

О процессах доортогонализации некоторых семейств векторов, возникающих при построении характеристических полиномов матриц и используемых при решении систем линейных алгебраических уравнений. 2

© О. А. Черная, А. И. Якимчик, 2005

Институт геофизики НАН Украины, Киев, Украина
Поступила 8 февраля 2005 г.
Представлено членом редколлегии В. И. Старостенко

Охарактеризовано сфери застосування методів, розглянутих в першій частині статті, які використовують у визначенні власних векторів і під час відшукування розв'язків систем лінійних рівнянь. Дослідження запропонованих методів відрізняється від традиційного і містить побудову ортонормованих або A -ортонормованих послідовностей векторів. Указані оптимальні способи вибору початкових векторів сукупностей, що будуються. Знайдено стійкі процедури доортогоналізації ортонормованої, A -ортонормованої або біортонормованої систем векторів, що дає змогу визначити з високою точністю як самі вектори конструйованих множин, так і коефіцієнти характеристичного полінома.

The areas of applying the methods presented in the first part of this work and used while proper vectors are being determined and solutions of the system of linear equations are being looked for have been characterized. The study of proposed methods differs from conventional one and involves constructing orthonormalized or A -orthonormalized vector sequences. Optimal ways for the choice of initial vectors of totalities being constructed have been shown. Stable procedures of diorthogonalization of orthonormalized, A -orthonormalized or biorthonormalized vector systems have been found, which allow to determine with high precision both vectors of multitudes being constructed and coefficients of characteristic polynome.

Предлагаемую вниманию работу следует рассматривать как вторую часть статьи [1], которая состоит из введения и четырех пунктов. Для удобства и краткости изложения в каждой части статьи, включая и данную, нумерация формул и теорем в пределах каждого пункта самостоятельная; ссылки нумеруются двумя числами. Например, теорема 8.2, формулы (1.7), (6.5), (7.4), (8.11) и т.д. Причем, первое число обозначает номер пункта, а второе — порядковый номер формулы в пределах того или иного пункта.

5. Определение собственных векторов [7]. Рассмотрим сначала построение собственных векторов по данным методов ортогонализации последовательных итераций и минимальных итераций К.Ланцоша. Будем считать, что корни характеристического полинома $\varphi_n(\lambda) = 0$ найдены (каким-нибудь из методов [2, 5]) и составляют совокупность $\{\lambda_i\}$, $i = 1, 2, \dots, n$. Для каждого из собственных векторов $\mathbf{u}^{(i)}$, отвечающих собственным значениям λ_i , можно записать представление его в ортонормированном базисе $\{\mathbf{u}_i\}$ в виде

$$\mathbf{u}^{(i)} = \sum_{k=1}^n c_{ik} \mathbf{u}_k, \quad i = 1, 2, \dots, n. \quad (5.1)$$

Компоненты $c_{ik} = (\mathbf{u}^{(i)}, \mathbf{u}_k)$ вектора $\mathbf{u}^{(i)}$ в базисе $\{\mathbf{u}_i\}$ можно найти следующим образом. Запишем разложение начального вектора \mathbf{u}_1 ортонормированного базиса по собственным векторам $\mathbf{u}^{(i)}$, $i = 1, 2, \dots, n$. Будем иметь

$$\mathbf{u}_1 = \sum_{i=1}^n \alpha_i \mathbf{u}^{(i)}.$$

Вспомним теперь формулы (1.7 А) и (2.8 А) из [1] и для всех последующих векторов \mathbf{u}_k , $k = 2, 3, \dots, n$, базиса найдем

$$\mathbf{u}_k = \frac{1}{\beta_{k-1k}} \varphi_{k-1}(A) \mathbf{u}_1 = \sum_{i=1}^n \frac{\alpha_i}{\beta_{k-1k}} \varphi_{k-1}(\lambda_i) \mathbf{u}^{(i)},$$

откуда сразу же получаем

$$c_{ij} = (\mathbf{u}^{(i)}, \mathbf{u}_j) = \frac{\alpha_i}{\beta_{j-1j}} \varphi_{j-1}(\lambda_i). \quad (5.2)$$

Попутно заметим, что значения полиномов $\varphi_{j-1}(\lambda_i)$ можно вычислять непосредственно по соответствующим рекуррентным формулам (1.8) и (2.10) из работы [1].

Таким образом, подставляя значения (5.2) коэффициентов c_{ij} в представление (5.1), получаем с точностью до некоторого постоянного множителя α_i (т.е. с той же точностью, с какой можно получить их, решая систему (2))

$$\mathbf{u}^{(i)} = \sum_{k=1}^n \frac{\varphi_{k-1}(\lambda_i)}{\beta_{k-1k}} \mathbf{u}_k, \quad i = 1, 2, \dots, n. \quad (5.3)$$

В случае нормального течения процесса вычислений по методу A -минимальных итераций собственные векторы $\mathbf{u}^{(i)}$ положительно определенной матрицы определяются так же в соответствии с приведенной схемой с тем только отличием, что в качестве базиса $\mathcal{R}^{(n)}$ выбирается совокупность $\{\mathbf{q}_i\}$ A -ортонормированных векторов, удовлетворяющих условиям $(\mathbf{q}_i, A\mathbf{q}_j) = \delta_{ij}$, $i, j = 1, 2, \dots, n$.

Отсюда с точностью до некоторой постоянной получаем

$$\mathbf{u}^{(i)} = \sum_{k=1}^n \frac{\varphi_{k-1}(\lambda_i)}{\beta_{k-1k}} \mathbf{q}_k, \quad i = 1, 2, \dots, n. \quad (5.4)$$

По аналогии с приведенными вычислениями так же просто определяются собственные векторы $\mathbf{u}^{(i)}$ и $\mathbf{v}^{(i)}$ несимметрических матриц A и A^* по данным метода биортогонализации, а именно:

$$\mathbf{u}^{(i)} = \sum_{k=1}^n \frac{\varphi_{k-1}(\lambda_i)}{(\mathbf{u}_k, \mathbf{v}_k)} \mathbf{u}_k, \quad \mathbf{v}^{(i)} = \sum_{k=1}^n \frac{\varphi_{k-1}(\lambda_i)}{(\mathbf{u}_k, \mathbf{v}_k)} \mathbf{v}_k, \quad i = 1, 2, \dots, n. \quad (5.5)$$

6. Решение систем линейных уравнений. Наличие данных, которые получаются с помощью методов, описанных в предыдущих пунктах (см. первую часть [1]), позволяет находить также и решения систем линейных уравнений

$$A\mathbf{x} = \mathbf{b},$$

$$A: \mathcal{R}^{(n)} \rightarrow \mathcal{R}^{(n)}; \quad \mathbf{x}, \mathbf{b} \in \mathcal{R}^{(n)}. \quad (6.1)$$

Действительно, осуществим преобразование

$$\xi = U^* \mathbf{x}, \quad (6.2)$$

где $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ — ортогональная матрица, составленная из векторов-столбцов \mathbf{u}_i , $i = 1, 2, \dots, n$, полученных реализацией либо метода ортогональных итераций (п. 1), либо метода минимальных итераций (п. 2). Умножим уравнение (6.1) слева на матрицу U^* и с учетом равенств (1.10) или (2.9), а также подстановки (6.2) получим

$$B\xi = U^* \mathbf{b}. \quad (6.3)$$

Ясно, что система (6.3) проще исходной (6.1). Особенно эта простота проявляется в случае симметрической матрицы A . Тогда, как мы помним, матрица B будет трехдиагональной. В последнем случае систему (6.1) можно решить с помощью одного изящного приема [7], который распространяется на системы с несимметрической матрицей. Описание такого обобщения не встречается в работах [2 — 7]. Тем не менее, оно может оказаться важным для приложений. Поэтому приведем здесь этот результат.

