2)^2}{2}+2\sigma^2\right]^{1/2}, \tag{3.12}$$

причем в случае строгого неравенства в (3.12) выполнено строгое неравенство в (3.11). Поэтому существует такое $\delta>0$, что

$$\max_{i\in I,\ 2<i<n}\frac{|\rho_i|}{\rho_2}<1-\delta. \tag{3.13}$$

Аналогично (3.9) можно показать, что

$$\theta_{i,\,2k+2}=c_k(\rho_i+\varepsilon_k^i)\theta_{i2k},\qquad i=2,3,\ldots,n, \tag{3.14}$$

где $\varepsilon_k^i\to 0$ при $k\to\infty$, c_k — некоторое положительное число. Так как

$$\|y_{2k}\|^2<\frac{\theta_{12k}^2}{\theta_{22k}^2}+\sum_{i=3}^{n-1}\frac{\theta_{i2k}^2}{\theta_{22k}^2},$$

то, используя (3.13), (3.14), убеждаемся в справедливости теоремы. Теорема доказана.

3. Будем понимать под асимптотической скоростью сходимости выражение

$$V(v_0)=\lim_{k\to\infty}\frac{\mu_k-\lambda_1}{\mu_{k+1}-\lambda_1}.$$

Имеет место

Теорема 5. *Если* $\tilde{v}_0$ — *такое начальное приближение, что* $\mu_0<\lambda_2$, $\alpha_{20}\neq0$, $\alpha_{n0}\neq0$, *то* $V(v_0)$ *существует и*

$$V(v_0)=1+\frac{(\lambda_2-\lambda_1)(\lambda_n-\lambda_1)}{\beta^2}, \tag{3.15}$$

где

$$\beta=\lim_{k\to\infty}\beta_k.$$

Доказательство. Используя теорему 3, нетрудно показать, что

$$\lim_{k\to\infty}\frac{\|w_{k+1}\|^2(\mu_k-\lambda_1)\|\bar{v}_{k+1}\|^2}{(\mu_{k+1}-\lambda_1)\|w_k\|^2\|\bar{v}_k\|^2}=(q-\lambda_1)(\bar{q}-\lambda_1),$$

где

$$q=\lim_{k\to\infty}q_{2k},\qquad \bar{q}=\lim_{k\to\infty}q_{2k+1}.$$

Тогда

$$\lim_{k\to\infty}\frac{\mu_k-\lambda_1}{\mu_{k+1}-\lambda_1}=\frac{(q-\lambda_1)(\bar{q}-\lambda_1)}{\beta^2}. \tag{3.16}$$

Так как $(q-\lambda_1)(\bar{q}-\lambda_1)=\beta^2+(\lambda_2-\lambda_1)(\lambda_n-\lambda_1)$, то из (3.16) следует справедливость теоремы. Теорема доказана.

Так как β является весьма сложной функцией от начального приближения $\tilde{v}_0$, то непосредственное применение формулы (3.15) затруднено. Однако с помощью этой формулы можно получить полезную в приложениях оценку сверху $V(v_0)$ через β_0.

Действительно, в силу неравенства (2.5), имеем

$$\beta\geqslant\beta_0\prod_{k=0}^{\infty}(1-\varepsilon_k). \tag{3.17}$$

Так как

$$1-\varepsilon_k=\frac{\tau_k^{-2}}{\tau_k^{-2}+\|w_k\|^2}=(1+\tau_k^2\|w_k\|^2)^{-1},$$

а

$$\tau_k\|w_k\|^2=\mu_k-\mu_{k+1},$$

то

$$P=\prod_{k=0}^{\infty}(1-\varepsilon_k)>\exp\left[-\sum_{k=0}^{\infty}\tau_k(\mu_k-\mu_{k+1})\right]. \tag{3.18}$$

Предположим, что $\mu_0<\lambda_2$, $\alpha_{20}\neq0$, $\alpha_{n0}\neq0$. В силу неравенств (2.8), (3.18), имеем

$$P>\exp\left\{\frac{\lambda_1-\mu_0}{\lambda_2-\mu_0}\right\},$$

тогда, в силу (3.17),

$$\beta\geqslant\beta_0P>\beta_0\exp\left\{\frac{\lambda_1-\mu_0}{\lambda_2-\mu_0}\right\}. \tag{3.19}$$

Используя теорему 5 и неравенство (3.19), окончательно получаем

$$V(v_0)<1+\left\{(\lambda_n-\lambda_1)(\lambda_2-\lambda_1)\exp\left[\frac{2(\mu_0-\lambda_1)}{\lambda_2-\mu_0}\right]\right\}\beta_0^{-2}.$$

4. К методу наискорейшего спуска применимы приемы ускорения сходимости А. А. Абрамова и Л. А. Люстерника [5].

