А.Г.Нестеренко

УНИТАРНАЯ ОРТОГОНАЛИЗАЦИЯ СИСТЕМЫ ВЕКТОРОВ[1]

Abstract

Приводится итерационный метод ортогонализации произвольного набора комплекснозначных векторов, использующий унитарные матрицы.

1. Элементарное вращение. Для заданного набора комплексных векторов $z_1 = (z_{11}, ..., z_{n1})^T, ..., z_m = (z_{1m}, ..., z_{nm})^T$, образующих столбцы матрицы

$$A = \begin{pmatrix} z_{11}, & ..., & z_{1m} \\ . & . & . \\ z_{n1}, & ..., & z_{nm} \end{pmatrix} \in \mathbf{C}^{n\times m},$$

и заранее выбранных номеров $i, j \in \{1, 2, ..., m\}$ построим унитарную $m \times m$-матрицу элементарного вращения [2] $U_{ij}(\varphi, \psi) = (u_{kl}), \quad i < j \quad :$

$$u_{ii} = u_{jj} = \cos\varphi = \alpha; \quad u_{ji} = -\bar{u}_{ij} = \mathrm{e}^{-i\psi} \cdot \sin\varphi = \beta;$$

прочие диагональные элементы матрицы U_{ij} равны единице, а внедиагональные - нулю.

Вещественные параметры φ, ψ подберем так, чтобы в произведении матриц $A \cdot U_{ij} =: A^1 = (z^1_{ij})$ вектор-столбцы

$$z^1_i = \alpha \cdot z_i + \beta \cdot z_j, \qquad z^1_j = -\bar{\beta} \cdot z_i + \alpha z_j$$

отвечали следующим условиям:

$$I)\ (z^1_i, \bar{z}^1_j) = 0, \qquad II)\ \|z^1_i\| \geq \|z_i\|.$$

Остальные столбцы в матрице A^1, очевидно, не меняются, $z^1_k = z_k, \quad k \neq i, j$.

Здесь и далее для произвольных комплекснозначных векторов $v_i, v_j \in \mathbf{C}^n$ используются обозначения

$$(v_i, \bar{v}_j) = \sum_{s=1}^{n} v_{si} \cdot \bar{v}_{sj}, \qquad \|v_i\|^2 = (v_i, \bar{v}_i) = \sum_{s=1}^{n} |v_{si}|^2.$$

[1]Исследования проводятся при финансовой поддержке РФФИ, грант N 98-01-03628.

Мы имеем:

$$\begin{aligned}(z_i^1, \bar z_j^1) &= (\alpha z_i + \beta z_j, -\beta \bar z_i + \alpha \bar z_j)\\ &= -\alpha\beta\|z_i\|^2 + \alpha^2(z_i, \bar z_j) - \beta^2(z_j, \bar z_i) + \alpha\beta\|z_j\|^2\\ &= -\alpha\beta(\|z_i\|^2 - \|z_j\|^2) + \alpha^2(z_i, \bar z_j) - \beta^2(z_j, \bar z_i).\end{aligned}$$

Преобразуем разность

$$\begin{aligned}\|z_i^1\|^2 - \|z_i\|^2 &= (\alpha z_i + \beta z_j, \alpha \bar z_i + \bar\beta \bar z_j) - \|z_i\|^2\\ &= \alpha^2\|z_i\|^2 + \alpha\bar\beta(z_i, \bar z_j) + \alpha\beta(z_j, \bar z_i) + |\beta|^2\|z_j\|^2 - \|z_i\|^2\\ &= -|\beta|^2(\|z_i\|^2 - \|z_j\|^2) + \alpha\bar\beta(z_i, \bar z_j) + \alpha\beta(z_j, \bar z_i).\end{aligned}$$

Обозначим:

$$\|z_i\|^2 - \|z_j\|^2 =: D; \quad (z_i, \bar z_j) =: r\,\mathrm{e}^{i\theta}; \quad (r > 0).$$