Теорема 6.1. Если начальные векторы \mathbf{u}_1 и $\mathbf{v}_1 = A\mathbf{v}_0$, $\mathbf{u}_1, \mathbf{v}_1, \mathbf{v}_0 \in \mathcal{R}^{(n)}$, и структура невырожденной несимметрической матрицы $A: \mathcal{R}^{(n)} \rightarrow \mathcal{R}^{(n)}$ таковы, что обеспечивают нормальное течение процесса вычислений по рекуррентным формулам:

$$\mathbf{u}_k = A\mathbf{u}_{k-1} - \beta_{k-1,k-1}\mathbf{u}_{k-1} - \beta_{k-1,k-2}\mathbf{u}_{k-2},$$

$$\mathbf{v}_k = A^* \mathbf{v}_{k-1} - \beta_{k-1,k-1}\mathbf{v}_{k-1} - \beta_{k-1,k-2}\mathbf{v}_{k-2},$$

$$k = 2, 3, \dots, n,$$

$$\mathbf{u}_2 = A\mathbf{u}_1 - \beta_{11}\mathbf{u}_1, \quad \mathbf{v}_2 = A^* \mathbf{v}_1 - \beta_{11}\mathbf{v}_1$$

векторов двойственного базиса $\{\mathbf{u}_i\}$, $\{\mathbf{v}_i\}$ в $\mathcal{R}^{(n)}$, где

$$\beta_{kk} = \frac{(A\mathbf{u}_k, \mathbf{v}_k)}{(\mathbf{u}_k, \mathbf{v}_k)}, \quad k = 1, 2, \dots, n, \quad \beta_{k,k-1} = \frac{(\mathbf{u}_k, \mathbf{v}_k)}{(\mathbf{u}_{k-1}, \mathbf{v}_{k-1})}, \quad k = 2, 3, \dots, n,$$

то решение системы линейных уравнений $A\mathbf{x} = \mathbf{b}$, которое будем отыскивать в виде

$$\mathbf{x} = \sum_{i=1}^n \xi_{0i} \mathbf{u}_i,$$

может быть получено за n последовательных шагов приближений по схеме

$$\mathbf{x}^{(1)} = \frac{(\mathbf{b}, \mathbf{v}_0)}{(\mathbf{u}_1, \mathbf{v}_1)} \mathbf{u}_1, \quad \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \frac{(\mathbf{r}_k, \mathbf{v}_{k-1})}{(\mathbf{u}_k, \mathbf{v}_k)} \mathbf{u}_k,$$

$$\mathbf{r}_k = \mathbf{b} - A\mathbf{x}^{(k)}, \quad \mathbf{x}^{(k)} = \sum_{i=1}^{k-1} \xi_{ki} \mathbf{u}_i, \quad k = 2, 3, \dots, n.$$

Доказательство теоремы вытекает из следующих непротиворечивых рассуждений. Поскольку справедливы выражения

$$(\mathbf{x}, \mathbf{v}_k) = \sum_{i=1}^n \xi_{0i} (\mathbf{u}_i, \mathbf{v}_k) = \xi_{0k} (\mathbf{u}_k, \mathbf{v}_k), \quad k = 1, 2, \dots, n,$$

$$(\mathbf{x}^{(k)}, \mathbf{v}_k) = \sum_{i=1}^{k-1} \xi_{ki} (\mathbf{u}_i, \mathbf{v}_k) \equiv 0, \quad k = 2, 3, \dots, n,$$

то вектор $\mathbf{x} - \mathbf{x}^{(k)}$ удовлетворяет условиям

$$\mathbf{x} - \mathbf{x}^{(k)} \in \mathcal{L}(\mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n) \perp \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}).$$

Поэтому для k -й компоненты вектора-решения \mathbf{x} в базисе $\{\mathbf{u}_i\}$ можно записать

$$\begin{aligned}\xi_k(\mathbf{u}_k, \mathbf{v}_k) &= (\mathbf{x} - \mathbf{x}^{(k)}, \mathbf{v}_k) = (\mathbf{x} - \mathbf{x}^{(k)}, A^* \mathbf{v}_{k-1} - \beta_{k-1k-1} \mathbf{v}_{k-1} - \beta_{k-1k-2} \mathbf{v}_{k-2}) = \\ &= (\mathbf{x} - \mathbf{x}^{(k)}, A^* \mathbf{v}_{k-1}) = (\mathbf{r}_k, \mathbf{v}_{k-1}), \quad k = 2, 3, \dots, n,\end{aligned}$$

где $\mathbf{r}_k = A\mathbf{x} - A\mathbf{x}^{(k)} = \mathbf{b} - A\mathbf{x}^{(k)}$. Отсюда получаем, что

$$\xi_k = \frac{(\mathbf{r}_k, \mathbf{v}_{k-1})}{(\mathbf{u}_k, \mathbf{v}_k)}, \quad k = 2, 3, \dots, n.$$

Следовательно,

$$\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} = \xi_k \mathbf{u}_k = \frac{(\mathbf{r}_k, \mathbf{v}_{k-1})}{(\mathbf{u}_k, \mathbf{v}_k)} \mathbf{u}_k, \quad (6.4)$$

что с учетом определения первого приближения $\mathbf{x}^{(1)}$ в соответствии с выбором начального вектора $\mathbf{v}_1 = A\mathbf{v}_0$ полностью доказывает теорему.

Заметим, что невязку вектора $\mathbf{x}^{(k)}$ можно вычислять рекуррентно по формуле

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \frac{(\mathbf{r}_k, \mathbf{v}_{k-1})}{(\mathbf{u}_k, \mathbf{v}_k)} A\mathbf{u}_k,$$

которая выводится элементарно со ссылкой на (6.4):

$$\mathbf{r}_{k+1} = \mathbf{b} - A\mathbf{x}^{(k+1)} = \mathbf{b} - A\mathbf{x}^{(k)} - A(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = \mathbf{r}_k - \frac{(\mathbf{r}_k, \mathbf{v}_{k-1})}{(\mathbf{u}_k, \mathbf{v}_k)} A\mathbf{u}_k, \quad k = 2, 3, \dots, n+1.$$

В заключение этого параграфа отметим, что при наличии данных метода A -минимальных итераций (п. 3) система линейных уравнений

$$A\mathbf{x} = \mathbf{b} \quad (6.5)$$

(с положительно определенной матрицей) решается проще, чем в предыдущих случаях. В самом деле, при реализации метода A -минимальных итераций получим матрицу $Q = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n]$, составленную из A -ортонормированных векторов — столбцов $\mathbf{q}_k = (q_{1k}, q_{2k}, \dots, q_{nk})^*$ и поэтому $Q^*AQ = E^{(n)}$.

Отсюда, если положить $\mathbf{x} = Q\xi$, $\xi \in \mathcal{R}^{(n)}$, и умножить систему (6.5) слева на матрицу Q^* , то немедленно получим $Q^*AQ\xi = \xi = Q^*\mathbf{b}$. Следовательно, решение системы (6.5) определяется формулой $\mathbf{x} = QQ^*\mathbf{b}$, что, между прочим, дает для $A^{-1} = QQ^*$.

7. Выбор начальных векторов и построение базиса в $\mathcal{R}^{(n)}$. Рассматривая методы косвенного определения коэффициентов характеристического многочлена, аннулирующего матрицу (см. пункты 1 — 4 статьи [1]), каждый раз подчеркивалось, что нормальное течение процесса вычислений зависит не только от структуры исходной матрицы, но и от удачного выбора начального вектора (или пары векторов, если речь идет о методе биортогонализации). Для того чтобы удача всегда сопутствовала вычислениям по схеме того или иного метода, выясним возможные причины прекращения вычислений и наметим пути выхода из затруднительных ситуаций. Впрочем, основная причина прекращения вычислений ясна. Это — линейная зависимость системы векторов, полученной итерациями матрицей начального вектора. В свою очередь для обнаружения связи характеристик этой системы векторов со структурой матрицы и с выбором исходного вектора используем специальные базисы в $\mathcal{R}^{(n)}$.

Сконцентрируем наше внимание прежде всего на алгоритмах нахождения характеристического полинома для симметрической матрицы. Эти алгоритмы проще других в том отношении, что, как известно из линейной алгебры, характеристические многочлены для симметрических матриц всегда разрешимы в поле вещественных чисел. А это означает, что для всякой симметрической матрицы существует (над полем вещественных чисел) хотя бы один собственный вектор. Впрочем, количество этих векторов не меньше, чем количество простых собственных чисел. Это вытекает из определения (2) собственных векторов.