ε	k	λ_1^k	λ_2^k	λ_{19}^k
1	14	10.686530	89.824655	1561.9711
10^{-1}	25	9.9458139	86.771042	1581.7707
10^{-2}	38	9.8576849	84.493418	1587.4895
10^{-3}	52	9.8502060	74.733621	1589.0522
10^{-4}	73	9.8494257	47.201028	1589.7438
10^{-5}	104	9.8493366	39.333939	1590.1311
	128	9.8493290	39.163064	1590.1477

Так, если на каждой итерации метода наискорейшего спуска не производить нормировку последовательных приближений $\tilde{v}_k$, то, в силу теоремы 3, получившийся итерационный процесс будет асимптотически линейным. Это позволяет применить к нему прием Абрамова ускорения сходимости линейных итерационных процессов.

При нормальном же течении метода наискорейшего спуска происходит «накапливание» коэффициентов α_{2k}, α_{nk}. Для ликвидации этого отрицательного явления можно использовать идею Люстерника ускорения сходимости линейных итерационных процессов, прерывая нормальное течение метода наискорейшего спуска итерациями с шагами

$$\tau_k=(\lambda_n^k-\mu_k)^{-1},\qquad \tau_{k+1}=(\lambda_2^k-\mu_{k+1})^{-1},$$

где λ_n^k, λ_2^k определены в п. 2.

5. Скорость сходимости λ_2^k к λ_2, λ_n^k к λ_n, вообще говоря, меньше скорости сходимости μ_k к λ_1. Проиллюстрируем это на простейшей задаче на собственные значения:

$$v_{x\bar{x}}+\lambda v=0,\qquad v_0=v_n=0$$

$$\text{на } D=\{x_i=ih,\ i=0,1,\dots,n,\ h=1/n\},\quad n=20.$$

Точные значения λ_1, λ_2, λ_{19} таковы: $\lambda_1=9.8493271$, $\lambda_2=39.154785$, $\lambda_{19}=1590.1506$.

Через λ_1^k, λ_2^k, λ_{19}^k обозначим соответствующие приближения к λ_1, λ_2, λ_{19} на k-й итерации. В качестве начального приближения возьмем $\tilde{v}_0=(0,2,\dots,1,0)$. Результаты счета вынесены в таблицу. Номер k в таблице — это первый номер, для которого $\lambda_1^k-\lambda_1\leqslant\varepsilon$.

Автор выражает глубокую благодарность А. А. Самарскому, В. Л. Макарову и А. В. Гулину за обсуждение результатов и ценные замечания.

Литература

1. *Самарский А. А., Николаев Е. С.* Методы решения сеточных уравнений. М.: Наука, 1978.

2. *Потапова А. Ф.* Об ускорении сходимости метода скорейшего спуска.— Ж. вычисл. матем. и матем. физ., 1971, т. 11, № 3, с. 749—752.
3. *Forsythe G. E.* On the asymptotic directions of the s-dimensional optimum gradient method.— Numer. Math., 1968, v. 11, № 1, p. 57—76.
4. *Akaike H.* On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method.— Ann. Inst. Statist. Math., Tokyo, 1959, v. 11, p. 1—16.
5. *Фаддеев Д. К., Фаддеева В. Н.* Вычислительные методы линейной алгебры. М.: Физматгиз, 1963.
6. *Приказчиков В. Г.* Строгие оценки скорости сходимости итерационного метода вычисления собственных значений.— Ж. вычисл. матем. и матем. физ., 1975, т. 15, № 5, с. 1330—1333.
7. *Канторович Л. В.* Функциональный анализ и прикладная математика.— Успехи матем. наук, 1948, т. 3, № 6, с. 89—185.

Поступила в редакцию 16.VII.1979