ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 28, 1988 № 9

УДК 519.61

КОМБИНИРОВАННЫЕ ИТЕРАЦИОННЫЕ МЕТОДЫ ВАРИАЦИОННОГО ТИПА

БОНДАРЕНКО Л. Н., ЖУК П. Ф.

(Херсон)

Предложены итерационные методы решения линейных уравнений в гильбертовом пространстве, обобщающие s-шаговые методы наискорейшего спуска, минимальных невязок, α-процессы и метод двухступенчатого градиентного спуска. Получены оценки осредненных скоростей их сходимости.

Введение

1. В [1] был предложен итерационный метод двухступенчатого градиентного спуска (д. г. с.) решения систем линейных уравнений с положительно-определенной матрицей, каждая нечетная итерация которого выполняется по методу минимальных невязок, каждая четная — по методу наискорейшего спуска.

Скорость сходимости этого метода изучалась в [1], [2]. Поскольку теоретически ее оценить не удалось, были проведены численные эксперименты. Основные выводы состоят в следующем: для симметричных матриц метод д. г. с. предпочтительнее методов наискорейшего спуска и минимальных невязок, взятых в отдельности, и его скорость сходимости близка к скорости сходимости метода сопряженных градиентов; для несимметричных матриц д. г. с., вообще говоря, не эффективен.

В настоящей статье вводится понятие комбинированных итерационных методов вариационного типа, обобщающих метод д. г. с. в самосопряженном случае, и проводится теоретический анализ их на основе осредненных скоростей сходимости [3].

2. Пусть A — линейный самосопряженный ограниченный оператор с границами спектра $0<m<M$, действующий в вещественном гильбертовом пространстве H со скалярным произведением (u, v) и нормой $\|u\|=(u, u)^{0.5}$.

Последовательные приближения s-шагового α-процесса к решению u^* уравнения $Au=f$ задаются формулой

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

где α, s — соответственно, действительное и натуральное числа, $u_0^{(\alpha)}$ — произвольное начальное приближение, $z_k^{(\alpha)}=Au_k^{(\alpha)}-f$ — невязка, а действительные числа $\{\gamma_{ik}^{(\alpha)},\ i=1, 2, \ldots, s\}$ выбираются из условия минимума величины

$$F_\alpha(u_{k+1}^{(\alpha)})=\|u_{k+1}^{(\alpha)}-u^*\|_{A^{\alpha+1}}=(A^{\alpha+1}(u_{k+1}^{(\alpha)}-u^*), u_{k+1}^{(\alpha)}-u^*).\tag{2}$$

При $s=1$ получим α-процесс [4], при $\alpha=0$, 1 — это s-шаговые методы, соответственно, наискорейшего спуска и минимальных невязок.

Рассмотрим по аналогии с [5] операторный многочлен $P_s^{(\alpha)}(A, z)=$ $=E+\gamma_1^{(\alpha)}(z)A+\ldots+\gamma_s^{(\alpha)}(z)A^s$, удовлетворяющий условию $z_{k+1}^{(\alpha)}=P_s^{(\alpha)}(A, z_k^{(\alpha)})z_k^{(\alpha)}$, $k=0, 1, \ldots$. Из (1), (2) следует, что если $z, Az, \ldots, A^{s-1}z$ линейно независимы, то, аналогично [5],

$$P_s^{(\alpha)}(t, z)=(-1)^s \det \begin{vmatrix} \mu_\alpha & \mu_{\alpha+1} \cdots \mu_{\alpha+s-1} & 1 \\ \mu_{\alpha+1} & \mu_{\alpha+2} \cdots \mu_{\alpha+s} & t \\ \cdots & \cdots & \cdots \\ \mu_{\alpha+s} & \mu_{\alpha+s+1} \cdots \mu_{\alpha+2s-1} & t^s \end{vmatrix} \times \tag{3}$$

$$\times \left(\det \begin{vmatrix} \mu_{\alpha+1} \cdots \mu_{\alpha+s} \\ \cdots\cdots \\ \mu_{\alpha+s} \cdots \mu_{\alpha+2s-1} \end{vmatrix} \right)^{-1},$$

$$\mu_{\alpha+j}=(A^{\alpha+j}z, z), \qquad j=0, 1, \ldots, 2s-1;$$

кроме того, для любого $z\in H$

$$(A^{\alpha+j}P_s^{(\alpha)}(A, z)z, z)=0, \quad j=0, 1, \ldots, s-1. \tag{4}$$

Перейдем к определению комбинированного метода. Зафиксируем две последовательности: последовательность $S=(s_1, s_2, \ldots)$ натуральных чисел и последовательность действительных чисел $\Lambda=(\alpha_1, \alpha_2, \ldots)$.

Под (S, Λ)-комбинированным итерационным методом вариационного типа будем понимать итерационный процесс, k-я итерация которого совершается при помощи s_k-шагового α_k-процесса.

В частности, комбинированный метод с $S=S^{(2)}=(s_1, s_2, s_1, s_2, \ldots)$, $\Lambda=$ $=\Lambda^{(2)}=(\alpha_1, \alpha_2, \alpha_1, \alpha_2, \ldots)$ (т. е. с последовательностями пар s_1, s_2 и α_1, α_2 соответственно) назовем двухступенчатым.

Пусть $u_0, u_1, \ldots$ — итерационная последовательность, порожденная (S, Λ)-методом, а $z_k=Au_k-f$. Тогда

$$z_k = T_{s_k}^{(\alpha)k} \ldots T_{s_1}^{(\alpha_1)} z_0, \qquad k=1, 2, \ldots, \tag{5}$$

где $T_s^{(\alpha)}$ — оператор, определенный на H по формуле

$$T_s^{(\alpha)}z=P_s^{(\alpha)}(A, z)z. \tag{6}$$

В дальнейшем предполагаем, что $T_{s_k}^{(\alpha_k)} \ldots T_{s_1}^{(\alpha_1)} \neq 0$, $k=1, 2, \ldots$.

По аналогии с [3, с. 18, 94] величину

$$R_k=R_k(A, S, \Lambda, \lambda)=-\frac{\ln \| T_{s_k}^{(\alpha_k)} \ldots T_{s_1}^{(\alpha_1)} \|_\lambda}{s_1+\ldots+s_k}, \qquad k=1, 2, \ldots, \tag{7}$$

назовем осредненной скоростью сходимости (s, Λ)-метода за k итераций, а величину

$$R_\infty=R_\infty(A, S, \Lambda)=\liminf_{k\to\infty} R_k(A, S, \Lambda, \lambda)$$

назовем осредненной асимптотической скоростью сходимости. Здесь $\lambda\in R$, а

$$\| T \|_\lambda = \sup_{0\neq z\in H} \| Tz \|_{A^{\lambda-1}} \| z \|_{A^{\lambda-1}}^{-1}$$

является нормой оператора T.

В § 1 получены неулучшаемые оценки снизу для R_k (см. теорему 1 и замечание 2): если $\inf \Lambda\geqslant\lambda$, то $R_k\geqslant-\ln(1-m/M)$, $k=1, 2, \ldots, \infty$. Отсюда следует, в частности, что (S, Λ)-метод сходится по крайней мере со скоростью геометрической прогрессии.