Пусть $\mathbf{u}^{(i)} \in \mathcal{R}^{(n)}$ — собственные векторы симметрической матрицы A , отвечающие простым собственным значениям λ_i , $i = 1, 2, 3, \dots, k \leq n$. Разложим начальный вектор $\mathbf{p}^{(1)}$ (в методе минимальных итераций $\mathbf{p}^{(1)} = \mathbf{u}_1$ и в методе A -минимальных итераций $\mathbf{p}^{(1)} = \mathbf{q}_1$) по базису $\{\mathbf{u}^{(i)}\}$ из собственных векторов матрицы A :

$$\mathbf{p}^{(1)} = \sum_{i=1}^k p_{1i} \mathbf{u}^{(i)} + \tilde{\mathbf{p}}^{(1)}, \quad (7.1)$$

где $\tilde{\mathbf{p}}^{(1)}$ — $(n-k)$ -мерный вектор, принадлежащий дополнению $\mathcal{R}^{(n)} \setminus \mathcal{L}(\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(k)})$ линейной оболочки $\mathcal{L}(\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(k)})$ до всего пространства $\mathcal{R}^{(n)}$. Заметим, что если среди собственных чисел λ_i содержатся кратные, то $k < n$. Для итераций вектора $\mathbf{p}^{(1)}$ матрицей A получаем

$$A^m \mathbf{p}^{(1)} = \sum_{i=1}^{k_0} p_{1i} \lambda_i^m \mathbf{u}^{(i)}, \quad k_0 \leq k \leq n, \quad (7.2)$$

где $k_0 < k$, если среди собственных значений существуют равные нулю. Представления (7.1) и (7.2) выписаны для самого благоприятного выбора вектора, когда $p_{1i} \neq 0$, $i = 1, 2, 3, \dots, k \leq n$. Однако может случиться, что неудачу выбранный вектор $\mathbf{p}^{(1)}$ принадлежит $\mathcal{L}(\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(l)})$, $l < k$, и тогда получим

$$\mathbf{p}^{(1)} = \sum_{i=1}^l p_{1i} \mathbf{u}^{(i)}, \quad A^m \mathbf{p}^{(1)} = \sum_{i=1}^{k_1} p_{1i} \lambda_i^m \mathbf{u}^{(i)}, \quad k_1 = \min(l, k_0).$$

Таким образом, итерации

$$\mathbf{p}^{(1)}, A\mathbf{p}^{(1)}, \dots, A^{n-1}\mathbf{p}^{(1)} \quad (7.3)$$

вектора $\mathbf{p}^{(1)}$ симметрической матрицей A будут линейно независимыми тогда и только тогда, когда среди чисел λ_i , $i = 1, 2, \dots, n$ нет кратных ($k \equiv n$) и равных нулю ($k_0 \equiv k \equiv n$), а вектор $\mathbf{p}^{(1)}$ выбран так, что ни одна из компонент p_{1i} , $i = 1, 2, \dots, n$ не равна нулю. Если же $\mathbf{p}^{(1)} \in \mathcal{L}(\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(l)})$, $l < n$, то в последовательности (7.3) независимыми окажутся только векторы $A^m \mathbf{p}^{(1)}$, $m = 0, 1, \dots, l-1$, тогда как векторы $A^m \mathbf{p}^{(1)}$, $m = l, l+1, \dots, n-1$, будут линейно от них зависеть. Это случится уже для $l=1$, если $\mathbf{p}^{(1)} = \mathbf{u}^{(1)}$ или для $k=2$, если $\mathbf{p}^{(1)} = p_{11}\mathbf{u}^{(1)} + p_{12}\mathbf{u}^{(2)}$ и т.п. Этими рассуждениями еще не найден критерий для благоприятного выбора начального вектора $\mathbf{p}^{(1)}$, но по крайней мере уже ясно, от чего зависит прекращение вычислений. Это может быть связано либо с наличием в спектре¹ матрицы кратных собственных чисел или чисел, равных нулю, либо с тем, что начальный вектор принадлежит линейной оболочке собственных векторов матрицы, не совпадающей со всем пространством.

Вернемся к равенствам (7.2) и отметим, что при наличии в спектре матрицы одного или нескольких собственных чисел, превосходящих по модулю остальные, слагаемые $p_{1i} \lambda_i^m \mathbf{u}^{(i)}$,

¹ Спектром матрицы называют дискретное множество $\{\lambda_i\}$ корней ее характеристического полинома.

содержащие эти большие по модулю числа, окажутся преобладающими. Отсюда выводим следующее правило выбора начального вектора $\mathbf{p}^{(1)}$. Возьмем произвольный вектор, например — правую часть системы уравнений, если таковая задана, или какую-нибудь строку $\mathbf{a}_{(k)} = (a_{k1}, a_{k2}, \dots, a_{kn})$ заданной матрицы, и последовательно вычислим

$$\mathbf{p}_{(1)} = A\mathbf{a}_{(k)}, \quad \mathbf{p}_{(2)} = A\mathbf{p}_{(1)}, \dots, \mathbf{p}_{(m)} = A\mathbf{p}_{(m-1)}.$$

Если $m = 5 + 6$, то наибольшие по модулю собственные значения уже будут играть основную роль и вектор $\mathbf{p}_{(m)}$ будет близким к линейной оболочке, натянутой на собственные векторы, отвечающие большим по модулю собственным значениям. Пусть среди компонент p_{mi} , $i = 1, 2, \dots, n$, вектора $\mathbf{p}_{(m)}$ компоненты p_{mi_k} , $i = 1, 2, 3, \dots, l \leq n$, оказались «малыми» по модулю. Изменим их произвольным образом, прибавив к ним некоторые числа γ_{ik} . В результате полу-

$$\text{чим } \tilde{\mathbf{p}}_{(m)} = (\tilde{p}_{m1}, \dots, \tilde{p}_{mn})^*, \quad \tilde{p}_{mi} = \begin{cases} p_{mi}, & i = i_k \\ p_{mi_k} + \gamma_{ik}, & i = i_k, k = 1, 2, \dots, l. \end{cases}$$

Далее, вычислим линейную комбинацию

$$\tilde{\mathbf{p}}^{(1)} = \sum_{i=1}^{m-1} \gamma_i \mathbf{p}_{(i)} + \tilde{\mathbf{p}}_{(m)}$$

при заданных γ_i , $i = 1, 2, \dots, m-1$, и в качестве начального вектора $\mathbf{p}^{(1)}$ примем соответствующим образом пронормированный вектор $\tilde{\mathbf{p}}^{(1)}$.

Однако такой выбор начального вектора вовсе не гарантирует того, что процесс вычислений в любом из методов итераций (см. п. 2 — 3 из [1]) завершится нормально. Теперь его протекание будет определяться структурой матрицы. Ниже для краткости поведем речь только о процедурах, регулирующих процесс вычисления в методе минимальных итераций. Если вычисления по этому методу завершаются досрочно на k -м шаге при $k < n$, то это означает, что $\varphi_k(\lambda)$ будет минимальным аннулирующим вектор $\mathbf{p}^{(1)} = \mathbf{u}_1$ полиномом. Следуя предложениям [7], покажем, как можно достроить систему векторов $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$ до ортонормированного базиса всего пространства $\mathcal{R}^{(n)}$, если вектор \mathbf{u}_k оказался линейно зависимым от векторов оболочки $\mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1})$, т.е. если на k -м шаге получим

$$0 = -\sum_{i=1}^{k-1} \beta_{k-1,i} \mathbf{u}_i + A\mathbf{u}_{k-1}, \quad \beta_{k-1,k} = 0, \quad k < n.$$

Продолжать вычисления будем по следующей схеме.