Тогда первое условие $I)$ имеет вид:

$$-\frac{\sin 2\varphi}{2}\,\mathrm{e}^{-i\psi}\cdot D + r(\cos^2\varphi\,\mathrm{e}^{i\theta} - \sin^2\varphi\,\mathrm{e}^{-i2\psi}\,\mathrm{e}^{-i\theta}) = 0,$$

или

$$-\sin 2\varphi\frac{D}{2} + r\left(\cos^2\varphi\,\mathrm{e}^{i(\psi+\theta)} - \sin^2\varphi\,\mathrm{e}^{-i(\psi+\theta)}\right) = 0.$$

Выделяя отдельно уравнения для вещественной и мнимой частей получим:

$$\begin{aligned}&\sin 2\varphi\frac{D}{2} = r\cos 2\varphi\cos(\psi+\theta),\\ &\sin(\psi+\theta) = 0.\end{aligned}$$

Из условия $II)$ в новых обозначениях имеем

$$\|z_i^1\|^2 - \|z_i\|^2 = -\sin^2\varphi\cdot D + r\frac{\sin 2\varphi}{2}\left(\mathrm{e}^{i(\psi+\theta)} + \mathrm{e}^{-i(\psi+\theta)}\right) > 0,$$

что можно представить в виде

$$\|z_i^1\|^2 - \|z_i\|^2 = -\sin^2\varphi\cdot D + r\sin 2\varphi\cdot\cos(\psi+\theta) > 0.$$

Рассмотрим отдельные альтернативы:

$a^o)$ $(z_i, \bar z_j) = 0;$ $\|z_i\| > \|z_j\|,$ т.е. $r = 0; D > 0;$

$b^o)$ $(z_i, \bar z_j) = 0;$ $\|z_i\| = \|z_j\|,$ т.е. $r = 0; D = 0;$

$c^o)$ $(z_i, \bar z_j) = 0;$ $\|z_i\| < \|z_j\|,$ т.е. $r = 0; D < 0;$

$d^o)$ $(z_i, \bar z_j) \neq 0;$ $\|z_i\| = \|z_j\|,$ т.е. $r \neq 0; D = 0;$

$e^o)$ $(z_i, \bar z_j) \neq 0;$ $\|z_i\| \neq \|z_j\|,$ т.е. $r \neq 0; D \neq 0.$

Подставляя соответствующие величины r и D в полученные ранее формулы, после несложных упрощений имеем:
$a^o)$ $\quad r=0, \quad D>0$; выбираем $\varphi=0, \psi=0; \quad \alpha=1, \beta=0; \quad U_{ij}=E$;
$b^o)$ $\quad r=0, \quad D=0$ для простоты $\varphi=0, \psi=0; \quad \alpha=1, \beta=0; \quad U_{ij}=E$;
$c^o)$ $\quad r=0, \quad D<0$; принимаем $\alpha=0, \beta=1$; т.е. $z_i^{k+1}=z_j^k; \quad z_j^{k+1}=z_i^k$;

$$d^o) \quad r\neq 0, \quad D=0; \quad \Rightarrow \quad \begin{cases} \cos 2\varphi=0, \\ \sin(\psi+\theta)=0, \\ \sin 2\varphi\cos(\psi+\theta)>0; \end{cases}$$

этой системе удовлетворяют значения $\varphi=\pi/4; \psi=-\theta$, что влечет $\alpha=\sqrt{2}/2; \beta=\sqrt{2}/2\cdot \mathrm{e}^{i\theta}$;

$$e^o) \quad r\neq 0, \quad D\neq 0; \quad \Rightarrow \quad \begin{cases} \operatorname{tg} 2\varphi=\dfrac{2r}{D}, \\ \sin(\psi+\theta)=0, \\ \operatorname{sgn}(D)=\operatorname{sgn}(\cos 2\varphi), \end{cases}$$