В § 2 получены оценки R_k сверху для двухступенчатого метода $(S^{(2)}, \Lambda^{(2)})$ в конечномерном пространстве. Пусть $\pi_s(t)=1+\pi_1 t+\ldots+\pi_s t^s$ — многочлен степени $s=s_1+s_2$, наименее уклоняющийся от нуля на спектре оператора A, а ρ_s — величина уклонения. Из теоремы 2 следует, что

$$R_{2k}\leqslant R_\infty\leqslant -\frac{\ln\rho_s}{s},\qquad k=1,2,\ldots. \tag{8}$$

В частности, для д. г. с. ($s_1=s_2=1$, $\alpha_1=1$, $\alpha_2=0$, $s=2$)

$$R_{2k}\leqslant R_\infty\leqslant -0.5\ln\rho_2.$$

Оценка (8) позволяет сравнить двухступенчатый метод с s-шаговым методом. Пусть $s=s_1+s_2$, $\alpha\in R$. Из [6] следует, что $\|T_s^{(\alpha)}\|_\alpha\leqslant\rho_s$, поэтому

$$\|(T_s^{(\alpha)})^k\|_\alpha\leqslant\rho_s^k,\quad k=1,2,\ldots. \tag{9}$$

Если $n>s$, то из [7, с. 447] следует существование $z\in H$, $z\neq 0$, такого, что

$$\|(T_s^{(\alpha)})^k z\|_{A^{\alpha-1}}=\rho_s^k\|z\|_{A^{\alpha-1}},\quad k=1,2,\ldots. \tag{10}$$

Из (9), (10) вытекает, что при $n>s$

$$\|(T_s^{(\alpha)})^k\|_\alpha=\rho_s^k,\quad k=1,2,\ldots. \tag{11}$$

Рассматривая s-шаговый α-процесс как комбинированный метод с $S^{(1)}=(s,s,\ldots)$, $\Lambda^{(1)}=(\alpha,\alpha,\ldots)$, из (11) получаем

$$R_k(A,S^{(1)},\Lambda^{(1)},\alpha)=R_\infty(A,S^{(1)},\Lambda^{(1)})=-\frac{\ln\rho_s}{s},\qquad k=1,2,\ldots. \tag{12}$$

Из сравнения (8) и (12) можно сделать вывод о том, что s-шаговый метод с $s=s_1+s_2$ предпочтительнее (с точки зрения осредненных скоростей сходимости) двухступенчатого метода. В частности, двухшаговый метод наискорейшего спуска (или минимальных невязок) предпочтительнее д. г. с.

§ 1. Оценки снизу

Здесь и в дальнейшем используются обозначения из Введения.

Лемма 1. *Если $\lambda\leqslant\alpha$, то для любого $s\in\{1,2,\ldots\}$*

$$\|T_s^{(\alpha)}\|_\lambda\leqslant\rho^s,\quad \rho=1-m/M. \tag{1.1}$$

Доказательство. Зафиксируем произвольный элемент $z\in H$ и рассмотрим линеал L_z, состоящий из элементов вида $Q(A)z$, где $Q(t)$ — произвольный многочлен от t. Пополнив L_z предельными элементами, образуем подпространство H_z пространства H. Из (6) следует, что $\tilde z=T_s^{(\alpha)}z\in\in L_z\subseteq H_z$.

Рассмотрим два случая.

Случай 1. Подпространство H_z конечномерно. Сужение оператора A на H_z является линейным самосопряженным оператором, действующим из H_z в H_z. Поэтому z можно представить в виде

$$z=\eta_1 e_1+\ldots+\eta_n e_n, \tag{1.2}$$

где $\eta_i\neq 0$, $e_i\in H_z$ $(i=1,2,\ldots,n)$ — ортонормированные собственные элементы оператора A, причем действие оператора A^φ, $\varphi\in R$, на z таково:

(1.3) $$A^\varphi z=\lambda_1^\varphi\eta_1 e_1+\ldots+\lambda_n^\varphi\eta_n e_n,$$

где $m\leqslant\lambda_1<\ldots<\lambda_n\leqslant M$.

Случай 2. Подпространство H_z бесконечномерно. Это возможно лишь при условии, что элементы z, $Az,\ldots$ линейно независимы. Но тогда решение проблемы моментов (см., например, [8, с. 29]) определяет последовательность операторов

(1.4) $$A_n=E_nAE_n,\quad n=1,2,\ldots,$$

где E_n есть оператор проектирования в подпространство H_n — линейную оболочку элементов z, $Az,\ldots,A^{n-1}z$. Известно [8, с. 74], что сужение A_n на H_n является линейным самосопряженным оператором (действующим из H_n в H_n) с простыми собственными значениями $\{\lambda_i,\ i=1,2,\ldots,n\}$:

(1.5) $$m\leqslant\lambda_1<\ldots<\lambda_n\leqslant M.$$

Поэтому, как в случае 1, элемент z может быть представлен в виде (1.2), а $A_n^\varphi z$ — в виде (1.3).

Используем далее следующий факт [8, с. 30]: последовательность операторов A_n сильно сходится к A на подпространстве H_z (т. е. $\|A_nx-Ax\|\to\to0$, $n\to\infty$, для всех $x\in H_z$). Но тогда [8, с. 74] и последовательность A_n^φ сильно сходится к A^φ на H_z для любого $\varphi\in R$; поэтому из (3) следует, что $P_s^{(\alpha)}(A_n,\ z)z\to P_s^{(\alpha)}(A,\ z)z$, $n\to\infty$, в H. Таким образом, если для каждого $n=1,2,\ldots$ будет выполнено неравенство

(1.6) $$\|P_s^{(\alpha)}(A_n,z)z\|_{A_n^{\lambda-1}}\leqslant\rho^s\|z\|_{A_n^{\lambda-1}},$$

то справедлива оценка $\|T_s^{(\alpha)}z\|_{A^{\lambda-1}}\leqslant\rho^s\|z\|_{A^{\lambda-1}}$, а следовательно, и оценка (1.1) (так как z — произвольный элемент H).

Из рассмотренных выше случаев 1, 2 и из формул (1.2)—(1.6) заключаем, что неравенство (1.1) достаточно доказать в предположении, что $H=R^n$ есть n-мерное арифметическое пространство со скалярным произведением

$$(u,v)=\sum_{i=1}^n u_iv_i,$$

$$A=\operatorname{diag}(\lambda_1,\ldots,\lambda_n),\qquad m\leqslant\lambda_1<\ldots<\lambda_n\leqslant M,\quad n=1,2,\ldots.$$

Итак, имеем $z=(\eta_1,\ldots,\eta_n)\in R^n$, $\tilde z=T_s^{(\alpha)}z=P_s^{(\alpha)}(A,\ z)z=(P_1\eta_1,\ldots$ $\ldots,P_n\eta_n)$, где $P_i=P_s^{(\alpha)}(\lambda_i,\ z)$.

Для доказательства (1.1) достаточно показать, что

(1.7) $$\|\tilde z\|_{A^{\lambda-1}}\leqslant\rho^s\|z\|_{A^{\lambda-1}}.$$

Обозначим через $X=\{(x_1,\ldots,x_n)\}\subseteq R_+^n$ множество неотрицательных решений системы линейных уравнений

(1.8) $$\sum_{i=1}^n\lambda_i^{\lambda-1}x_i=\|z\|^2_{A^{\lambda-1}},$$

(1.9) $$\sum_{i=1}^n\lambda_i^{\alpha+j}P_ix_i=0,\qquad j=0,1,\ldots,s-1.$$

Пусть $x_0=(\eta_1^2,\ldots,\eta_n^2)$. Из (4) имеем

$$\sum_{i=1}^{n}\lambda_i^{\alpha+j}P_i\eta_i^2=(A^{\alpha+j}\tilde{z},z)=0,\qquad j=0,1,\ldots,s-1,$$

поэтому $X\neq\varnothing$ и представляет собой, как следует из определения, симплекс. Отметим, что любое базисное решение (1.8), (1.9), а соответственно, и любая вершина симплекса X содержит не более $s+1$ ненулевой компоненты.