Выберем вектор $\tilde{\mathbf{p}}^{(2)}$, не принадлежащий $\mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1})$, в виде

$$\tilde{\mathbf{p}}^{(2)} = \sum_{i=1}^{k-1} \gamma_i \mathbf{u}_i + \gamma_k \mathbf{e}_k, \quad \mathbf{p}^{(2)} = \left\| \tilde{\mathbf{p}}^{(2)} \right\|^{-1} \tilde{\mathbf{p}}^{(2)},$$

где γ_i , $i = 1, 2, \dots, k$, — заданные числа, а \mathbf{e}_k — единичный вектор с компонентами, равными нулю, за исключением k -й, равной единице. Разложим вектор $\mathbf{p}^{(2)}$ на две составляющие:

$$\mathbf{p}_1^{(2)} \perp \mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}) \quad \text{и} \quad \mathbf{p}_2^{(2)} \in \mathcal{L}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}),$$

а в качестве \mathbf{u}_k строящейся последовательности берем вектор

$$\mathbf{u}_k = \frac{\mathbf{p}_1^{(2)}}{\|\mathbf{p}_1^{(2)}\|}$$

Проекцию $\mathbf{p}_1^{(2)}$ отыскиваем тем же способом ортогонализации, который применялся для построения всех предыдущих векторов $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$. Следующий $(k+1)$ -й вектор вычисляется точно так же, как определялся вектор \mathbf{u}_2 , т. е. находится сначала

$$\mathbf{u}_{k+1}' = -\beta_{kk} \mathbf{u}_k + A\mathbf{u}_k$$

и т. д., и т. п.

Итерации вектора \mathbf{u}_k матрицей A могут оказаться снова линейно зависимыми на некотором l -м шаге вычислений. Если $l < n$, то придется повторить описанную процедуру, начиная с выбора нового вектора $\tilde{\mathbf{p}}^{(3)}$, и т. д. Для того чтобы завершить вычисления, возможно придется построить несколько линейных оболочек:

$$\mathcal{L}_{p^{(1)}}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}), \mathcal{L}_{p^{(2)}}(\mathbf{u}_k, \dots, \mathbf{u}_{l-1}), \dots, \mathcal{L}_{p^{(m)}}(\mathbf{u}_l, \dots, \mathbf{u}_n),$$

ортогональных друг другу. При этом все пространство $\mathcal{R}^{(n)}$ развивается естественным образом в прямую сумму этих оболочек:

$$\mathcal{R}^{(n)} = \mathcal{L}_{p^{(1)}}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}) \oplus \mathcal{L}_{p^{(2)}}(\mathbf{u}_k, \dots, \mathbf{u}_{l-1}) \oplus \dots \oplus \mathcal{L}_{p^{(m)}}(\mathbf{u}_l, \dots, \mathbf{u}_n). \quad (7.4)$$

А так как каждая из этих оболочек получена на основании итераций некоторого начального вектора $\tilde{\mathbf{p}}^{(i)}$, $i = 1, 2, \dots, m \leq n$, аннулируемого соответствующим полиномом $\varphi^{(i)}(t)$, то характеристический полином матрицы A с точностью до некоторой постоянной совпадает с произведением минимальных полиномов $\varphi^{(i)}(t)$, аннулирующих последовательные «начальные» векторы $\tilde{\mathbf{p}}^{(i)}$, $i = 1, 2, \dots, m$, т. е.

$$\varphi_n(t) = \prod_{i=1}^m \varphi^{(i)}(t),$$

где каждый из полиномов $\varphi^{(i)}(t)$, соответствующий вектору $\tilde{\mathbf{p}}^{(i)}$, находится по схеме (1.8). Выпишем для обозрения матрицу B перехода от системы векторов $\{\mathbf{u}_i\}$ к системе $\{A^m \mathbf{u}_i\}$, отвечающую случаю

$$\mathcal{R}^{(n)} = \mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}) \oplus \mathcal{L}(\mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n).$$

Получим

$$B^* = \begin{vmatrix} \beta_{11} & \beta_{12} & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \beta_{21} & \beta_{22} & \beta_{23} & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \beta_{k-2,k-3} & \beta_{k-2,k-2} & \beta_{k-2,k-1} & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & \beta_{k-1,k-2} & \beta_{k-1,k-1} & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \beta_{kk} & \beta_{k,k+1} & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \beta_{k+1,k} & \beta_{k+1,k+1} & \beta_{k+1,k+2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \beta_{n-1,n-2} & \beta_{n-1,n-1} & \beta_{n-1,n} \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 0 & \beta_{n,n-1} & \beta_{n,n} \end{vmatrix}.$$

Как видим, в описанном случае матрица B состоит из двух трехдиагональных клеток, каждая из которых совпадает по своей структуре с матрицей преобразования (2.8) из [1]. Очевидно, в общем случае, отвечающему разбиению (7.4), матрица B будет состоять из m трехдиагональных клеток.

Построение базиса $\{\mathbf{q}_i\}$ в $\mathcal{R}^{(n)}$ и отыскание коэффициентов характеристического многочлена $\varphi_n(\lambda) = 0$ с помощью метода A -минимальных итераций осуществляется почти в соответствии с описанной выше схемой за исключением одного обстоятельства, когда вместо ортонормированной строится A -ортонормированная совокупность векторов.

Описанный способ выбора начального вектора и схема достройки базиса в $\mathcal{R}^{(n)}$ легко обобщаются на случай несимметрической матрицы. Несколько сложнее дело обстоит с обоснованием этого обобщения, поскольку не каждый характеристический полином (с вещественными коэффициентами) несимметрической матрицы разрешим в поле вещественных чисел и, в связи с этим, не каждая несимметрическая матрица имеет хотя бы один собственный вектор над этим полем. Из-за недостатка места и времени не будем касаться здесь этого вопроса, а только заметим следующее. Процесс вычислений в методе ортогонализации последовательных итераций (п. 1), если он прерывается досрочно на k -м шаге, $k < n$, можно продолжить, выбирая в качестве вектора \mathbf{u}_k вновь подобранный вектор $\mathbf{p}^{(i)}$, $i = 1, 2, \dots, m \leq n$, по описанной выше схеме без каких бы то ни было изменений.

Что же касается метода биортогонализации (п. 4), то при односторонних или тупиковых обрывах вычисления приходится начинать сначала с вновь выбранными исходными векторами \mathbf{u}_1 и \mathbf{v}_1 . При двустороннем обрыве полученная последовательность векторов взаимного базиса

$$\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}; \quad \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1} \quad (7.5)$$

достраивается соответствующим образом. При выборе пары начальных векторов $\mathbf{u}_1, \mathbf{v}_1$, как и при выборе начального вектора в методах минимальных итераций, предварительно можно вычислить последовательности

$$\begin{aligned} \mathbf{u}_{(2)} &= A\mathbf{u}_{(1)}, \quad \mathbf{u}_{(3)} = A\mathbf{u}_{(2)}, \quad \dots, \quad \mathbf{u}_{(m)} = A\mathbf{u}_{(m-1)}, \\ \mathbf{v}_{(2)} &= A^*\mathbf{v}_{(1)}, \quad \mathbf{v}_{(3)} = A^*\mathbf{v}_{(2)}, \quad \dots, \quad \mathbf{v}_{(m)} = A^*\mathbf{v}_{(m-1)}. \end{aligned}$$

где в качестве $\mathbf{u}_{(1)}$ и $\mathbf{v}_{(1)}$ можно, например, взять соответственно какую-либо строку $\mathbf{a}^{(k)}$ и столбец $\mathbf{a}^{(k)}$ матрицы A и на этой основе определить \mathbf{u}_1 и \mathbf{v}_1 . Достройка системы (7.5) начинается с выбора таких векторов $\mathbf{p}^{(i)}$ и $\mathbf{q}^{(i)}$, $i = 1, 2, \dots, m$, $m \leq n$, которые удовлетворяют условиям

$$\begin{aligned} \mathbf{p}^{(i)} &\notin \mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}) \quad \text{и} \quad \mathbf{p}^{(i)} \perp \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}), \\ \mathbf{q}^{(i)} &\notin \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1}) \quad \text{и} \quad \mathbf{q}^{(i)} \perp \mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}). \end{aligned} \quad (7.6)$$

Если на k -м шаге имеет место двусторонний обрыв, а векторы $\mathbf{p}^{(1)}$ и $\mathbf{q}^{(1)}$ подобраны в соответствии с указанными требованиями, то полагаем $\mathbf{u}_k = \mathbf{p}^{(1)}$, $\mathbf{v}_k = \mathbf{q}^{(1)}$. Процесс постройки последовательности (7.5) продолжаем до следующего l -го обрыва. Он может оказаться последним, если $l = n$. При $l < n$ вычисления продолжают с новыми векторами $\mathbf{u}_l = \mathbf{p}^{(2)}$, $\mathbf{v}_l = \mathbf{q}^{(2)}$, удовлетворяющими условиям (7.6) при $k = l$, и т. д. и т. п.

