ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 22 Март 1982 Апрель № 2

УДК 517.988.8

АСИМПТОТИЧЕСКИЕ СВОЙСТВА s-ШАГОВОГО МЕТОДА СКОРЕЙШЕГО СПУСКА

ЖУК П. Ф.

(Херсон)

Указан способ построения векторов, инвариантных по направлению относительно двух итераций s-шагового метода скорейшего спуска. Найдено достаточное условие, при котором справедлива гипотеза Форсайта об асимптотическом поведении s-шагового метода скорейшего спуска.

Введение

1. Пусть дано уравнение

$$Au=f, \tag{1}$$

где A — линейный положительно-определенный и самосопряженный оператор, действующий в вещественном гильбертовом пространстве H. Решение уравнения (1) ищется с помощью s-шагового метода скорейшего спуска (м.с.с.). Последовательные приближения $u_0, u_1, \ldots$ к решению уравнения (1) строятся по правилу

$$u_{k+1}=u_k+\sum_{i=1}^{s}\gamma_i^{(k)}A^{i-1}z_k, \qquad k=0,1,\ldots, \tag{2}$$

где $z_k=Au_k-f$, а числа $\gamma_i^{(k)}$ выбираются из условия минимума величины

$$F(u_{k+1})=0.5(Au_{k+1}, u_{k+1})-(u_{k+1}, f).$$

Доказательство сходимости и оценки скорости сходимости s-шагового м.с.с. получены в [1] и [2]. Неулучшаемая оценка скорости сходимости этого метода в конечномерном пространстве H найдена в [3].

Теоретический и практический интерес представляет изучение асимптотического поведения s-шагового м.с.с. (т. е. изучение сходимости последовательности векторов $y_k=z_k/\|z_k\|$). Для одношагового ($s=1$) м.с.с. в конечномерном пространстве в [4], [5] установлено, что последовательности $\{y_{2k}\}_{k=0}^{\infty}$, $\{y_{2k+1}\}_{k=0}^{\infty}$ сходятся. Обобщение некоторых результатов об асимптотическом поведении s-шагового м.с.с. в конечномерных пространствах на бесконечномерные приведено в [6]. Вопросам практического применения асимптотических свойств и изучению асимптотической скорости сходимости одношагового м.с.с. посвящены работы [7]—[9].

Наиболее известные результаты об асимптотическом поведении s-шагового м.с.с. ($s>1$) в конечномерном пространстве принадлежат Форсайту. Так, им доказано (см. [5]), что множество предельных векторов $R(y_0, A)$ последовательности $\{y_{2k}\}_{k=0}^{\infty}$ связно и замкнуто в H и что каждый предельный вектор инвариантен по направлению относительно двух последовательных итераций.

В настоящей работе сформулировано условие, при котором последовательность $\{y_{2k}\}_{k=0}^{\infty}$ сходится, а также исследуется множество векторов $F(A)$ из единичной сферы, инвариантных по направлению относительно двух последовательных итераций s-шагового м.с.с. В § 1 указан критерий, позволяющий только по многочлену $P_s(t, y)$ ответить на вопрос, принадлежит ли вектор y множеству $F(A)$. В § 2 найдено необходимое и достаточное условие, которому должен удовлетворять произвольный вещественный многочлен $P(t)$, для того чтобы существовал хотя бы один вектор $y \in H$ такой, что $P(t) = P_s(t, y)$ (множество $F(P)$ всех y, удовлетворяющих данному равенству, определяется множеством положительных решений указанной в § 2 системы линейных уравнений). Далее определяется совокупность всех многочленов, удовлетворяющих критерию § 1 и условию § 2, и строится множество $F(A)$ как объединение множеств $F(P)$ по всем многочленам этой совокупности. В частности, показано, что множества $F_m(A)$ (см. п. 2) непусты при любых $0<\lambda_1<\ldots<\lambda_m$ и m, $s+1 \leqslant m \leqslant 2s$. В § 3 рассматривается специальное множество $V(I)$ векторов из $(s+1)$-мерного подпространства, натянутого на собственные векторы оператора A. Доказано, что если $R(y_0, A) \cap V(I) \neq \varnothing$, то последовательность нормированных градиентов $v_0, v_2, \ldots$ сходится, если только v_0 принадлежит достаточно малой окрестности вектора y_0. Более того, предел последовательности $v_0, v_2, \ldots$ непрерывно зависит от v_0 из данной окрестности вектора y_0. В частности, для любого s доказано существование векторов $y_0 \notin F(A)$ таких, что последовательность $y_0, y_2, \ldots$ сходится.

2. Здесь и в дальнейшем полагаем, что A — диагональная матрица с различными собственными значениями:

$$A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \qquad 0<\lambda_1<\ldots<\lambda_n$$

(общий случай рассмотрен в § 3). Так как случай $2s>n$ сводится к случаю $2s \leqslant n$ путем добавления к векторам нулевых компонент (не влияющих на свойства s-шагового м.с.с.), то всюду предполагается также $2s \leqslant n$. Положим

$$P_s(A, z_0) = E + \sum_{i=1}^{s} \gamma_i^{(0)} A^i, \tag{3}$$

где коэффициенты $\gamma_i^{(0)}$ определяются из соотношения (2) при $k=0$, E — единичная матрица. Пусть Σ^* — множество векторов из единичной сферы Σ пространства H, у которых по крайней мере $s+1$ ненулевых компонент. Каждому $y \in \Sigma^*$ поставим в соответствие матрицу $T(y) = P_s(A, y)/\|P_s(A, y)y\|$, где $P_s(A, y)$ определено в (3) с $z_0 = y$. Множество $F(A)$ совпадает с множеством решений уравнения

$$T(y')y' = y, \tag{4}$$

где $y' = T(y)y$. Так как многочлен степени $2s$ имеет не более $2s$ корней, то решениями уравнения (4) могут являться векторы не более чем с $2s$ ненулевыми компонентами. Множества $F_m(A)$, $s+1<m \leqslant 2s$, решений уравнения (4) с m ненулевыми компонентами фактически не исследованы, и даже вопрос об их непустоте в общем случае является открытым (множество $F_{s+1}(A)$ изучено в [5]; оно совпадает с множеством векто-