Рассмотрим линейную функцию

$$\mathcal{F}(x)=\sum_{i=1}^{n}\lambda_i^{\lambda-1}P_i^2x_i.$$

Поскольку

$$\mathcal{F}(x_0)=\sum_{i=1}^{n}\lambda_i^{\lambda-1}P_i^2\eta_i^2=\|\tilde{z}\|^2_{A^{\lambda-1}},$$

то неравенство (1.7) эквивалентно неравенству

$$(1.10)\qquad \mathcal{F}(x_0)\leqslant\rho^{2s}\sum_{i=1}^{n}\lambda_i^{\lambda-1}\eta_i^2,$$

а так как $\max\limits_{x\in X}\mathcal{F}(x)$ достигается по крайней мере в одной из вершин симплекса X, то (1.10) следует (с учетом (1.8)) из неравенства

$$(1.11)\qquad \mathcal{F}(x)\leqslant\rho^{2s}\sum_{i=1}^{n}\lambda_i^{\lambda-1}x_i,$$

где x — произвольная вершина X.

Так как любая вершина симплекса содержит не более $s+1$ ненулевой компоненты, то достаточно рассмотреть следующие два случая.

Случай а. Вершина $x=(x_1,\ldots,x_n)\in X$ содержит не более s ненулевых компонент. Из (1.9) следует, что $P_ix_i=0$, $i=1,2,\ldots,n$; таким образом, $\mathcal{F}(x)=0$ и неравенство (1.11) верно.

Случай б. Вершина x содержит ровно $s+1$ ненулевую компоненту. Не ограничивая общности рассуждений для простоты записи, считаем, что $x_i>0$ для $i=1,2,\ldots,s+1$, $x_i=0$ для $i>s+1$. Так как

$$(1.12)\qquad P_i=P_s^{(\alpha)}(\lambda_i,z)=1+\gamma_1^{(\alpha)}(z)\lambda_i+\ldots+\gamma_s^{(\alpha)}(z)\lambda_i^s,$$

то, подставляя вместо P_i в (1.9) правую часть (1.12), получаем систему s линейных уравнений относительно $\{\gamma_i^{(\alpha)}(z),\ i=1,2,\ldots,s\}$; разрешив ее и подставив вместо $\{\gamma_i^{(\alpha)}(z),\ i=1,2,\ldots,s\}$ в (1.12) их выражения через $\{\lambda_i,x_i,\ i=1,2,\ldots,s+1\}$, получим

$$(1.13)\qquad P_i=(-1)^s\det\begin{vmatrix}\mu_\alpha & \mu_{\alpha+1}\cdots\mu_{\alpha+s-1} & 1\\ \mu_{\alpha+1} & \mu_{\alpha+2}\cdots\mu_{\alpha+s} & \lambda_i\\ \cdots & \cdots\cdots & \cdots\\ \mu_{\alpha+s} & \mu_{\alpha+s+1}\cdots\mu_{\alpha+2s-1} & \lambda_i^s\end{vmatrix}\times$$

$$\times\left(\det\begin{vmatrix}\mu_{\alpha+1}\cdots\mu_{\alpha+s}\\ \cdots\cdots\\ \mu_{\alpha+s}\cdots\mu_{\alpha+2s-1}\end{vmatrix}\right)^{-1},$$

где

$$\mu_{\alpha+j}=\sum_{i=1}^{s+1}\lambda_i^{\alpha+j}x_i \qquad j=0,1,\ldots,2s-1.$$

Первый определитель в правой части (1.13) легко преобразуется к виду

(1.14) $$(-1)^{s+i+1}\lambda_1^{\alpha}\ldots\lambda_{i-1}^{\alpha}\lambda_{i+1}^{\alpha}\ldots\lambda_{s+1}^{\alpha}x_1\ldots x_{i-1}x_{i+1}\ldots$$
$$\ldots x_{s+1}\Delta_{1,2,\ldots,i-1,i+1,\ldots,s+1}\Delta_{1,2,\ldots,s+1};$$

непосредственное вычисление второго определителя дает

(1.15) $$\sum_{j=1}^{s+1}\lambda_1^{\alpha+1}\ldots\lambda_{j-1}^{\alpha+1}\lambda_{j+1}^{\alpha+1}\ldots\lambda_{s+1}^{\alpha+1}x_1\ldots x_{j-1}x_{j+1}\ldots$$
$$\ldots x_{s+1}\Delta^2_{1,2,\ldots,j-1,j+1,\ldots,s+1},$$

где

$$\Delta_{i_1\ldots i_N}=\prod_{1\leqslant p<q\leqslant N}(\lambda_{i_q}-\lambda_{i_p}), \qquad \Delta_i=1.$$

Заменяя определители в (1.13) выражениями (1.14), (1.15), получаем

$$P_i=(-1)^{i+1}\lambda_i^{-\alpha}x_i^{-1}\Delta_{1,2,\ldots,i-1,i+1,\ldots,s+1}\Delta_{1,2,\ldots,s+1}\times$$
$$\times\left(\lambda_1\ldots\lambda_{s+1}\sum_{j=1}^{s+1}\lambda_j^{-\alpha-1}x_j^{-1}\Delta^2_{1,2,\ldots,j-1,j+1,\ldots,s+1}\right)^{-1},$$
$$i=1,2,\ldots,s+1.$$

Полагая найденные выражения для P_i, $i=1, 2, \ldots, s+1$, в

$$\mathcal{F}(x)=\sum_{i=1}^{s+1}\lambda_i^{\lambda-1}P_i^2x_i,$$

имеем

(1.16) $$\mathcal{F}(x)=\left(\sum_{i=1}^{s+1}\lambda_i^{\lambda-2\alpha-1}x_i^{-1}\Delta^2_{1,2,\ldots,i-1,i+1,\ldots s+1}\Delta^2_{1,2,\ldots,s+1}\right)\times$$
$$\times\left(\sum_{k,j=1}^{s+1}\lambda_k^{-\alpha}\lambda_j^{-\alpha}x_k^{-1}x_j^{-1}Q_kQ_j\right)^{-1},$$

где $Q_j=\lambda_1\ldots\lambda_{j-1}\lambda_{j+1}\ldots\lambda_{s+1}\Delta^2_{1,2,\ldots,j-1,j+1,\ldots,s+1}$, $j=1,2,\ldots,s+1$. После подстановки полученного значения $\mathcal{F}(x)$ в (1.11) и умножения обеих частей этого неравенства на знаменатель правой части в (1.16), получаем эквивалентное неравенство

(1.17) $$\Delta^2_{1,2,\ldots,s+1}\sum_{i=1}^{s+1}\lambda_i^{\lambda-2\alpha-1}x_i^{-1}\Delta^2_{1,2,\ldots,i-1,i+1,\ldots,s+1}\leqslant$$
$$\leqslant\rho^{2s}\sum_{i,k,j=1}^{s+1}\lambda_i^{\lambda-1}\lambda_k^{-\alpha}\lambda_j^{-\alpha}x_k^{-1}x_j^{-1}Q_kQ_jx_i.$$

Для доказательства неравенства (1.17) сравним i-е слагаемое его левой части со слагаемым $\lambda_1^{\lambda-\alpha-1}\lambda_i^{-\alpha}Q_1Q_ix_i^{-1}$ правой части:

$$\lambda_i^{\lambda-2\alpha-1}\Delta^2_{1,2,\ldots,i-1,i+1,\ldots,s+1}\Delta^2_{1,2,\ldots,s+1}\times$$
$$\times(\lambda_1^{\lambda-\alpha-1}\lambda_i^{-\alpha}\lambda_2\ldots\lambda_{s+1}\Delta^2_{2,3,\ldots,s+1}\lambda_1\ldots\lambda_{i-1}\lambda_{i+1}\ldots$$
$$\ldots\lambda_{s+1}\Delta^2_{1,2,\ldots,i-1,i+1,\ldots,s+1})^{-1}=\frac{\lambda_i^{\lambda-\alpha}(\lambda_2-\lambda_1)^2\ldots(\lambda_{s+1}-\lambda_1)^2}{\lambda_1^{\lambda-\alpha}\lambda_2^2\ldots\lambda_{s+1}^2}\leqslant\rho^{2s}$$

при $\lambda\leqslant\alpha$. Лемма доказана.