8. Процессы доортогонализации при численной реализации канонических представлений (1.10), (2.9), (3.5) и (4.6). В предыдущих пунктах описаны некоторые методы косвенного определения коэффициентов характеристических полиномов симметрических и несимметрических

ких вещественных матриц, а также основанные на этих методах способы вычисления собственных векторов (при заданных собственных числах) и решений систем линейных уравнений с такими матрицами. Как было показано, для определения коэффициентов аннулирующей матрицу полинома требовалось выполнить определенный объем вычислений. Так, в методе ортогонализации последовательных итераций приходилось вычислять линейные комбинации (1.3) — (1.4) вида

$$\mathbf{u}'_{k+1} = A\mathbf{u}_k - \sum_{i=1}^k \beta_{ki} \mathbf{u}_i, \quad \beta_{ki} = (A\mathbf{u}_k, \mathbf{u}_i), \quad \beta_{k,k+1} = \|\mathbf{u}_{k+1}\|, \quad (8.1)$$

$$\mathbf{u}_{k+1} = \beta_{k,k+1}^{-1} \mathbf{u}'_{k+1}, \quad i = 1, 2, \dots, k; \quad k = 1, 2, \dots, n-1.$$

В методе минимальных итераций (п. 2) проводились аналогичные вычисления (см. (2.1), (2.6), (2.7) статьи [1]), только в меньшем объеме. Более сложные вычисления (см. (3.8) — (3.10)) выполнялись в методе A -минимальных итераций по формулам

$$\mathbf{q}'_{k+1} = -b_{k,k-1} \mathbf{q}_{k-1} - b_{kk} \mathbf{q}_k + A\mathbf{q}_k,$$

$$b_{k,k-1} = (A\mathbf{q}_k, \mathbf{q}_{k-1}), \quad k = 2, 3, \dots, n, \quad b_{kk} = (A\mathbf{q}_k, \mathbf{q}_k), \quad k = 1, 2, \dots, n, \quad (8.2)$$

$$b_{k,k+1} = \sqrt{(A\mathbf{q}'_{k+1}, \mathbf{q}'_{k+1})}, \quad \mathbf{q}_{k+1} = b'_{k,k+1} \mathbf{q}'_{k+1}, \quad k = 1, 2, \dots, n-1,$$

и, наконец, в методе биортогонализации (п. 4) — по формулам

$$\mathbf{u}_{k+1} = -b_{k,k-1} \mathbf{u}_{k-1} - b_{kk} \mathbf{u}_k + A\mathbf{u}_k, \quad \mathbf{v}_{k+1} = -b_{k,k-1} \mathbf{v}_{k-1} - b_{kk} \mathbf{v}_k + A^* \mathbf{v}_k, \quad (8.3)$$

$$b_{kk} = \frac{(A\mathbf{u}_k, \mathbf{v}_k)}{(\mathbf{u}_k, \mathbf{v}_k)}, \quad k = 1, 2, \dots, n; \quad b_{k,k-1} = \frac{(\mathbf{u}_k, \mathbf{v}_k)}{(\mathbf{u}_{k-1}, \mathbf{v}_{k-1})}, \quad k = 2, 3, \dots, n.$$

Далее, для отыскания собственных векторов матриц в виде (5.3), (5.4) и (5.5) прежде всего должны быть определены коэффициенты β_{ki} , b_{ki} по соответствующим формулам (8.1) — (8.3), а для получения решения систем линейных уравнений необходимо иметь помимо этих коэффициентов еще и значения векторов \mathbf{u}_i , \mathbf{v}_i , $A\mathbf{u}_i$, $A^* \mathbf{v}_i$, \mathbf{q}_i , $i = 1, 2, \dots, n$, по возможности более точные. Однако в связи с тем, что при реализации на ЭВМ почти каждой арифметической операции неизбежно возникают, а при большом числе операций накапливаются ошибки округления, нет никаких оснований надеяться на получение хороших результатов при реализации схем (8.1) — (8.3). Это подтверждается исследованиями [3 — 8], где также сказано, что для получения с некоторой гарантированной точностью результатов типа скалярных произведений (при умножении матрицы на вектор) необходимо специальным образом организовать весь вычислительный процесс. В частности, чтобы получить минимальную погрешность округления суммы некоторых чисел, необходимо суммировать эти числа, ранжированные по величине их модуля, начиная с наименьшего (по абсолютной величине). В дальнейшем, не оговаривая каждый раз особо, будем широко пользоваться «упорядоченным суммированием». Впрочем, использование одного только «упорядоченного суммирования» не всегда приводит к желаемому результату. Так, например, уже при нормировке вектора $\mathbf{u}^{(k)}$ по схеме

$$\mathbf{u}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}^{(k)}\|}, \quad \|\mathbf{u}^{(k)}\|^2 = (\mathbf{u}^{(k)}, \mathbf{u}^{(k)}) = \sum_{i=1}^n \mathbf{u}_{ki}^2 \quad (8.4)$$

часто не удается достигнуть равенства [3, 8]. Для того чтобы вычислить орт \mathbf{u} в направлении вектора $\tilde{\mathbf{u}} = (\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n)^*$, следует, как это рекомендуется в [8], провести циклические вычисления по схеме (8.4) для $k = 1, 2, \dots, m$, начиная вычисления с вектора $\mathbf{u}^{(1)} = \tilde{\mathbf{u}}$ и заканчивая их получением вектора $\mathbf{u}_m = \mathbf{u}$, для которого имеет место неравенство $|\|\mathbf{u}_m\| - 1| < \varepsilon$, где $\varepsilon > 0$ — машинная точность изображения чисел. На приведенном примере проиллюстрирована основная идея вычислений того или иного результата: для восстановления потерянной при округлениях точности исходных данных необходимо строить некоторый (быстросходящийся) итерационный процесс с учетом неподчинения машинной арифметики обычным законам коммутативности, ассоциативности и дистрибутивности. Идея использования последовательных приближений (с указанием необходимого числа итераций), направленных на восстановление значащих цифр искомого результата, является основным, если не единственно возможным, средством решения поставленной в данной работе задачи. Без ее реализации нельзя определить ни одним из методов, рассмотренных в п. 1 — 4 работы [1], коэффициенты характеристического полинома с высокой точностью.

Первым шагом в реализации данной идеи является построение с помощью соответствующих итераций (донормировки по [8]) единичных или A -единичных векторов, позволяющих более надежно определить как сами векторы \mathbf{u}_k и \mathbf{q}_k , так и коэффициенты β_{ki} и b_{ki} рекуррентных связей между ними в виде (8.1) — (8.3). С самого начала было предусмотрено применение этой процедуры и поэтому приведено описание методов в пунктах 1 — 3 в отличие от традиционного их изложения в работах [2, 5, 7] с учетом нормировки строящейся системы векторов.

Второй шаг состоит в построении итерационного процесса для вычисления ортогональной (A -ортогональной или биортогональной) системы векторов. Описанные в пунктах 1 — 4 статьи [1] процессы ортогонализации являются очень неустойчивыми по отношению к ошибкам округлений процедурами, особенно в условиях плохой обусловленности матриц. В этой связи напомним, что спектр плохо обусловленных² матриц принадлежит довольно широкому интервалу, включающему как очень большие, так и очень малые по модулю значения собственных чисел [8]. А это означает, что при плохой обусловленности матрицы A в последовательности

$$\mathbf{u}_1, A\mathbf{u}_1, \dots, A^k\mathbf{u}_1, \quad k = 1, 2, \dots, n-1,$$

(или в последовательностях $\{A^k\mathbf{q}_1\}$, $\{(A^*)^k\mathbf{v}_1\}$) обязательно присутствуют близкие по норме $\mathcal{R}^{(n)}$ (или $\mathcal{R}_A^{(n)}$) векторы. Последнее, в свою очередь, может привести к тому, что при определении с невысокой точностью k -го ортогонального (A -ортогонального или биортогонального) вектора \mathbf{u}_k линейной оболочки $\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ (или ей соответствующих в других методах)

исходные векторы $A\mathbf{u}_{k-1}$ и $\sum_{i=1}^{k-1} \beta_{ki}\mathbf{u}_i$ (или им соответствующие) для $k = 2, 3, \dots, n+1$, могут оказаться почти линейно зависимыми и, как следствие этого, вектор \mathbf{u}_k не будет ортогональным к линейной оболочке $\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1})$ (или ей соответствующих). Поэтому особое значение приобретают следующие предложения.