ров, имеющих в точности $s+1$ ненулевую компоненту). Отметим, что в [5] установлена непустота некоторых из множеств $F_m(A)$, $s+1<m\leqslant 2s$, лишь для спектров, симметричных относительно середины.

§ 1. Условие включения в множество $F(A)$

В § 1, 2 рассматриваются векторы y из $F(A)$, имеющие только первые m ненулевых компонент (т. е. $y=(y_1\ldots,y_m,\ 0,\ldots,0)$, $y_i\neq 0$, $i=1,2,\ldots,m$), что, по существу, не ограничивает общности утверждений. Уравнение (4) запишем в виде

$$(1.1)\qquad P_s(A,y')P_s(A,y)y=\|P_s(A,y')P_s(A,y)y\|y.$$

Так как многочлен $P_s(t,y)$ с точностью до постоянного множителя равен определителю (см. [5])

$$D(t,y)=\det\begin{vmatrix}\mu_0 & \mu_1 & \ldots & \mu_{s-1} & 1\\ \mu_1 & \mu_2 & \ldots & \mu_s & t\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \mu_s & \mu_{s+1} & \ldots & \mu_{2s-1} & t^s\end{vmatrix},$$

где

$$\mu_\alpha=(A^\alpha y,y)=\sum_{i=1}^m \lambda_i^\alpha y_i^2,\qquad \alpha=0,1,\ldots,2s-1,$$

то уравнение (1.1) эквивалентно следующей системе нелинейных алгебраических уравнений относительно $y_1,\ldots,y_m$:

$$(1.2)\qquad P_1'P_1=P_k'P_k,\qquad k=1,2,\ldots,m,$$

где $P_k=P_k(y)=D(\lambda_k,y)$,

$$(1.3)\qquad P_k'=P_k'(y)=\det\begin{vmatrix}\mu_0' & \mu_1' & \ldots & \mu_{s-1}' & 1\\ \mu_1' & \mu_2' & \ldots & \mu_s' & \lambda_k\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \mu_s' & \mu_{s+1}' & \ldots & \mu_{2s-1}' & \lambda_k^s\end{vmatrix},$$

$$\mu_\alpha'=\sum_{i=1}^m \lambda_i^\alpha P_i^2 y_i^2,\qquad \alpha=0,1,\ldots,2s-1.$$

Непосредственное решение системы (1.2) затруднительно, однако верна

Теорема 1. *Вектор y принадлежит множеству $F(A)$ тогда и только тогда, когда существует многочлен $P_s(t)=\gamma_0+\gamma_1 t+\ldots+\gamma_s t^s$ такой, что для всех $k=1,2,\ldots,m$ справедливо равенство*

$$(1.4)\qquad P_k(y)P_s(\lambda_k)=1.$$

Предварительно докажем вспомогательное утверждение.

Лемма 1. *Пусть $y=(y_1,\ldots,y_m,\ 0,\ldots,0)$, $y_i\neq 0$, $i=1,2,\ldots,m$. Тогда*

$$(1.5)\qquad P_1'P_1-P_m'P_m=-\sum_{1<j_1<\ldots<j_s<m} P_{j_1}y_{j_1}^2\ldots P_{j_s}y_{j_s}^2\Delta_{j_1\ldots j_s}B_{1j_1\ldots j_s m},$$

где

$$\Delta_{j_1\ldots j_s}=\det\begin{vmatrix}1 & \ldots & 1\\ \lambda_{j_1} & \ldots & \lambda_{j_s}\\ \cdot & \cdot & \cdot\\ \lambda_{j_1}^{s-1} & \ldots & \lambda_{j_s}^{s-1}\end{vmatrix},$$

$$B_{1j_1\ldots j_s m}=\det\begin{vmatrix}1 & 1 & \ldots & 1 & 1\\ P_1 & P_{j_1} & \ldots & P_{j_s} & P_m\\ \lambda_1P_1 & \lambda_{j_1}P_{j_1} & \ldots & \lambda_{j_s}P_{j_s} & \lambda_mP_m\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ \lambda_1^sP_1 & \lambda_{j_1}^sP_{j_1} & \ldots & \lambda_{j_s}^sP_{j_s} & \lambda_m^sP_m\end{vmatrix}.$$

Суммирование в (1.5) ведется по всем наборам $j_1,\ldots,j_s$, удовлетворяющим указанным там условиям.