Замечание 1. Пусть заданы $0<m<M$, λ, $\bar\rho\in]0,\rho[$ и гильбертово пространство H. Тогда если $\dim H\geqslant2$, то существуют числа $s\in\{1,2,\ldots\}$, $\alpha\geqslant\lambda$ и линейный самосопряженный оператор $A:H\to H$ с границами спектра m, M такие, что

(1.18) $\quad \|T_s^{(\alpha)}\|_\lambda>\bar\rho^s.$

Для доказательства возьмем ортонормированные элементы $e_1,e_2\in H$, т. е. $\|e_1\|=\|e_2\|=1$, $(e_1,e_2)=0$ (это возможно, так как $\dim H\geqslant2$). Обозначим через H_2 линейную оболочку e_1,e_2, а через $H_2^\perp$ – ортогональное дополнение к H_2. Известно (см., например, [8, с. 14]), что всякий элемент $z\in H$ однозначно представим в виде $z=v+w$, где $v=\xi e_1+\eta e_2\in H_2$, $w\in H_2^\perp$. Зададим оператор A на пространстве H формулой $Az=Av+Aw=m\xi e_1+M\eta e_2+Mw$; тогда A – линейный оператор и $(z,z)=\xi^2+\eta^2+(w,w)$, $(Az,z)=m\xi^2+M\eta^2+M(w,w)$, $m(z,z)\leqslant(Az,z)\leqslant M(z,z)$, $(Az_1,z_2)=m\xi_1\xi_2+M\eta_1\eta_2+M(w_1,w_2)=(z_1,Az_2)$, где $z_i=\xi_ie_1+\eta_ie_2$, $i=1,2$. Следовательно, $mE\leqslant A=A^*\leqslant ME$.

Положим $s=1$, $z(\eta)=e_1+\eta e_2$, $\delta=m/M$,

(1.19) $\quad g(\alpha,\eta)=\|T_1^{(\alpha)}z(\eta)\|_{A^{\lambda-1}}/\|z(\eta)\|_{A^{\lambda-1}}.$

При $\eta\neq0$ получим

(1.20а) $\quad T_1^{(\alpha)}z(\eta)=\rho(\alpha,\eta)(e_1-\delta^\alpha\eta^{-1}e_2),$

(1.20б) $\quad g(\alpha,\eta)=\rho(\alpha,\eta)\dfrac{(1+\eta^{-2}\delta^{2\alpha-\lambda+1})^{0.5}}{(1+\eta^2\delta^{1-\lambda})^{0.5}},$

где $\rho(\alpha,\eta)=\rho(1+\eta^{-2}\delta^{\alpha+1})^{-1}$. Выберем $\varepsilon>0$ из условия

(1.21) $\quad \rho>\bar\rho(1+\varepsilon^2\delta^{1-\lambda})^{0.5}$

и положим $U_\varepsilon=\{\eta:0<|\eta|\leqslant\varepsilon\}$. Так как $0<\delta<1$, то при $\eta\in U_\varepsilon$ из (1.20), (1.21) следует

$$\lim_{\alpha\to\infty}g(\alpha,\eta)>\bar\rho;$$

поэтому для всех $\eta\in U_\varepsilon$ существует $\alpha(\eta)$ такое, что для всех $\alpha\geqslant\alpha(\eta)$

(1.22) $\quad g(\alpha,\eta)>\bar\rho.$

Тогда $U=\{(\alpha,\eta):\eta\in U_\varepsilon,\ \alpha\geqslant\alpha(\eta)\}\neq\varnothing$ и для всякого $(\alpha,\eta)\in U$ из (1.19), (1.22) будет $\|T_1^{(\alpha)}\|_\lambda\geqslant g(\alpha,\eta)>\bar\rho$, т. е. неравенство (1.18) для $s=1$ и указанных α верно.

Теорема 1. *Если* $\inf\Lambda>-\infty$, *то* (S,Λ)-*метод сходится, причем*

(1.23) $\quad R_k(A,S,\Lambda,\lambda)\geqslant-\ln\rho \qquad \forall\lambda\leqslant\inf\Lambda,\quad k=1,2,\ldots,\infty.$

Доказательство. Пусть $\inf\Lambda>-\infty$. Возьмем $\lambda\leqslant\inf\Lambda$. Из леммы 1 следует, что

(1.24) $\quad \|T_{s_k}^{(\alpha_k)}\|_\lambda\leqslant\rho^{s_k},\qquad k=1,2,\ldots.$

Поэтому из (5) получаем

$$\|z_k\|_{A^{\lambda-1}}\leqslant\rho^{s_1+\ldots+s_k}\|z_0\|_{A^{\lambda-1}}\to0,\qquad k\to\infty.$$

Следовательно, $\|u_k-u^*\|\to0$, $k\to\infty$, для произвольного начального приближения u_0 и (S, Λ)-метод сходится. Из (1.24) имеем

$$R_k = (A, S, \Lambda, \lambda) = -\frac{\ln \| T_{s_k}^{(\alpha_k)} \ldots T_{s1}^{(\alpha_1)} \|_\lambda}{s_1 + \ldots + s_k} \geqslant -\ln \rho,$$

$$k = 1, 2, \ldots,$$

поэтому и

$$R_\infty = \liminf_{k\to\infty} R_k \geqslant -\ln \rho.$$

Теорема доказана.

Замечание 2. Пусть заданы числа $0<m<M$, λ, $\bar\rho\in]0, \rho[$ и гильбертово пространство H с $\dim H\geqslant 2$. Тогда существует линейный оператор $A: H\to H$ $(mE\leqslant A = A^*\leqslant ME)$ и (S, Λ)-метод с $\inf\Lambda\geqslant\lambda$ такие, что

(1.25) $\quad R_k\leqslant -\ln\bar\rho, \qquad k=1, 2, \ldots, \infty.$

Пусть A, $z(\eta)$, δ, $g(\alpha, \eta)$, $\rho(\alpha, \eta)$, U_ε, U означают то же, что и в замечании 1. Положим $S=(1, 1, \ldots)$ и укажем способ построения $\Lambda=(\alpha_1, \alpha_2, \ldots)$. Для $\eta_0\in U_\varepsilon$ положим

(1.26) $\quad \alpha_i = \max\{\alpha(\eta_{i-1}), \nu(\eta_{i-1}), \lambda\}, \qquad \eta_i = -\delta^{\alpha_i}\overset{-1}{\eta_{i-1}}, \quad i=1, 2, \ldots,$

где $\nu(\eta) = \ln|\varepsilon\eta|/\ln\delta$. Докажем, что $\eta_i\in U_\varepsilon$, $i=1, 2, \ldots$.