Теорема 8.1. Пусть начальный орт $\mathbf{u}_1 \in \mathcal{R}^{(n)}$ и структура матрицы $A: \mathcal{R}^{(n)} \rightarrow \mathcal{R}^{(n)}$ таковы, что последовательность $\mathbf{u}_1, A\mathbf{u}_1, \dots, A^{n-1}\mathbf{u}_1$ состоит из линейно независимых векторов n -мерного евклидова пространства $\mathcal{R}^{(n)}$. Система векторов $\{\mathbf{u}_i\}$ будет практически ортогональной эквивалентной системе векторов линейной оболочки

² Плохо обусловленными называют [6 — 8] матрицы A , для которых мера $c(A) = \|A\| \cdot \|A^{-1}\|$ принимает значения, значительно превышающие единицу.

$$\mathcal{L}(\mathbf{u}_1, A\mathbf{u}_1, \dots, A^k\mathbf{u}_1), \quad k = 0, 1, \dots, n-1,$$

тогда и только тогда, когда каждый вектор \mathbf{u}_k , $k = 1, 2, \dots, n$, будет определяться с помощью следующего процесса доортогонализации:

$$\mathbf{u}_k = \lim_{i \rightarrow \infty} \mathbf{u}_k^{(i)}, \quad \mathbf{u}_k^{(i)} = \tilde{\mathbf{u}}_k^{(i)} \left\| \tilde{\mathbf{u}}_k^{(i)} \right\|_{R^{(n)}}^{-1},$$

$$\tilde{\mathbf{u}}_k^{(i)} = \mathbf{u}_k^{(i-1)} - \sum_{j=1}^{k-1} \beta_{k-1j}^{(i-1)} \mathbf{u}_j, \quad \beta_{k-1j}^{(i-1)} = (\mathbf{u}_k^{(i-1)}, \mathbf{u}_j), \quad j \leq k-1, \quad (8.5)$$

$$\beta_{k-1k}^{(i-1)} = \left\| \tilde{\mathbf{u}}_k^{(i)} \right\| = \left\| \mathbf{u}_k^{(i-1)} - \sum_{j=1}^{k-1} \beta_{k-1j}^{(i-1)} \mathbf{u}_j \right\|, \quad k \leq n,$$

$$\mathbf{u}_k^{(0)} \equiv A\mathbf{u}_k, \quad k = 2, 3, \dots, n,$$

который для получения хорошего приближения $\mathbf{u}_k^{(i)}$ вектора \mathbf{u}_k , удовлетворяющего неравенству

$$\left\| \mathbf{u}_k - \mathbf{u}_k^{(i)} \right\| \leq \varepsilon,$$

достаточно прекратить на l -м шаге, когда будет достигнуто неравенство

$$\left| \sum_{j=1}^{k-1} \beta_{k-1j}^{(l)} u_{jp} \right| \leq \varepsilon_0,$$

где в качестве «порога» $\varepsilon_0 = \frac{\varepsilon}{\sqrt{n}}$ берется число $\varepsilon_0 = 2^{-s} |u_{kp}^{(0)}| = 2^{-s} \min_i |u_{ki}^{(0)}|$, а s — количество значащих цифр мантиссы чисел (в двоичном коде).

Теорема 8.2. Пусть начальный вектор $\mathbf{q}_1 \in \mathcal{R}_A^{(n)}$ выбран так, что совокупность $\mathbf{q}_1, A\mathbf{q}_1, \dots, A^{n-1}\mathbf{q}_1$, полученная итерациями вектора \mathbf{q}_1 положительно определенной матрицей A , состоит из линейно независимых векторов n -мерного евклидова пространства $\mathcal{R}_A^{(n)}$. Система векторов $\{\mathbf{q}_i\}$ будет A -ортогональной и эквивалентной системе векторов оболочки

$$\mathcal{L}(\mathbf{q}_1, A\mathbf{q}_1, \dots, A^k\mathbf{q}_1), \quad k = 0, 1, \dots, n-1,$$

тогда и только тогда, когда каждый вектор \mathbf{q}_k , $k = 2, 3, \dots, n$, будет определяться с помощью следующего процесса A -доортогонализации:

$$\mathbf{q}_k = \lim_{i \rightarrow \infty} \mathbf{q}_k^{(i)}, \quad \mathbf{q}_k^{(i)} = \tilde{\mathbf{q}}_k^{(i)} \left\| \tilde{\mathbf{q}}_k^{(i)} \right\|_{R_A^{(n)}}^{-1}, \quad \tilde{\mathbf{q}}_k^{(i)} = \mathbf{q}_k^{(i-1)} - b_{k-1,k-1}^{(i-1)} \mathbf{q}_{k-1} - b_{k-1,k-2}^{(i-1)} \mathbf{q}_{k-2},$$

$$b_{k-1,k-2}^{(i-1)} = (\mathbf{q}_k^{(i-1)}, A\mathbf{q}_{k-2}), \quad b_{k-1,k-1}^{(i-1)} = (\mathbf{q}_k^{(i-1)}, A\mathbf{q}_{k-1}),$$

$$b_{k-1,k}^{(i-1)} = \left\| \tilde{\mathbf{q}}_k^{(i)} \right\|_{R_A^{(n)}} = \sqrt{(A\tilde{\mathbf{q}}_k^{(i)}, \tilde{\mathbf{q}}_k^{(i)})}, \quad \mathbf{q}_k^{(0)} \equiv A\mathbf{q}_{k-1}, \quad k = 2, 3, \dots, n,$$

который для получения хорошего приближения $\mathbf{q}_k^{(i)}$ вектора \mathbf{q}_k , удовлетворяющего неравенству

$$\|\mathbf{q}_k - \mathbf{q}_k^{(l)}\|_{\mathcal{R}^{(n)}} \leq \varepsilon,$$

достаточно прекратить на l -м шаге при достижении неравенства

$$\left| b_{k-1,k-1}^{(l)} q_{k-1,p} + b_{k-1,k-2}^{(l)} q_{k-2,p} \right| \leq \varepsilon_0,$$

где в качестве «порога» $\varepsilon_0 = \frac{\varepsilon}{\sqrt{n}}$ берется число $\varepsilon_0 = 2^{-s} |q_{kp}^{(0)}| = 2^{-s} \min_i |q_{ki}^{(0)}|$.

Теорема 8.3. Пусть начальные векторы $\mathbf{u}_1 \in \mathcal{R}^{(n)}$, $\mathbf{v}_1 \in \mathcal{R}^{*(n)}$ и структура матрицы $A: \mathcal{R}^{(n)} \rightarrow \mathcal{R}^{(n)}$ таковы, что каждая из последовательностей

$$\mathbf{u}_1, A\mathbf{u}_1, \dots, A^{n-1}\mathbf{u}_1 \text{ и } \mathbf{v}_1, A^*\mathbf{v}_1, \dots, (A^*)^{n-1}\mathbf{v}_1$$

состоят из линейно независимых векторов n -мерных евклидовых пространств $\mathcal{R}^{(n)}$ и $\mathcal{R}^{*(n)}$, где $\mathcal{R}^{*(n)}$ — сопряженное с $\mathcal{R}^{(n)}$ пространство. Совокупности векторов $\{\mathbf{u}_i\}$ и $\{\mathbf{v}_i\}$ будут эквивалентными соответственно линейным оболочкам

$$\mathcal{L}(\mathbf{u}_1, A\mathbf{u}_1, \dots, A^k\mathbf{u}_1) \text{ и } \mathcal{L}(\mathbf{v}_1, A^*\mathbf{v}_1, \dots, (A^*)^k\mathbf{v}_1), \quad k = 0, 1, \dots, n-1$$