Доказательство. Из определения P_k', $k=1, 2,\ldots, m$, следует, что

$$(1.6)\qquad P_1'=\sum_{1<j_1<\ldots<j_s<m} P_{j_1}^2y_{j_1}^2\ldots P_{j_s}^2y_{j_s}^2\Delta_{j_1\ldots j_s1}\Delta_{j_1\ldots j_s}+$$

$$+\sum_{1<j_1<\ldots<j_{s-1}<m} P_m^2y_m^2P_{j_1}^2y_{j_1}^2\ldots P_{j_{s-1}}^2y_{j_{s-1}}^2\Delta_{mj_1\ldots j_{s-1}}\Delta_{mj_1\ldots j_{s-1}1},$$

$$(1.7)\qquad P_m'=\sum_{1<j_1<\ldots<j_s<m} P_{j_1}^2y_{j_1}^2\ldots P_{j_s}^2y_{j_s}^2\Delta_{j_1\ldots j_sm}\Delta_{j_1\ldots j_s}+$$

$$+\sum_{1<j_1<\ldots<j_{s-1}<m} P_1^2y_1^2P_{j_1}^2y_{j_1}^2\ldots P_{j_{s-1}}^2y_{j_{s-1}}^2\Delta_{1j_1\ldots j_{s-1}}\Delta_{1j_1\ldots j_{s-1}m}.$$

С помощью (1.6), (1.7) получаем

$$(1.8)\qquad P_1'P_1-P_m'P_m=-\sum_{1<j_1<\ldots<j_{s-1}<m} P_1P_mP_{j_1}^2y_{j_1}^2\ldots P_{j_{s-1}}^2y_{j_{s-1}}^2\times$$

$$\times\Delta_{1j_1\ldots j_{s-1}m}(P_1y_1^2\Delta_{1j_1\ldots j_{s-1}}+P_my_m^2\Delta_{mj_1\ldots j_{s-1}})+$$

$$+\sum_{1<j_1<\ldots<j_s<m} P_1P_{j_1}^2y_{j_1}^2\ldots P_{j_s}^2y_{j_s}^2\Delta_{j_1\ldots j_s}\Delta_{j_1\ldots j_s1}-$$

$$-\sum_{1<j_1<\ldots<j_s<m} P_mP_{j_1}^2y_{j_1}^2\ldots P_{j_s}^2y_{j_s}^2\Delta_{j_1\ldots j_s}\Delta_{j_1\ldots j_sm}.$$

Рассмотрим выражение

$$(1.9)\qquad P_1y_1^2\Delta_{1j_1\ldots j_{s-1}}+P_my_m^2\Delta_{mj_1\ldots j_{s-1}}=$$

$$=\det\begin{vmatrix}P_1y_1^2+P_my_m^2 & 1 & \ldots & 1\\ \lambda_1P_1y_1^2+\lambda_mP_my_m^2 & \lambda_{j_1} & \ldots & \lambda_{j_{s-1}}\\ \cdot & \cdot & \cdot & \cdot\\ \lambda_1^{s-1}P_1y_1^2+\lambda_m^{s-1}P_my_m^2 & \lambda_{j_1}^{s-1} & \ldots & \lambda_{j_{s-1}}^{s-1}\end{vmatrix}.$$

В силу построения, вектор y' ортогонален векторам $y, Ay,\ldots,A^{s-1}y$, т. е.

$$\sum_{i=1}^m\lambda_i^\alpha P_iy_i^2=0,\qquad \alpha=0, 1,\ldots, s-1.$$

Используя это соотношение, преобразуем (1.9) к виду

$$(1.10)\qquad P_1 y_1^2 \Delta_{1j_1\ldots j_{s-1}} + P_m y_m^2 \Delta_{mj_1\ldots j_{s-1}} = -\sum_{1<j<m} P_j y_j^2 \Delta_{jj_1\ldots j_{s-1}}.$$

Подставляя (1.10) в (1.8), убеждаемся в справедливости леммы. Лемма доказана.

Замечание 1. Если в левой и правой частях тождества (1.5) поменять одновременно ролями λ_m и λ_k, y_m и y_k, то левая часть (1.5) преобразуется в разность $P_1'P_1-P_k'P_k$, а измененная правая часть (1.5) даст представление этой разности, аналогичное представлению разности $P_1'P_1-P_m'P_m$ правой частью (1.5).

Доказательство теоремы 1. Пусть существует многочлен $P_s(t)$, удовлетворяющий условию (1.4). Тогда для любого набора целых чисел $1<j_1<\ldots<j_{s+1}\leqslant m$ определитель $B_{1j_1\ldots j_{s+1}}=0$ и, следовательно, в силу леммы 1 и замечания к ней, $P_1'P_1-P_k'P_k=0$, $k=1, 2,\ldots, m$. Итак, вектор y — решение системы (1.2), т. е. $y\in F(A)$.

Наоборот, если $y\in F(A)$, то многочлен $P_s(t)$ с точностью до постоянного множителя равен многочлену $P_s(t, y')$. Теорема доказана.

§ 2. Построение множества $F(A)$

1. Пусть $P(t)=t^s+q_{s-1}t^{s-1}+\ldots+q_0$ — произвольный многочлен с вещественными коэффициентами. Обозначим через $N_m(P)$ совокупность действительных решений $(y_1,\ldots,y_m)$, $y_i\neq0$, $i=1, 2,\ldots, m$, системы уравнений

$$(2.1)\qquad \sum_{i=1}^{m} y_i^2=1,\qquad \sum_{i=1}^{m} \lambda_i^\alpha P(\lambda_i) y_i^2=0,\quad \alpha=0, 1,\ldots, s-1,$$

линейной относительно $y_1^2,\ldots, y_m^2$. Имеет место

Лемма 2. *Множество $N_m(P)$ непусто тогда и только тогда, когда количество перемен знака последовательности $P(\lambda_1),\ldots, P(\lambda_m)$ равно s. Если $(y_1,\ldots, y_m)\in N_m(P)$, то вектор $y=(y_1,\ldots, y_m, 0,\ldots, 0)\in H$ таков, что $P_s(t, y)=cP(t)$, где c — некоторая постоянная.*

Множества $N_m(P)$, $s+1\leqslant m\leqslant 2s$, играют роль «кирпичиков», из которых будет построено множество $F(A)$, и могут быть найдены с помощью методов линейного программирования (см., например, [10]).