Действительно, если $\eta_{i-1}\in U_\varepsilon$, то из (1.26) следует, что

$$\alpha_i\geqslant\nu(\eta_{i-1})\Rightarrow 0<|\delta^{\alpha_i}\overset{-1}{\eta_{i-1}}|\leqslant\varepsilon\Rightarrow 0<|\eta_i|\leqslant\varepsilon\Rightarrow\eta_i\in U_\varepsilon;$$

так как $\eta_0\in U_\varepsilon$, то $\eta_i\in U_\varepsilon$ для всех $i=1, 2, \ldots$. Но тогда из (1.26) и из определения U следует, что

(1.27) $\quad (\alpha_i, \eta_{i-1})\in U, \qquad i=1, 2, \ldots$.

Используя однородность оператора $T_1^{(\alpha)}$, из (1.20), (1,26) получаем $T_1^{(\alpha_i-1)}\ldots$ $\ldots T_1^{(\alpha_1)} z(\eta_0) = \rho(\alpha_{i-1}, \eta_{i-2})\ldots\rho(\alpha_1, \eta_0) z(\eta_{i-1})$, $i=2, 3, \ldots$, поэтому (см. (1.19))

(1.28) $$\frac{\| T_1^{(\alpha_i)}\ldots T_1^{(\alpha_1)} z(\eta_0)\|_{A^{\lambda-1}}}{\| T_1^{(\alpha_i-1)}\ldots T_1^{(\alpha_1)} z(\eta_0)\|_{A^{\lambda-1}}} = \frac{\| T_1^{(\alpha_i)} z(\eta_{i-1})\|_{A^{\lambda-1}}}{\| z(\eta_{i-1})\|_{A^{\lambda-1}}} = g(\alpha_i, \eta_{i-1}), \quad i = 1, 2, \ldots.$$

Из этого соотношения вытекает, что

$$\frac{\| T_1^{(\alpha_k)}\ldots T_1^{(\alpha_1)} z(\eta_0)\|_{A^{\lambda-1}}}{\| z(\eta_0)\|_{A^{\lambda-1}}} = g(\alpha_k, \eta_{k-1})\cdots g(\alpha_1, \eta_0), \qquad k = 1, 2, \ldots,$$

поэтому из (1.27), (1.22) следует $\|T_1^{(\alpha_k)}\ldots T_1^{(\alpha_1)}\|_\lambda > \bar\rho^k$, т. е. $R_k < -\ln\bar\rho$, $k=1, 2, \ldots$, и неравенство (1.25) для указанного оператора A и (S, Λ)-метода верно.

§ 2. Оценки сверху

Рассмотрим двухступенчатый метод $(S^{(2)}, \Lambda^{(2)})$ с $S^{(2)} = (s_1, s_2, s_1, s_2, \ldots)$, $\Lambda^{(2)} = (\alpha_1, \alpha_2, \alpha_1, \alpha_2, \ldots)$. Положим

(2.1) $$T_i = T_{s_i}^{(\alpha_i)}, \qquad i = 1, 2, \quad T = T_2 T_1.$$

Осредненные скорости сходимости двухступенчатого метода определены при $T\neq 0$ выражениями (см. формулу (7))

(2.2a) $$R_{2k-1} = R_{2k-1}(A, S^{(2)}, \Lambda^{(2)}, \lambda) = -\frac{\ln\| T_1 T^{k-1}\|_\lambda}{ks_1 + (k-1)s_2},$$

$$(2.2б)\qquad R_{2k}=R_{2k}(A, S^{(2)}, \Lambda^{(2)}, \lambda)=-\frac{\ln\|T^k\|_\lambda}{k(s_1+s_2)},\qquad k=1, 2, \ldots,$$

$$(2.2в)\qquad R_\infty=R_\infty(A, S^{(2)}, \Lambda^{(2)})=\liminf_{k\to\infty} R_k$$

и при

$$(2.3)\qquad \lambda\leqslant\min\{\alpha_1, \alpha_2\}$$

удовлетворяют оценке (1.23) (отметим, что R_∞ не зависит от λ).

Так как $\|T_2\|_\lambda^{-1}\|T^k\|_\lambda\leqslant\|T_1T^{k-1}\|_\lambda\leqslant\|T_1\|_\lambda\|T^{k-1}\|_\lambda$, то

$$(2.4а)\qquad R_{2k-1}\geqslant-\frac{\ln\|T_1\|_\lambda}{ks_1+(k-1)s_2}+\left[1-\frac{s_1}{ks_1+(k-1)s_2}\right]R_{2k-2},$$

$$(2.4б)\qquad R_{2k-1}\leqslant\frac{\ln\|T_2\|_\lambda}{ks_1+(k-1)s_2}+\left[1+\frac{s_2}{ks_1+(k-1)s_2}\right]R_{2k}.$$

В частности, при условии (2.3) получаем

$$(2.5)\qquad \left[1-\frac{s_1}{ks_1+(k-1)s_2}\right]R_{2k-2}\leqslant R_{2k-1}\leqslant\left[1+\frac{s_2}{ks_1+(k-1)s_2}\right]R_{2k}.$$

Рассмотрим последовательность $a_k=\ln\|T^k\|_\lambda$, $k=1, 2, \ldots$. Так как $a_{k+i}\leqslant a_k+a_i$, k, $i=1, 2, \ldots$, то последовательность $\{a_k/k\}$ либо сходится к своей нижней грани, либо расходится к $-\infty$ (см., например, [9, с. 37]); поэтому для всех $\lambda\in R$ (см. также [3, с. 19])

$$R_{2k}\leqslant\lim_{k\to\infty} R_{2k},\qquad k=1, 2, \ldots$$

(если R_{2k} расходится к ∞, то полагаем $\lim\limits_{k\to\infty} R_{2k}=\infty$). Но тогда из (2.4) имеем

$$(2.6)\qquad R_{2k}\leqslant R_\infty=\lim_{k\to\infty} R_k\qquad \forall\lambda\in R,\quad k=1, 2, \ldots.$$

Здесь R_k, R_∞ оцениваются сверху в предположении конечномерности пространства H; при этом используется установленное ниже свойство общего (S, Λ)-метода, справедливое, вообще говоря, для произвольных гильбертовых пространств.

Зафиксируем $s\in\{1, 2, \ldots\}$ и $\alpha\in R$ и обозначим через $\mathcal{N}$ множество всех $z\in H$, для которых $T_s^{(\alpha)}z\neq0$ или, что эквивалентно, при которых элементы z, $Az, \ldots, A^sz$ линейно независимы.

Положим $\mathcal{P}(L)=\{P_s^{(\alpha)}(t, z): z\in\mathcal{N}\cap L\}$, где L — произвольное подмножество H. Отметим, что если $z\in\mathcal{N}$, то $P_s^{(\alpha)}(t, z)$ — многочлен от t — для данного z определен однозначно и его степень равна s.

Пусть $P(t)=1+\gamma_1t+\ldots+\gamma_st^s$. Установим необходимое и достаточное условие принадлежности $P(t)$ к $\mathcal{P}(L)$.

Лемма 2. *Для того чтобы $P(t)\in\mathcal{P}(L)$, необходимо и достаточно существования элемента $z\in\mathcal{N}\cap L$ такого, что*

$$(2.7)\qquad (A^{\alpha+j}\bar z_1, z)=0,\qquad j=0, 1, \ldots, s-1,$$

где $\bar z_1=P(A)z$.

Доказательство. Необходимость. Пусть $P(t)\in\mathcal{P}(L)$; следовательно, $\mathcal{P}(L)\neq\varnothing$ и существует элемент $z\in\mathcal{N}\cap L$ такой, что $P_s^{(\alpha)}(t, z)\equiv\equiv P(t)$. Тогда $\bar z_1=P_s^{(\alpha)}(A, z)z$. В силу (4),

$$(2.8)\qquad (A^{\alpha+j}P_s^{(\alpha)}(A, z)z, z)=0,\qquad j=0, 1, \ldots, s-1,$$

а значит, для указанного z соотношение (2.7) выполнено.