и биортогональными тогда и только тогда, когда каждый из векторов $\mathbf{u}_k \in \mathcal{R}^{(n)}$, $\mathbf{v}_k \in \mathcal{R}^{*(n)}$, $k = 2, 3, \dots, n$, будут определяться с помощью следующего итерационного процесса:

$$\mathbf{u}_k = \lim_{i \rightarrow \infty} \mathbf{u}_k^{(i)}, \quad \mathbf{v}_k = \lim_{i \rightarrow \infty} \mathbf{v}_k^{(i)}, \quad \mathbf{u}_k^{(i)} = -\beta_{k-1,k-2}^{(i-1)} \mathbf{u}_{k-2} - \beta_{k-1,k-1}^{(i-1)} \mathbf{u}_{k-1} + \mathbf{u}_k^{(i-1)},$$

$$\mathbf{v}_k^{(i)} = -\beta_{k-1,k-2}^{(i-1)} \mathbf{v}_{k-2} - \beta_{k-1,k-1}^{(i-1)} \mathbf{v}_{k-1} + \mathbf{v}_k^{(i-1)},$$

$$\beta_{k-1,k-2}^{(i-1)} = \frac{(\mathbf{v}_k^{(i-1)}, \mathbf{u}_{k-2})}{(\mathbf{u}_{k-2}, \mathbf{v}_{k-2})}, \quad \beta_{k-1,k-1}^{(i-1)} = \frac{(\mathbf{u}_k^{(i-1)}, \mathbf{v}_{k-1})}{(\mathbf{u}_{k-1}, \mathbf{v}_{k-1})},$$

$$\mathbf{u}_k^{(0)} \equiv A\mathbf{u}_{k-1}, \quad \mathbf{v}_k^{(0)} \equiv A^*\mathbf{v}_{k-1},$$

который для получения хороших приближений

$$\|\mathbf{u}_k - \mathbf{u}_k^{(l)}\|_{\mathcal{R}^{(n)}} \leq \varepsilon, \quad \|\mathbf{v}_k - \mathbf{v}_k^{(l)}\|_{\mathcal{R}^{*(n)}} \leq \varepsilon$$

векторов \mathbf{u}_k и \mathbf{v}_k достаточно оборвать на l -м шаге при достижении каждого из неравенств

$$\left| \beta_{k-1,k-2}^{(l)} u_{k-2,p} + \beta_{k-1,k-1}^{(l)} u_{k-1,p} \right| \leq \varepsilon_0, \quad \left| \beta_{k-1,k-2}^{(l)} v_{k-2,p} + \beta_{k-1,k-1}^{(l)} v_{k-1,p} \right| \leq \varepsilon_0,$$

где в качестве «порога» $\varepsilon_0 = \varepsilon/\sqrt{n}$ берется число $\varepsilon_0 = 2^{-s} w_{kp} = 2^{-s} \min_{i,j} \left(|u_{ki}^{(0)}|, |v_{kj}^{(0)}| \right)$.

Доказательства каждой из этих трех конструктивных теорем отличаются одно от другого незначительными нюансами. Поэтому достаточно привести хотя бы одно из них. Ограничимся доказательством теоремы 8.1.

Покажем, прежде всего, что процесс доортогонализации сходится, т. е. при $i = 1, 2, 3, \dots$

$$\mathbf{u}_k^{(i)} \rightarrow \mathbf{u}_k \perp \mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}).$$

Обозначим через матрицу B_{ki} переход от системы векторов $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}, \mathbf{u}_k^{(i+1)}$ к системе $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}, \mathbf{u}_k^{(i)}$, которая вследствие рекуррентных соотношений (8.1) имеет вид

$$B_{ki} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \beta_{k-1,1}^{(i)} & \beta_{k-1,2}^{(i)} & \dots & \beta_{k-1,k}^{(i)} \end{bmatrix}. \quad (8.6)$$

Так как векторы $\tilde{\mathbf{u}}_k^{(i)}$ и \mathbf{u}_j , $j = 1, 2, \dots, k-1$, почти ортогональны, то элементы $\beta_{k-1,j}^{(i)}$ матрицы B_{ki} удовлетворяют неравенствам $|\beta_{k-1,j}^{(i)}| \ll 1$ при $j \leq k-1$, $i = 1, 2, 3, \dots$

Далее, из очевидного равенства

$$\|\tilde{\mathbf{u}}_k^{(i+1)}\|^2 = \|\mathbf{u}_k^{(i)}\|^2 - \sum_{j=1}^{k-1} (\mathbf{u}_j, \mathbf{u}_k^{(i)})^2 = 1 - \sum_{j=1}^{k-1} (\beta_{kj}^{(i)})^2$$

вытекает

$$0 \leq |\beta_{kj}^{(i)}| < \beta_{k-1,k}^{(i)} \leq 1. \quad (8.7)$$

Поэтому каждая из матриц B_{ki} , $i = 0, 1, 2, \dots$, невырождена и имеет обратную:

$$B_{ki}^{-1} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \gamma_{k-1,1}^{(i)} & \gamma_{k-1,2}^{(i)} & \dots & \gamma_{k-1,k}^{(i)} \end{bmatrix}, \quad (8.8)$$

где, как нетрудно получить $\gamma_{k-1,j}^{(i)} = -\frac{\beta_{k-1,j}^{(i)}}{\beta_{k-1,k}^{(i)}}$ при $j \leq k-1$ и $\gamma_{k-1,k}^{(i)} = -\frac{1}{\beta_{k-1,k}^{(i)}}$. Из соотношений (8.3) следует, что элементы $\gamma_{k-1,j}^{(i)}$ обратной матрицы B_{ki}^{-1} подчиняются неравенствам

$$0 \leq |\gamma_{k-1,j}^{(i)}| < 1 \leq \gamma_{k-1,k}^{(i)}. \quad (8.9)$$

Обозначим теперь через $U_p^{(n,k)}$ почти полуортогональную матрицу, составленную из векторов-столбцов $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}, \mathbf{u}_k^{(p)}$, а через C -матрицу из векторов-столбцов $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}, \mathbf{A}\mathbf{u}_{k-1}$. Тогда процесс доортогонализации (8.1) можно записать в виде следующего матричного равенства:

$$U_p^{(n,k)} = C \left(\left[\prod_{i=0}^p B_{ki} \right]^{-1} \right)^*, \quad (8.10)$$

а доказательство сходимости процесса (8.1) сводится, следовательно, к определению предела $\lim_{p \rightarrow \infty} U_p^{(n,k)} = U^{(n,k)}$ по норме $R^{(k)}$, $k \leq n$.

Оценим разность $\|U_q^{(n,k)} - U_p^{(n,k)}\|$, полагая для определенности, что $q > p$. Из равенства (8.10) находим

$$U_q^{(n,k)} - U_p^{(n,k)} = C \left(\left[\prod_{i=0}^q B_{ki}^* \right]^{-1} - \left[\prod_{i=0}^p B_{ki}^* \right]^{-1} \right) = C \left[\prod_{i=0}^p B_{ki}^* \right]^{-1} \left(\left[\prod_{i=p}^q B_{ki}^* \right]^{-1} - E^{(k)} \right) = U_p^{(n,k)} \left(\left[\prod_{i=p}^q B_{ki}^* \right]^{-1} - E^{(k)} \right).$$

Отсюда, ссылаясь на неравенства (8.7), получаем

$$\|U_q^{(n,k)} - U_p^{(n,k)}\| \leq \|U_p^{(n,k)}\| \| (B_{kp}^{-1})^{q-p} - E^{(k)} \|.$$

Последнее неравенство с учетом соотношений (8.9) при $q, p \rightarrow \infty$ показывает, что последовательность матриц $\{U_p^{(n,k)}\}$ сходится в себе по норме пространства матриц. Следовательно, при $q, p \rightarrow \infty$

$$\| \|U_q^{(n,k)}\| - \|U_p^{(n,k)}\| \| \rightarrow 0,$$

т.е. числовая последовательность $\{ \|U_p^{(n,k)}\| \}$ сходится в себе и, значит, ограничена. А это, в свою очередь, означает, что $U^{(n,k)} = \lim_{p \rightarrow \infty} U_p^{(n,k)}$ по норме $R^{(k)}$, $k \leq n$.