2. Укажем сначала способ построения множества $F_{2s}(A)$. Пусть

$$\beta^*=\min_{i=1,2,\ldots,s}\ \max_{t\in(\lambda_{2i-1},\lambda_{2i})} |Q_{2s}(t)|,\qquad Q_{2s}(t)=\prod_{j=1}^{2s}(\lambda_j-t).$$

Зафиксируем β, $0<\beta\leqslant\beta^*$, и обозначим через $\tau_j^1\leqslant\tau_j^2$, $j=1, 2,\ldots, s$, корни многочлена $Q_{2s}(t)-\beta$, расположенные в интервале $(\lambda_{2j-1}, \lambda_{2j})$. Пусть M — множество конечных последовательностей $(i_1,\ldots, i_s)$, где i_j, $j=1, 2,\ldots, s$, могут принимать только значения 1 или 2. Рассмотрим многочлен

$$P_{\beta i_1\ldots i_s}(t)=\prod_{j=1}^{s}(\tau_j^{i_j}-t),\qquad (i_1,\ldots, i_s)\in M.$$

Пусть $N_{2s}(P_{\beta i_1\ldots i_s})$ — множество действительных решений системы (2.1) при $P(t)=P_{\beta i_1\ldots i_s}(t)$. Обозначим через $F_{2s}(\beta, i_1,\ldots,i_s, A)$ совокупность векторов $(y_1,\ldots,y_{2s}, 0,\ldots,0)\in H$ таких, что $(y_1,\ldots,y_{2s})\in N_{2s}(P_{\beta i_1\ldots i_s})$. В силу леммы 2, $F_{2s}(\beta, i_1,\ldots, i_s, A)\neq\varnothing$, и если $\beta<\beta^*$, $(\beta, i_1,\ldots, i_s)\neq$ $\neq(\bar\beta, \bar i_1,\ldots,\bar i_s)$, то $F_{2s}(\beta, i_1,\ldots,i_s, A)\cap F_{2s}(\bar\beta, \bar i_1,\ldots,\bar i_s, A)=\varnothing$. Из построения множеств $N_{2s}(P_{\beta i_1\ldots i_s})$ и теоремы 1 следует

$$(2.2)\qquad F_{2s}(A)=\bigcup_{0<\beta\leqslant\beta^*}\ \bigcup_{(i_1,\ldots,i_s)\in M} F_{2s}(\beta, i_1,\ldots,i_s, A).$$

3. Для $m<2s$ рассмотрим последовательность из $2s$ плюсов и минусов, образованную чередованием групп $+-$ и $-+$. Например, если $s=5$, то последовательность знаков такова: $+--++--++-$. Пусть $n_1,\ldots,n_{m-1}$ — набор целых неотрицательных чисел. Оставим первый слева знак последовательности знаков и сотрем n_1 следующих за ним знаков. Из группы знаков, расположенных справа от стертых, оставим первый слева (если он существует) и сотрем n_2 следующих за ним знаков, и т. д. Будем говорить, что последовательность знаков, полученная в результате описанной процедуры, порождается набором $n_1,\ldots,n_{m-1}$. Например, если $s=6$, то набор 1, 0, 1, 2, 0, 1, 1 порождает последовательность знаков $+-+-+-+$.

Набор $n_1,\ldots,n_{m-1}$ назовем допустимым, если порожденная им последовательность имеет в точности s перемен знака.

Пусть $\{\alpha_i\}_{i=1}^{2s-m}$ — не равные $\lambda_1,\ldots,\lambda_m$, $\lambda_1<\alpha_1<\ldots<\alpha_{2s-m}<\lambda_m$, числа. Обозначим через n_j, $j=1, 2,\ldots, m-1$, количество чисел α_i, принадлежащих интервалу $(\lambda_j, \lambda_{j+1})$. Конечную последовательность $\alpha_1,\ldots,\alpha_{2s-m}$ назовем допустимой, если допустимым является построенный по ней набор $n_1,\ldots$ $\ldots, n_{m-1}$.

Пусть $\lambda_1'<\ldots<\lambda_{2s}'$ — упорядоченная по возрастанию совокупность чисел $\lambda_1,\ldots,\lambda_m, \alpha_1,\ldots,\alpha_{2s-m}$, где $\alpha_1,\ldots,\alpha_{2s-m}$ — допустимая конечная последовательность. Так же как и в п. 2, рассмотрим многочлен

$$Q_{2s}^{\alpha_1\ldots\alpha_{2s-m}}(t)=\prod_{i=1}^{2s}(\lambda_i'-t)$$

и определим

$$\beta^*_{\alpha_1\ldots\alpha_{2s-m}}=\min_{i=1,2,\ldots,s}\ \max_{t\in(\lambda'_{2i-1},\lambda'_{2i})}|Q_{2s}^{\alpha_1\ldots\alpha_{2s-m}}(t)|.$$