Достаточность. Пусть для некоторого $z\in\mathcal{N}\cap L$ соотношение (2.7) выполнено. Рассмотрим $P_s^{(\alpha)}(t, z)-P(t)=tQ(t)$, где $Q(t)=(\gamma_1^{(\alpha)}(z)-\gamma_1)+\ldots+(\gamma_s^{(\alpha)}(z)-\gamma_s)t^{s-1}$.

Найдем $\{q_i=\gamma_i^{(\alpha)}(z)-\gamma_i,\ i=1, 2, \ldots, s\}$. Из (2.7), (2.8) следует $(A^{\alpha+j+1}Q(A)z, z)=0$, $j=0, 1, \ldots, s-1$, или

$$\sum_{i=1}^{s}\mu_{\alpha+j+i}q_i=0, \qquad j=0, 1, \ldots, s-1, \quad \mu_{\alpha+j+i}=(A^{\alpha+j+i}z, z). \tag{2.9}$$

Так как элементы $z, Az, \ldots, A^s z$ линейно независимы, то

$$\det|\mu_{\alpha+j+i}|_{\substack{i=1,2\ldots,s\\ j=0,1,\ldots,s-1}}\neq 0;$$

следовательно, система линейных относительно $\{q_i,\ i=1, 2, \ldots, s\}$ уравнений (2.9) имеет лишь тривиальное решение, т. е. $P(t)\equiv P_s^{(\alpha)}(t, z)\in\in\mathcal{P}(L)$. Лемма доказана.

Замечание 3. Из доказательства леммы 2 следует, что если многочлен $P(t)$ указанного выше вида и $z\in\mathcal{N}$ удовлетворяют соотношению (2.7), то $P(t)\equiv P_s^{(\alpha)}(t, z)$, $\deg P(t)=s$.

Лемма 2 допускает обобщение на (S, Λ)-метод. Обозначим через $\mathcal{N}_k$, $k=1, 2, \ldots$, множество всех z, для которых

$$z_1=T_{s_1}^{(\alpha_1)}z\neq 0, \ldots, z_k=T_{s_k}^{(\alpha_k)}z_{k-1}\neq 0. \tag{2.10}$$

Отметим, что $\mathcal{N}_k\neq\varnothing$, так как ранее предполагалось, что $T_{s_k}^{(\alpha_k)}\ldots T_{s_1}^{(\alpha_1)}\neq 0$. Положим

$$\mathcal{P}_k(L)=\{(P_{s_1}^{(\alpha_1)}(t, z), \ldots, P_{s_k}^{(\alpha_k)}(t, z_{k-1})): z\in\mathcal{N}_k\cap L\}. \tag{2.11}$$

Множество $\mathcal{P}_k(L)$ представляет собой совокупность упорядоченных наборов, состоящих из k многочленов от переменной t, порожденных первыми k итерациями (S, Λ)-метода из начальной невязки $z\in\mathcal{N}_k\cap L$. В частности, $\mathcal{P}_1(L)=\mathcal{P}(L)$ (с $s=s_1$, $\alpha=\alpha_1$). Отметим, что если $z\in\mathcal{N}_k$, то $P_{s_i}^{(\alpha_i)}(t, z_{i-1})$, $i=1, 2, \ldots, k$, $z_0=z$ (многочлен от t), имеет степень s_i и определен однозначно, следовательно, каждому $z\in\mathcal{N}_k$ соответствует лишь один упорядоченный набор из k многочленов указанного вида.

Пусть $P_i(t)$ — многочлен от t степени не выше s_i и $P_i(0)=1$, $i=1, 2, \ldots, k$.

Лемма 3. *Для того чтобы упорядоченный набор* $(P_1(t), \ldots, P_k(t))\in\in\mathcal{P}_k(L)$, *необходимо и достаточно существования элемента* $z\in\mathcal{N}_k\cap L$ *такого, что*

$$(A^{\alpha_i+j}\bar z_i, \bar z_{i-1})=0, \qquad i=1, 2, \ldots, k, \quad j=0, 1, \ldots, s_i-1, \tag{2.12}$$

где $\bar z_0=z$, $\bar z_i=P_i(A)\bar z_{i-1}$, $i=1, 2, \ldots, k$.

Доказательство. Необходимость. Пусть $(P_1(t), \ldots, P_k(t))\in\in\mathcal{P}_k(L)$. Тогда существует элемент $z\in\mathcal{N}_k\cap L$ такой, что $P_i(t)\equiv P_{s_i}^{(\alpha_i)}(t, z_{i-1})$, $i=1, 2, \ldots, k$, где $z_0=z$, z_i, $i=1, 2, \ldots, k$, определены в (2.10). Тогда $\bar z_1=$ $=P_1(A)z=P_{s_1}^{(\alpha_1)}(A, z)z=z_1$, $\bar z_2=P_2(A)\bar z_1=P_{s_2}^{(\alpha_2)}(A, z_1)z_1=z_2$ и т. д., $\bar z_i=z_i$, $i=$ $=1, 2, \ldots, k$. Равенства (2.12) теперь следуют из (4).

Достаточность. Пусть существует $z\in\mathcal{N}_k\cap L$, для которого соотношения (2.12) выполнены. Тогда при $i=1$ из (2.12) имеем

$$(2.13)\qquad (A^{\alpha_1+j}P_1(A)z,\, z)=0,\qquad j=0,1,\ldots,s_1-1.$$

Так как $z\in\mathcal{N}_k\cap L\subseteq\mathcal{N}_1\cap L$, то, в силу леммы 2 и замечания 3, из (2.13) следует $P_1(t)\in\mathcal{P}(L)$ (с $s=s_1$, $\alpha=\alpha_1$) и $P_1(t)\equiv P_{s_1}^{(\alpha_1)}(t,\,z)$. Но тогда $\bar z_1=z_1$, поэтому, полагая $i=2$ в (2.12), аналогично предыдущему получаем $P_2(t)\in\in\mathcal{P}(L)$ (с $s=s_2$, $\alpha=\alpha_2$), $P_2(t)\equiv P_{s_2}^{(\alpha_2)}(t,\,z_1)$, $\bar z_2=z_2$. Рассматривая аналогично $i=3,4,\ldots,k$, имеем $(P_1(t),\ldots,P_k(t))\equiv(P_{s_1}^{(\alpha_1)}(t,z),\ldots,P_{s_k}^{(\alpha_k)}(t,z_{k-1}))\in\in\mathcal{P}_k(L)$. Лемма доказана.