Для завершения доказательства теоремы заметим, что в условиях ограниченной точности представления чисел нет необходимости в получении приближений $\mathbf{u}_k^{(i)}$ с большими значениями индекса i (при этом вычисления на компьютере «зацикливаются»). Достаточно ограничиться таким приближением $\mathbf{u}_k^{(l)}$, для которого поправка $\sum_{j=1}^{k-1} \beta_{kj}^{(l)} \mathbf{u}_j$ в любую из компонент $u_{kp}^{(l)}$, $p = 1, 2, \dots, n$, вектора $\mathbf{u}_k^{(l)}$ не превышает по абсолютной величине точности ее изображения в ЭВМ. Теорема доказана полностью.

В ходе доказательства теоремы обнаружено (в частности из соотношения (8.10)), что элементы β_{kj} матрицы B в представлении (1.10) необходимо вычислять по формулам

$$\beta_{kj} = \lim_{p \rightarrow \infty} (\beta_{kj}^{(0)} + \sum_{i=1}^p \beta_{kj}^{(i)} \prod_{m=0}^{i-1} \beta_{k,k+1}^{(m)}), \quad j \leq k; \quad \beta_{k,k+1} = \lim_{p \rightarrow \infty} \left(\prod_{i=0}^p \beta_{k,k+1}^{(i)} \right), \quad k = 1, 2, \dots, n-1, \quad (8.11)$$

причем рационально при этом ограничиться значениями $p=l$. Аналогично определяются элементы матриц B и в канонических разложениях (2.9), (3.5) и (4.6) из [1].

Для того чтобы оценить точность получаемых таким образом элементов матриц B во всех представлениях (1.10), (2.9), (3.5) и (4.6) из [1], введем в рассмотрение матрицы погрешности $\Delta_B, \Delta_U, \Delta_V, \Delta_Q$ соответственно матриц B, U, V, Q и определим относительные погрешности последних соотношениями

$$\delta B = \frac{\|\Delta_B\|}{\|B\|}, \quad \delta U = \frac{\|\Delta_U\|}{\|U\|}, \quad \delta V = \frac{\|\Delta_V\|}{\|V\|}, \quad \delta Q = \frac{\|\Delta_Q\|}{\|Q\|}.$$

Если, исходя из теорем 8.1 — 8.3, можно сказать, что точность вычисления ортогональной (A -ортогональной или биортогональной) системы векторов практически ограничена лишь точностью изображения чисел в данной ЭВМ, то уже в отношении точности элементов β_{ki} или b_{ki} матриц B справедливы следующие следствия из упомянутых теорем.

Следствие 8.1. Относительные погрешности матриц B из представлений (1.10) и (2.9) оцениваются неравенством $\delta B \leq c(A) \delta U$, где $c(A) = \|A\| \|A^{-1}\|$ — мера обусловленности исходной матрицы A .

Доказательство следует из очевидного равенства

$$U\Delta_B = A\Delta_U - \Delta_U B.$$

Следствие 8.2. Относительные погрешности матрицы B из разложения (3.5) оцениваются неравенством $\delta B \leq c^2(Q)(1+c(Q))\delta Q$, где при вычислении числа $c(Q)$ следует учитывать, что $\|Q\|^2 = \sum_{i=1}^n (q_i, q_i)$, $\|Q^{-1}\| = \|Q^* A\|$.

В самом деле, из легко определяемого выражения $Q\Delta_B = A\Delta_Q - \Delta_Q B$ с учетом равенств $Q^* A Q = E^{(n)}$, $A^{-1} = Q Q^*$ сразу же получаем $\|\Delta_b\| \leq c(Q)\|A\|(1+c(Q))\delta Q$.

А из матриц A и B следует

$$\|A\| \leq c(Q)\|B\|$$

откуда находим требуемый результат.

Следствие 8.3. Относительные погрешности матрицы B в представлениях (4.6) оцениваются неравенством $\delta B \geq c(V)\delta U$, где $c(V) = \|U\| \|V\| = c(U)$.

Доказательство аналогично предыдущим.

Таким образом, следствиями теорем 8.1 — 8.3 подтверждается известный в теории устойчивости численных решений возмущающих систем линейных уравнений факт: в вопросах устойчивости численных методов косвенного определения коэффициентов характеристического многочлена матрицы решающую роль играет мера обусловленности матрицы. Чем хуже обусловлена матрица, тем с меньшей точностью определяются коэффициенты характеристического полинома.

Заключение. В статье изложены согласно литературным данным некоторые методы косвенного определения коэффициентов характеристического полинома матрицы, основанные на ортогонализации некоторых последовательностей векторов. Охарактеризованы области применения данных методов при решении других проблем вычислительной алгебры, в частности при определении собственных векторов и отыскании решений систем линейных уравнений.

В развитии положений работы [8] установлено, что методы определения коэффициентов полинома в классическом их исполнении мало пригодны для практических вычислений на компьютерах, особенно в условиях плохой обусловленности матриц. В этой связи сформулирована задача определения с высокой точностью искомым коэффициентов и предложены способы устойчивой реализации классических методов на компьютере, учитывающие неподчинение машинной арифметики обычным законом коммутативности, ассоциативности и дистрибутивности.

Под этим углом зрения дана, в отличие от традиционной [2, 3 — 7] и первоначальная трактовка методов, включающая построение ортонормированных или A -ортонормированных последовательностей векторов. Указаны способы выбора исходных векторов строящихся совокупностей, благоприятствующих нормальному течению того или иного процесса вычислений. Попутно доказана теорема 6.1, обобщающая на случай несимметрической матрицы известный способ решения систем уравнений [7]. С помощью приемов, описанных в работе [8], найдены устойчивые процедуры доортогонализации ортонормированной, A -ортонормированной или биортонормированной систем векторов, позволяющие определить с высокой точностью как сами векторы конструируемых множеств, так и коэффициенты характеристического полинома. Доказаны теоремы 8.1 — 8.3 о сходимости процессов доортогонализации ортонормированных, A -ортонормированных или биортонормированных последовательностей векторов. Указаны гарантированные оценки точности искомым коэффициентов на ЭВМ, естественно завышенные, так как при их получении рассматривались самые неблагоприятные возможные погрешности машинных операций.

Список литературы

1. Черная О. А., Якимчик А. И. О процессах доортогонализации некоторых семейств векторов, возникающих при построении характеристических полиномов матриц и используемых при решении систем линейных алгебраических уравнений. 1 // Геофиз. журн. — 2005. — 27, №3. — С. 503—511.
2. Березин И. С., Жигков Н. П. Методы вычислений. Т. 2. — М.: Физматгиз, 1960. — 620 с.
3. Воеводин В. В. Вычислительные основы линейной алгебры. — М.: Наука, 1977. — 304 с.
4. Гомунов С. К. Решение систем линейных уравнений. — Новосибирск: Наука, 1980. — 177 с.
5. Ланцош К. Практические методы прикладного анализа. — М.: Физматгиз, 1961. — 524 с.
6. Уилкинсон Дж. Х. Алгебраическая проблема собственных значений. — М.: Наука, 1970. — 564 с.
7. Фаддеев Д. К., Фаддеева В. Н. Вычислительные методы линейной алгебры. — М. — Л.: Физматгиз, 1963, — 734 с.
8. Черный А. В. О точности численных решений некоторых задач геофизики // Теория и практика интерпретации гравитационных и магнитных полей в СССР. — Киев: Наук. думка, 1983. — С. 263—290.

Вниманию читателей!

В первой части статьи («Геофизический журнал», Т. 27, №3, 2005 г.) выражения $u_1 \in P^{(n)}$; $A: P^{(n)} \rightarrow P^{(n)}$; $P^{(n)}$; $x, y \in P^{(n)}$; $P_A^{(n)}$ на с. 506—510 следует читать так: $u_1 \in \mathcal{R}^{(n)}$; $A: \mathcal{R}^{(n)} \rightarrow \mathcal{R}^{(n)}$; $\mathcal{R}^{(n)}$; $x, y \in \mathcal{R}^{(n)}$; $\mathcal{R}_A^{(n)}$.