Пусть $0<\beta\leqslant\beta^*_{\alpha_1\ldots\alpha_{2s-m}}$, а $\tau^1_{j\alpha_1\ldots\alpha_{2s-m}}\leqslant\tau^2_{j\alpha_1\ldots\alpha_{2s-m}}$ — корни многочлена $Q_{2s}^{\alpha_1\ldots\alpha_{2s-m}}(t)-\beta$, расположенные в интервале $(\lambda'_{2j-1}, \lambda_{2j}')$. Образуем многочлен

$$P^{\alpha_1\ldots\alpha_{2s-m}}_{\beta i_1\ldots i_s}(t)=\prod_{j=1}^{s}(\tau^{i_j}_{j\alpha_1\ldots\alpha_{2s-m}}-t),$$

где $(i_1,\ldots,i_s)\in M$ — множеству, определенному в п. 2. Пусть $F_m(\beta, i_1,\ldots$ $\ldots, i_s, \alpha_1,\ldots,\alpha_{2s-m}, A)$ — совокупность векторов $(y_1,\ldots,y_m, 0,\ldots,0)\in H$ таких, что $(y_1,\ldots,y_m)\in N_m(P^{\alpha_1\ldots\alpha_{2s-m}}_{\beta i_1\ldots i_s})$ — множеству действительных решений системы (2.1) при $P(t)=P^{\alpha_1\ldots\alpha_{2s-m}}_{\beta i_1\ldots i_s}(t)$. Используя теорему 1 и лемму 2, можно показать следующее:

$$(2.3)\qquad F_m(\beta, i_1,\ldots,i_s, \alpha_1,\ldots,\alpha_{2s-m}, A)\neq\varnothing;$$

$$F_m(\beta, i_1, \ldots, i_s, \alpha_1, \ldots, \alpha_{2s-m}, A) \cap F_m(\bar\beta, \bar i_1, \ldots, \bar i_s, \bar\alpha_1, \ldots, \bar\alpha_{2s-m}, A) = \varnothing$$

при $(\beta, i_1, \ldots, i_s, \alpha_1, \ldots, \alpha_{2s-m}) \neq (\bar\beta, \bar i_1, \ldots, \bar i_s, \bar\alpha_1, \ldots, \bar\alpha_{2s-m})$ и $\beta < \beta^*_{\alpha_1 \ldots \alpha_{2s-m}}$;

$$F_m(A) = \bigcup_{(\alpha_1, \ldots, \alpha_{2s-m}) \in G} \; \bigcup_{0 < \beta \leqslant \beta^*_{\alpha_1 \ldots \alpha_{2s-m}}} \; \bigcup_{(i_1, \ldots, i_s) \in M} F_m(\beta, i_1 \ldots$$
$$\ldots, i_s, \alpha_1, \ldots, \alpha_{2s-m}, A), \tag{2.4}$$

где $G \neq \varnothing$ — множество всех допустимых конечных последовательностей $\alpha_1, \ldots, \alpha_{2s-m}$.

Замечание 2. Из соотношений (2.2)–(2.4) следует, что $F_m(A)$ представляет собой объединение несчетного семейства непустых множеств, следовательно, непусто. Тем самым доказано, что при произвольных $0 < \lambda_1 < \ldots < \lambda_m$ и m, $s+1 \leqslant m \leqslant 2s$, множество векторов, инвариантных по направлению относительно двух последовательных итераций s-шагового м.с.с. и имеющих в точности m ненулевых компонент, непусто.

§ 3. Асимптотическое поведение s-шагового м.с.с.

1. Пусть $y_k = z_k / \|z_k\|$, $k = 0, 1, \ldots$, — нормированные градиенты s-шагового м.с.с. В [5] высказана гипотеза, что для любого вектора $p_0 \in \Sigma^*$ последовательность $\{y_{2k}\}_{k=0}^{\infty}$ сходится. Справедливость этой гипотезы доказана в случае $s=1$ и очевидна для $y_0 \in F(A)$. При $s > 1$, по-видимому, неизвестно, существует ли хотя бы один вектор $y_0 \notin F(A)$ такой, что $\{y_{2k}\}_{k=0}^{\infty}$ сходится.

2. Обозначим через Σ_1 множество векторов $y \in \Sigma^*$, для которых соответствующие последовательности $y_0 = y, y_2, \ldots, y_{2k}, \ldots$ сходятся (векторы $y_2, y_4, \ldots$ зависят от y_0), а через f: $\Sigma_1 \to F(A)$ — функцию, определенную по правилу

$$f(y) = \lim_{k \to \infty} y_{2k}.$$

Пусть I — набор целых чисел $i_1, \ldots, i_{s+1}$, $1 \leqslant i_1 < \ldots < i_{s+1} \leqslant n$, а $H(I)$ — подпространство пространства H, натянутое на собственные векторы матрицы A, соответствующие собственным значениям $\{\lambda_i\}_{i \in I}$. Через $V(I)$ обозначим множество векторов $y \in \Sigma^* \cap H(I)$ таких, что для каждого $i \in I$ и $j \in \{1, 2, \ldots, n\} \setminus I$ справедливо неравенство

$$P_s(\lambda_i, y') P_s(\lambda_i, y) > |P_s(\lambda_j, y') P_s(\lambda_j, y)| \tag{3.1}$$

(здесь $y' = T(y) y$). С помощью леммы 2 можно показать, что существует по крайней мере один набор I, для которого $V(I) \neq \varnothing$. Отметим также, что если $V(I) \neq \varnothing$ для некоторого I, то $1, n \in I$.