Замечание 4. Из доказательства леммы 3 следует, что если элемент $z\in\mathcal{N}_k$ и упорядоченный набор многочленов $(P_1(t),\ldots,P_k(t))$ указанного выше вида удовлетворяют соотношению (2.12), то $\bar z_i=z_i$ и

$$(2.14)\qquad P_{s_i}^{(\alpha_i)}(t,z_{i-1})\equiv P_i(t),\qquad i=1,2,\ldots,k.$$

Рассмотрим теперь двухступенчатый метод в конечномерном пространстве H. Положим, что $s=s_1+s_2$; $m=\lambda_1<\ldots<\lambda_n=M$ — собственные значения (возможно, кратные) оператора A, $\pi_s(t)=1+\pi_1t+\ldots+\pi_st^s$ — многочлен степени s, наименее уклоняющийся от нуля на спектре $\{\lambda_1,\ldots,\lambda_n\}$ оператора A, величина уклонения

$$\rho_s=\max_{i=1,2,\ldots,n}|\pi_s(\lambda_i)|.$$

В дальнейшем предполагаем, что $n>s$; тогда $T\neq0$ на H (см. (2.1)), $\rho_s>0$ и существует множество индексов $I=\{i_1,\ldots,i_{s+1}\}$, $1=i_1<\ldots<i_{s+1}==n$, таких, что (см., например, [10])

$$(2.15)\qquad \pi_s(\lambda_{i_j})=(-1)^{j+1}\rho_s,\qquad j=1,2,\ldots,s+1.$$

Пусть e_i, $i\in I$,— некоторый нормированный собственный элемент оператора A, соответствующий собственному значению λ_i. Обозначим через H_I линейную оболочку элементов $\{e_i,\ i\in I\}$; очевидно, что H_I является подпространством H, а элементы $\{e_i,\ i\in I\}$ образуют ортонормированный базис в H_I. Так как $\dim H_I=s+1$, то $T\neq0$ на H_I и, следовательно, $\mathcal{N}_2\cap H_I\neq\varnothing$.

Лемма 4. *Существует элемент $z_0\in\mathcal{N}_2\cap H_I$ такой, что*

$$(2.16)\qquad P_{s_2}^{(\alpha_2)}(t,z_1)\,P_{s_1}^{(\alpha_1)}(t,z_0)\equiv\pi_s(t),$$

где $z_1=T_1z_0$ (см. (2.1)).

Доказательство. Рассмотрим множество $\mathcal{M}$ различных упорядоченных пар $(P_1(t),\,P_2(t))$ многочленов таких, что $P_i(0)=1$, $i=1,2$, и

$$(2.17)\qquad P_2(t)P_1(t)\equiv\pi_s(t).$$

Отметим, что между множеством $\mathcal{M}$ и множеством упорядоченных разбиений совокупности $\{1,2,\ldots,s\}$ на два класса J_1, J_2 (т. е. $J_1\cap J_2=\varnothing$, $J_1\cup J_2=\{1,2,\ldots,s\}$) существует взаимно однозначное соответствие:

$$(2.18)\qquad J_i\leftrightarrow P_i(t)=\prod_{j\in J_i}(1-t/\theta_j),\qquad i=1,2,$$

где $P_i(t)\equiv1$ при $J_i=\varnothing$, $\theta_1<\ldots<\theta_s$ — корни $\pi_s(t)$; из (2.15) следует, что

$$(2.19)\qquad \theta_j\in]\lambda_{i_j},\lambda_{i_{j+1}}[,\qquad j=1,2,\ldots,s.$$

Пусть $\mathscr{P}_2(H_I)$ — множество, определенное в (2.11) при $k=2$, $L=H_I$. Для доказательства леммы достаточно показать, что

(2.20) $$\mathscr{M}\cap\mathscr{P}_2(H_I)\neq\varnothing.$$

Покажем предварительно, что существует пара $(P_1(t),\ P_2(t))\in\mathscr{M}$ и $0\neq z\in H_I$, удовлетворяющие (2.12) при $k=2$, т. е.

(2.21) $$(A^{\alpha_1+j}\bar{z}_1,\ z)=0,\quad j=0,1,\ldots,s_1-1,$$

(2.22) $$(A^{\alpha_2+j}\bar{z}_2,\ \bar{z}_1)=0,\quad j=0,1,\ldots,s_2-1,$$

где $\bar{z}_1=P_1(A)z$, $\bar{z}_2=P_2(A)\bar{z}_1$ (отметим, что условие $\deg P_i(t)\leqslant s_i$, $i=1,2$, не требуется).

Пусть

(2.23) $$z=\sum_{i\in I}\eta_i e_i.$$

После подстановки z в (2.21), (2.22) и учета (2.15), (2.17) получаем

$$\sum_{i\in I}\lambda_i^{\alpha_1+j}P_1(\lambda_i)\eta_i^2=0,\qquad j=0,1,\ldots,s_1-1,$$

$$\sum_{i\in I}\lambda_i^{\alpha_2+j}\operatorname{sign}(\pi_s(\lambda_i))\,P_1(\lambda_i)\eta_i^2=0,\qquad j=0,1,\ldots,s_2-1,$$

или, вводя новые переменные

(2.24) $$y_i=P_1(\lambda_i)\eta_i^2,\qquad i\in I,$$

имеем

(2.25а) $$\sum_{i\in I}\lambda_i^{\alpha_1+j}y_i=0,\qquad j=0,1,\ldots,s_1-1,$$

(2.25б) $$\sum_{i\in I}\lambda_i^{\alpha_2+j}\operatorname{sign}(\pi_s(\lambda_i))\,y_i=0,\qquad j=0,1,\ldots,s_2-1.$$

Система линейных уравнений (2.25) содержит s_1+s_2 уравнений относительно s_1+s_2+1 переменных $\{y_i,\ i\in I\}$, следовательно, она имеет нетривиальное решение $y^*=(y_i^*,\ i\in I)$ (отметим, что $-y^*=-y_i^*$ $i\in I$, также является решением (2.25)). Положим $J=\{i\in I\colon y_i^*\neq 0\}$; количество элементов множества J таково:

(2.26) $$N(J)>\max\{s_1,\ s_2\}.$$

Действительно, пусть, например, $N(J)\leqslant s_1$; из (2.25) следует, что

(2.27) $$\sum_{i\in J}\lambda_i^{\alpha_1+j}y_i^*=0,\qquad j=0,1,\ldots,s_1-1,$$

что невозможно, так как система (2.27) имеет (в случае $N(J)\leqslant s_1$) лишь тривиальное решение $y_i^*=0$, $i\in J$. Аналогично доказывается, что $N(J)>s_2$.

Без ограничения общности можно считать, что $y_{i_1}^*\geqslant 0$ (иначе в качестве нетривиального решения нужно взять $-y^*$). Выделим из $y_{i_1}^*,\ldots,y_{i_{s+1}}^*$ ненулевые значения $y_{i_{j_1}}^*,\ldots,y_{i_{j_p}}^*$, где $1\leqslant j_1<\ldots<j_p\leqslant s+1$. Образуем множество $J_1\subseteq\{1,2,\ldots,s\}$ по следующему правилу: $j_l\in J_1$ тогда и только тогда, когда $y_{i_{j_l}}^*y_{i_{j_{l+1}}}^*<0$, $l=1,2,\ldots,p-1$. Пусть $J_2=\{1,2,\ldots,s\}\setminus J_1$; пара $(J_1,\ J_2)$ порождает разбиение $\{1, 2, \ldots, s\}$, и ему соответствует (см. (2.18)) пара многочленов $(P_1^*(t), P_2^*(t))\in\mathcal{M}$ таких, что $P_1^*(\lambda_i)y_i^*\geqslant 0$, $i\in I$ (в силу построения J_1 и (2.19)).

Так как (см. (2.19))

$$P_1^*(\lambda_i)\neq 0, \qquad i\in I, \tag{2.28}$$

то из (2.24) находим $\eta_i^*=[y_i^*/P_1^*(\lambda_i)]^{0,5}$, $i\in I$, причем $\eta_i^*\neq 0$ при $i\in J$. Рассмотрим

$$z^*=\sum_{i\in I}\eta_i^* e_i\in H_I,\quad \bar z_1^*=P_1^*(A)z^*,\quad \bar z_2^*=P_2^*(A)\bar z_1^*. \tag{2.29}$$

Повторяя рассуждения, начиная с формулы (2.23), для элемента z^*, убеждаемся, что элементы z^*, $\bar z_1^*$, $\bar z_2^*$ удовлетворяют соотношениям (2.21), (2.22). Таким образом, предварительная задача решена.

Покажем теперь, что

$$\deg P_i^*(t)\leqslant s_i, \qquad i=1, 2, \quad z^*\in\mathcal{N}_2. \tag{2.30}$$

Заметим, что поскольку $\deg P_1^*(t)+\deg P_2^*(t)=s=s_1+s_2$, то либо $\deg P_1^*(t)\leqslant s_1$, либо $\deg P_2^*(t)\leqslant s_2$.

Пусть $\deg P_1^*(t)\leqslant s_1$. В силу (2.26), z^* в своем разложении (2.29) имеет $N(J)>s_1$ ненулевых компонент. Следовательно, элементы z^*, $Az^*, \ldots, A^{s_1}z^*$ линейно независимы, т. е. $z^*\in\mathcal{N}$ (с $s=s_1$, $\alpha=\alpha_1$). Так как z^*, $\bar z_1^*$ удовлетворяют соотношению (2.21), то, в силу замечания 3, имеем

$$P_1^*(t)\equiv P_{s_1}^{(\alpha_1)}(t, z^*), \tag{2.31}$$

причем $\deg P_1^*(t)=s_1$. Тогда $\deg P_2^*(t)=s_2$.

В силу (2.26), (2.28), (2.31), разложение элемента $z_1^*=T_1z^*$ по базису $\{e_i,\ i\in I\}$ содержит $N(J)>s_2$ ненулевых компонент, т. е. элементы z_1^*, $Az_1^*, \ldots, A^{s_2}z_1^*$ линейно независимы. Но тогда $Tz^*=T_2z_1^*\neq 0$, т. е. $z^*\in\mathcal{N}_2$. Соотношения в (2.30) для первого случая доказаны.

Пусть $\deg P_2^*(t)\leqslant s_2$. Из (2.26), (2.28) следует, что разложение $\bar z_1^*$ содержит $N(J)>s_2$ ненулевых компонент, поэтому $T_2\bar z_1^*\neq 0$, т. е. $\bar z_1^*\in\mathcal{N}$ (с $s=s_2$, $\alpha=\alpha_2$). Так как, кроме того, элементы $\bar z_1^*$, $\bar z_2^*$ удовлетворяют соотношению (2.22), то из леммы 2 вытекает $\deg P_2^*(t)=s_2$. Но тогда $\deg P_1^*(t)=s_1$, и мы возвращаемся к первому случаю. Соотношения в (2.30) доказаны.

Таким образом, для набора $(P_1^*(t), P_2^*(t))$ существует элемент $z=z^*\in\mathcal{N}_2\cap H_I$ такой, при котором равенства (2.12) (при $k=2$) выполнены. Поэтому, в силу леммы 3, $(P_1^*(t), P_2^*(t))\in\mathcal{P}_2(H_I)$, т. е. соотношение (2.20) верно. В качестве z_0 можно взять z^*. Лемма доказана.

Лемма 5. *Если $n>s$, то существует элемент $0\neq z_0\in H$ такой, что*

$$\|T^k z_0\|_{A^{\lambda-1}}=\rho_s^k\|z_0\|_{A^{\lambda-1}}\qquad \forall\lambda\in R,\quad k=1, 2, \ldots.$$

Доказательство. Рассмотрим z_0 из леммы 4. Так как $z_0\in\mathcal{N}_2\cap H_I$, то

$$z_0=\sum_{i\in I}\eta_i e_i\neq 0,$$

и, в силу (2.15), (2.16), имеем

$$Tz_0=\pi_s(A)z_0=\rho_s\sum_{i\in I}\operatorname{sign}[\pi_s(\lambda_i)]\eta_i e_i. \tag{2.32}$$

Положим

$$\hat{z}_0 = \sum_{i \in I} \operatorname{sign}[\pi_s(\lambda_i)]\, \eta_i e_i.$$

Тогда

$$(A^{\alpha+j}\hat{z}_0, \hat{z}_0) = \sum_{i \in I} \lambda_i^{\alpha+j} \eta_i^2 = (A^{\alpha+j} z_0, z_0) \qquad \forall \alpha \in R, \quad j = 0, 1, \ldots$$

Следовательно, из (3) вытекает $P_{s_i}^{(\alpha_i)}(t, \hat{z}_0) \equiv P_{s_i}^{(\alpha_i)}(t, z_0)$, $i = 1, 2$. Но тогда из (2.15), (2.16) имеем

$$T\hat{z}_0 = \pi_s(A)\hat{z}_0 = \rho_s \sum_{i \in I} \eta_i e_i = \rho_s z_0.$$

Отсюда и из (2.32) следует

$$T^{2k-1} z_0 = \rho_s^{2k-1} \hat{z}_0, \qquad T^{2k} z_0 = \rho_s^{2k} z_0, \qquad k = 1, 2, \ldots .$$

Из этого выражения, учитывая, что $\|\hat{z}_0\|_{A^{\lambda-1}} = \|z_0\|_{A^{\lambda-1}}$, получаем утверждение леммы. Лемма доказана.

Рассмотрим осредненные скорости сходимости двухступенчатого метода R_k, R_∞, определенные в (2.2).

Теорема 2. *Если $n > s$, то*

$$R_{2k} \leqslant R_\infty \leqslant -s^{-1} \ln \rho_s \qquad \forall \lambda \in R, \quad k = 1, 2, \ldots .$$

Доказательство. Из леммы 5 следует, что $\|T^k\|_\lambda \geqslant \rho_s^k$ для всех $\lambda \in R$, поэтому $R_{2k} \leqslant -s^{-1} \ln \rho_s$, $k = 1, 2, \ldots$. Но тогда из (2.6) имеем

$$R_{2k} \leqslant R_\infty = \lim_{k \to \infty} R_{2k} \leqslant -s^{-1} \ln \rho_s \qquad \forall \lambda \in R, \quad k = 1, 2, \ldots .$$

Теорема доказана.

Замечание 5. Оценки для R_{2k-1} следуют из (2.4), (2.5).

Литература

1. *Ермаков В. В., Калиткин Н. Н.* Двухступенчатый градиентный спуск // Ж. вычисл. матем. и матем. физ. 1980. Т. 20. № 4. С. 1040–1045.
2. *Воскобойников С. П., Сениченков Ю. Б., Цукерман И. А.* К вопросу о скорости сходимости метода двухступенчатого градиентного спуска // Ж. вычисл. матем. и матем. физ. 1983. Т. 23. № 5. С. 1227–1229.
3. *Марчук Г. И., Кузнецов Ю. А.* Итерационные методы и квадратичные функционалы // Методы вычисл. матем. Новосибирск: Наука, 1975. С. 4–143.
4. *Красносельский М. А., Вайникко Г. М., Забрейко П. П. и др.* Приближенное решение операторных уравнений. М.: Наука, 1969.
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. *Бирман М. Ш.* Некоторые оценки для метода наискорейшего спуска // Успехи матем. наук. 1950. Т. 5. Вып. 3. С. 152–155.
7. *Жук П. Ф., Бондаренко Л. Н.* Об одной гипотезе Дж. Форсайта // Матем. сб. 1983. Т. 121(163). № 4(8). С. 435–453.
8. *Воробьев Ю. В.* Метод моментов в прикладной математике. М.: Физматгиз, 1958.
9. *Полиа Г., Сегё Г.* Задачи и теоремы из анализа. М.: Наука, 1978. Ч. 1.
10. *Березин И. С., Жидков Н. П.* Методы вычислений. М.: Физматгиз, 1962. Т. 1.

Поступила в редакцию 20.III.1986
Переработанный вариант 7.IX.1987