Основным результатом данного параграфа является

Теорема 2. *Если $R(y_0, A) \cap V(I) \neq \varnothing$, то на достаточно малой окрестности вектора y_0*

$$O(\delta, y_0) = \{v \mid v \in \Sigma, \|v - y_0\| < \delta\}$$

функция f определена и непрерывна.

Для произвольного вектора $v \in \Sigma^*$ положим $v' = T(v) v$, $v'' = T(v') v'$, а $\tilde v = (\tilde v_1, \ldots, \tilde v_n)$ — проекция вектора v на подпространство $H(I)$, $\bar v = (\bar v_1, \ldots, \bar v_n)$ — ортогональное дополнение (т. е. $v = \tilde v + \bar v$, $\tilde v \in H(I)$, $\bar v \perp H(I)$). Предварительно докажем вспомогательное утверждение.

Лемма 3. *Для всякого замкнутого множества $K \subseteq \Sigma^*$ существует неотрицательная, ограниченная на K функция $\alpha_K(v)$ такая, что для любого $v \in K$*

(3.2) $$\|v''-v\|^2=([T(v')T(v)-E]^2\bar{v}, \bar{v})+\alpha_K(v)\|\bar{v}\|^4.$$

Доказательство. Соотношение (3.2) эквивалентно равенству

(3.3) $$([T(v')T(v)-E]^2\tilde{v}, \tilde{v})=\alpha_K(v)\|\bar{v}\|^4.$$

Если $v \notin H(I)$, то из (3.3) следует, что

(3.4) $$\alpha_K(v)=([T(v')T(v)-E]^2\tilde{v}, \tilde{v})/\|\bar{v}\|^4.$$

Для $v \in H(I)$ положим $\alpha_K(v)=0$. Определенная таким образом функция $\alpha_K(v)$ неотрицательна. Покажем, что она ограничена на K.

Выражение (3.4) запишем в виде

(3.5) $$\alpha_K(v)=\sum_{i=1}^{n}[P_i'(v)P_i(v)-L^{1/2}(v)]^2\tilde{v}_i^2/L(v)\|\bar{v}\|^4,$$

где

$$L(v)=\sum_{i=1}^{n}[P_i'(v)P_i(v)]^2v_i^2,$$

$P_i'(v)$, $P_i(v)$ определены в (1.3). Функция $L(v)$ непрерывна и положительна на Σ^*, поэтому существует

$$l=\min_{v\in K} L(v)>0.$$

Пусть

$$P=\max_{i=1,2,\ldots,n}\ \sup_{v\in\Sigma^*}[P_i'(v)P_i(v)]^2,$$

$$Q=\sup_{v\in\Sigma^*}\sum_{i=1}^{n}[P_i'(v)P_i(v)-L^{1/2}(v)]^2v_i^2.$$

Из (1.3) следует, что $P<\infty$, $Q<\infty$. Если для некоторого $v\in K$

$$L'(v)=\sum_{i=1}^{n}[P_i'(v)P_i(v)]^2\tilde{v}_i^2<\sum_{i=1}^{n}[P_i'(v)P_i(v)]^2\bar{v}_i^2=L''(v),$$

то, в силу неравенства

(3.6) $$\|\bar{v}\|^4\geqslant[L''(v)/P]^2\geqslant l^2/4P^2,$$

получаем

(3.7) $$\alpha_K(v)\leqslant 4QP^2/l^3.$$

В противном случае ($L'(v)\geqslant L''(v)$) запишем

(3.8) $$P_i'(v)P_i(v)-L^{1/2}(v)=P_i'(v)P_i(v)-[L'(v)]^{1/2}-$$
$$-\theta(v)L''(v)/[L'(v)]^{1/2},$$

где $0<\theta(v)<1/2$. Применяя к разности $P_i'(v)P_i(v)-P_j'(v)P_j(v)$ замечание 1, можно показать, что при $i\in I$

$$|P_i'(v)P_i(v)-[L'(v)]^{1/2}|\leqslant a\|\bar{v}\|^2,$$

где a — число, не зависящее от v; поэтому из соотношений (3.6), (3.8) следует, что в случае $L'(v)\geqslant L''(v)$

(3.9) $\qquad \alpha_K(v)\leqslant[a(2l)^{1/2}+P]^2/2l^2.$

Справедливость леммы 2 вытекает из (3.7), (3.9). Лемма доказана.

Доказательство теоремы 2. Пусть $y\in R(y_0, A)\cap V(I)$. Тогда существует ε_1-окрестность точки y

$$O(\varepsilon_1, y)=\{v\,|\,v\in\Sigma,\ \|v-y\|<\varepsilon_1\}$$

такая, что для каждого $v\in O(\varepsilon_1, y)$ выполняются неравенства

(3.10) $\qquad \|v''-v\|\leqslant M\|\bar{v}\|,$

(3.11) $\qquad \|v''-v\|\leqslant\rho\|\bar{v}\|, \qquad \rho<1,$

где M, ρ — числа, не зависящие от v. Действительно, пусть ε_1 таково, что $\bar{O}(\varepsilon_1, y)\subseteq\Sigma^*$ ($\bar{O}(\varepsilon_1, y)$ — замыкание окрестности $O(\varepsilon_1, y)$ в H). Из равномерной ограниченности по v матриц $T(v)$ на множестве $\bar{O}(\varepsilon_1, y)$ и леммы 3 (при $K=\bar{O}(\varepsilon_1, y)$) следует справедливость неравенства (3.10). Уменьшая ε_1, можно добиться также выполнения соотношения (3.11) (в силу непрерывности $P_s(\lambda_i, v)$ по v на $\bar{O}(\varepsilon_1, y)$ и неравенства (3.1)).

Пусть ε_2 удовлетворяет условию $0<\varepsilon_2<(1-\rho)\varepsilon_1/(1-\rho+M)$. Докажем, что на $O(\varepsilon_2, y)$ функция f определена и непрерывна. Построим последовательность функций g_k: $\Sigma^*\to\Sigma^*$ по правилу

$$g_1(v)=T(v)v, \qquad g_k(v)=T(g_{k-1}(v))g_{k-1}(v), \quad k=2, 3, \ldots,$$

и подействуем ими на произвольный вектор $v_0\in O(\varepsilon_2, y)$:

$$v_1=g_1(v_0), \quad v_2=g_2(v_0), \ldots .$$

Величина $v_{2k}\in O(\varepsilon_1, y)$ для каждого $k=0, 1, \ldots$. Действительно, пусть $v_{2k}\in O(\varepsilon_1, y)$ для $k\leqslant p$. Так как

$$\|v_{2p+2}-y\|\leqslant\|v_{2p+2}-v_{2p}\|+\|v_{2p}-v_{2p-2}\|+\ldots+\|v_0-y\|,$$

а в силу неравенств (3.11), (3.10) для $k=0, 1, \ldots, p$

(3.12) $\qquad \|v_{2k+2}-v_{2k}\|\leqslant M\rho^k\varepsilon_2,$

то $\|v_{2p+2}-y\|\leqslant M\varepsilon_2(\rho^p+\rho^{p-1}+\ldots+1)+\varepsilon_2<\varepsilon_1$, т. е. $v_{2p+2}\in O(\varepsilon_1, y)$.

Таким образом, оценка (3.12) обеспечивает сходимость ряда

$$\sum_{k=0}^{\infty}\|v_{2k+2}-v_{2k}\|$$

и тем более сходимость ряда $v_0+(v_2-v_0)+\ldots+(v_{2k+2}-v_{2k})+\ldots$, т. е. $v_0\in\Sigma_1$.

Докажем непрерывность функции f в точке v_0. По заданному $\varepsilon>0$ выбираем такое натуральное число N, что

(3.13) $\qquad \rho^N<\varepsilon(1-\rho)/3M\varepsilon_2.$

В силу непрерывности функции $g_{2N}(v)$ на Σ^*, существует $\delta_1>0$ такое, что $O(\delta_1, v_0)\subseteq O(\varepsilon_2, y)$ и

(3.14) $\qquad g_{2N}(O(\delta_1, v_0))\leqslant O\left(\dfrac{\varepsilon}{3}, v_{2N}\right).$

Тогда $f(O(\delta_1, v_0))\subseteq O(\varepsilon, f(v_0))$. Действительно, пусть u — произвольный вектор из $O(\delta_1, v_0)$. Используя (3.12)—(3.14), оцениваем $\|f(v_0)-f(u)\|$:

$$\text{(3.15)}\qquad \|f(v_0)-f(u)\|\leqslant\sum_{i=N}^{\infty}(\|v_{2i+2}-v_{2i}\|+\|u_{2i+2}-u_{2i}\|)+$$

$$+\|v_{2N}-u_{2N}\|<2M\varepsilon_2(\rho^N+\rho^{N+1}+\ldots)+\frac{\varepsilon}{3}<\varepsilon,$$

т. е. $f(u)\in O(\varepsilon, f(v_0))$.

Итак, функция f определена и непрерывна на $O(\varepsilon_2, y)$.

Поскольку $y\in R(y_0, A)$, то $y_{2p}\in O(\varepsilon_2, y)$ для некоторого номера p. В силу непрерывности $g_{2p}(v)$ на Σ^*, существует $\delta>0$ такое, что $g_{2p}(O(\delta, y_0))\subseteq \subseteq O(\varepsilon_2, y)$. Окрестность $O(\delta, y_0)$ — искомая в теореме окрестность. В самом деле, функция f определена на $O(\delta, y_0)$:

$$f(v)=\lim_{k\to\infty}g_{2k}(v)=f(g_{2p}(v)),\qquad v\in O(\delta, y_0),$$

и непрерывна как суперпозиция двух непрерывных функций. Теорема доказана.

Если $y_0\in V(I)$, то условие теоремы 2 выполнено:

$$R(y_0, A)\cap V(I)=\{y_0\}\neq\varnothing.$$

Следовательно, существует окрестность $O(\delta, y_0)$ точки y_0, на которой определена и непрерывна функция f, причем из доказательства теоремы 2 видно, что $f(O(\delta, y_0))\subseteq H(I)$. В частности, отсюда вытекает существование векторов $y_0\in\Sigma_1$ таких, что $y_0\notin F(A)$.

3. До сих пор предполагалось, что A — диагональная матрица с различными собственными значениями. Однако все результаты могут быть распространены и на общий случай. Действительно, выбирая в качестве базиса пространства H ортонормированную систему собственных векторов оператора A, можно считать, что A задается диагональной матрицей, возможно, с кратными собственными значениями $\lambda_1=\ldots=\lambda_{k_1}<\lambda_{k_1+1}=\ldots=\lambda_{k_2}< <\lambda_{k_2+1}=\ldots<\lambda_{k_{n-1}+1}=\ldots=\lambda_{k_n}$. Пусть $B=\operatorname{diag}(\lambda_{k_1},\ldots,\lambda_{k_n})$, а R^n есть n-мерное арифметическое пространство. Установим связь между асимптотическими свойствами s-шагового м.с.с. в пространстве H (с матрицей A) и в R^n (с матрицей B). Положим, что Σ_n — единичная сфера в R^n, $\Sigma_n^*\subseteq\Sigma_n$ — множество единичных векторов, у которых по крайней мере $s+1$ ненулевых координат. Пусть $P_s(A, y)$, $Q_s(B, x)$ — матричные многочлены, реализующие итерацию s-шагового м.с.с. в пространствах H и R^n соответственно (см. (3)). Рассмотрим отображение, заданное формулой $\Lambda(\alpha_1,\ldots,\alpha_{k_1},\ldots \ldots,\alpha_{k_n})=(\xi_1,\ldots,\xi_n)$, где $\xi_i=\Lambda_i(\alpha_1,\ldots,\alpha_{k_1},\ldots,\alpha_{k_n})=(\alpha^2_{k_{i-1}+1}+\ldots+\alpha^2_{k_i})^{1/2}$, $i=1, 2,\ldots, n$, $k_0=0$. Имеет место следующее основное соотношение: если $x=\Lambda y\in\Sigma_n^*$, то

$$\text{(3.16)}\qquad Q_s(\lambda_{k_{i+1}}, x)=P_s(\lambda_{k_i+1}, y)=\ldots=P_s(\lambda_{k_{i+1}}, y),\quad i=0, 1,\ldots, n-1,\quad k_0=0.$$

Из (3.16), в частности, вытекает, что если $F(B)\subseteq\Sigma_n^*$ — множество векторов, инвариантных по направлению относительно двух последовательных итераций s-шагового м.с.с. в пространстве R^n, то

$$F(A)=\Lambda^{-1}(F(B)). \tag{3.17}$$

Соотношение (3.17) позволяет построить множество $F(A)$ в общем случае.

Для изучения асимптотического поведения s-шагового м.с.с. в пространстве H заметим, что если $y_0, y_1, \ldots$ — последовательность нормированных градиентов в H, то, в силу (3.16), $x_0=\Lambda y_0$, $x_m=(\tau_{1m}\Lambda_1 y_m, \ldots, \tau_{nm}\Lambda_n y_m)$, $m=1, 2, \ldots$, — последовательность нормированных градиентов в R^n (здесь $\tau_{im}=\operatorname{sign}(P_s(\lambda_{k_i}, y_m)x_{i,m-1})$, $x_{i,m-1}$ есть i-я координата вектора x_{m-1}). Более того, из сходимости $x_0, x_2, \ldots$ следует сходимость $y_0, y_2, \ldots$, причем если $y_0=(\alpha_1^0, \ldots, \alpha_{k_1}^0, \ldots, \alpha_{k_n}^0)$, $x_0=(\xi_1^0, \ldots, \xi_n^0)$, $x=\tilde f(x_0)=\lim_{m\to\infty} x_{2m}=(\xi_1, \ldots, \xi_n)$, то $y=f(y_0)=(\alpha_1, \ldots, \alpha_{k_1}, \ldots, \alpha_{k_n})$, где

$$\alpha_i=\begin{cases} 0, & \text{если } \xi_j^0=0,\\ \alpha_i^0\xi_j/\xi_j^0, & \text{если } \xi_j^0\neq 0, \end{cases} \qquad i=1,2,\ldots,k_n,$$

а j определяется из неравенства $k_{j-1}<i\leqslant k_j$, $k_0=0$. Таким образом, если функция

$$\tilde f(x_0)=\lim_{m\to\infty} x_{2m}$$

определена и непрерывна на некотором открытом в топологии Σ_n множестве O и $\Lambda^{-1}(O)$ — непустое открытое множество в топологии Σ, то на $\Lambda^{-1}(O)$ определена и непрерывна функция f.

Автор выражает благодарность В. Л. Макарову за внимание к работе.

Литература

1. *Канторович Л. В.* Функциональный анализ и прикладная математика.– Успехи матем. наук, 1948, т. 3, вып. 6, с. 89–185.
2. *Бирман М. Ш.* Некоторые оценки для метода наискорейшего спуска.– Успехи матем. наук, 1950, т. 5, вып. 3, с. 152–155.
3. *Ковригин А. Б.* Оценка быстроты сходимости K-шагового градиентного метода.– Вестн. Ленингр. ун-та, 1970, т. 13, с. 34–36.
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., 1959, v. 11, p. 1–16.
5. *Forsythe G. E.* On the asymptotic directions of the s-dimensional optimum gradient method.– Numer. Math., 1968, v. 11, № 1, p. 57–76.
6. *Заболоцкая А. Ф.* Асимптотическое поведение s-шагового метода скорейшего спуска в гильбертовом пространстве.– Ж. вычисл. матем. и матем. физ., 1979, т. 19, № 1, с. 228–232.
7. *Самарский А. А., Николаев Е. С.* Методы решения сеточных уравнений. М.: Наука, 1978.
8. *Потапова А. Ф.* Об ускорении сходимости метода скорейшего спуска.– Ж. вычисл. матем. и матем. физ., 1971, т. 11, № 3, с. 749–752.
9. *Емелин И. В.* О быстроте сходимости метода наискорейшего спуска.– Успехи матем. наук, 1977, т. 32, вып. 1(193), с. 163–164.
10. *Ерёмин И. И., Астафьев Н. Н.* Введение в теорию линейного и выпуклого программирования. М.: Наука, 1976.

Поступила в редакцию 6.V.1980
Переработанный вариант 25.VII.1980