что влечет

$$\begin{cases} \varphi=\dfrac{1}{2}\operatorname{arctg}\dfrac{2r}{D}, \\ \psi=-\theta, \end{cases} \text{при } D\geq 0, \text{ либо } \begin{cases} \varphi=\dfrac{1}{2}\operatorname{arctg}\dfrac{2r}{D}+\pi/2, \\ \psi=-\theta, \end{cases} \text{при } D<0.$$

После первого вращения повторим ту же процедуру для матрицы A^1, изменив один или оба индекса $i,j\in\{1,2,...,m\}$, тем самым организуя итерационный процесс $A^{k+1}=A^k\cdot U^k(\varphi_k,\psi_k); \quad k=0,1,2,...$, на каждом шаге которого соблюдаются условия:

$$I)(z_i^{k+1},\bar{z}_j^{k+1})=0, \qquad II)\|z_i^{k+1}\|\geq\|z_i^k\|,$$

и, следовательно, имеем систему

$$\begin{cases} \dfrac{D_k}{2}\sin 2\varphi_k=r_k\cos 2\varphi_k\cos(\psi_k+\theta_k), \\ \sin(\psi_k+\theta_k)=0, \end{cases}$$

и по аналогии записывается разность

$$\|z_i^{k+1}\|^2-\|z_i^k\|^2=-\sin^2\varphi_k\cdot D_k+r_k\sin 2\varphi_k\cdot\cos(\psi_k+\theta_k),$$

где

$$\|z_i^k\|^2-\|z_j^k\|^2=: D_k; \quad (z_i^k,\bar{z}_j^k)=: r_k\,\mathrm{e}^{i\theta_k}\,.$$

При этом справедлива

Лемма 1. Если для некоторого k выполняется равенство $\|z_i^{k+1}\|^2 = \|z_i^k\|^2$, то $(z_i^k, \bar{z}_j^k) = 0$.

Доказательство. Равенство нулю $(z_i^k, z_j^k) = 0$ в наших обозначениях равносильно равенству $r_k = 0$, поэтому из вышеобозначенных вариантов $a^o) - e^o)$ требуют рассмотрения только случаи $d^o), e^o)$.

Из условий леммы для соблюдения равенства $\|z_i^{k+1}\|^2 = \|z_i^k\|^2$, в варианте $d^o)$ получаем

$$\begin{cases} \cos 2\varphi = 0, \\ \sin(\psi + \theta) = 0, \\ r \sin 2\varphi \cos(\psi + \theta) = 0, \end{cases}$$

а по свойствам тригонометрических функций, из первого и второго уравнений, для третьего вытекает $r_k = 0$.

Рассмотрим вариант $e^o)$:

$$\begin{cases} D \operatorname{tg} 2\varphi = 2r \cos(\psi + \theta), \\ \sin(\psi + \theta) = 0, \\ 0 = \|z_i^{k+1}\|^2 - \|z_i^k\|^2 = \dfrac{D \sin^2 \varphi}{\cos 2\varphi}. \end{cases}$$

В третьем уравнении необходимо имеем $D \sin \varphi = 0$, тогда из первого получим $D \operatorname{tg} 2\varphi = 0$, следовательно, $2r \cos(\psi + \theta) = 0$, но из второго уравнения вытекает неравенство $\cos(\psi + \theta) \neq 0$, и, значит, $r = 0$, что и заканчивает доказательство леммы 1.

Смысл леммы 1 в том, что если только векторы z_i^k и z_j^k не являются взаимно ортогональными, то норма вектора z_i^k после преобразования вращения обязательно увеличивается, $\|z_i^{k+1}\| > \|z_i^k\|$.

Лемма 2. Если на некоторой итерации векторы z_i^k, z_j^k $(i < j)$ оказываются линейно зависимыми, то в результате очередного шага вектор z_j^{k+1} обращается в нуль.

Доказательство. Линейная зависимость пары векторов означает $z_j^k = \lambda z_i^k$; $\lambda \in \mathbf{C}$, тогда $z_i^{k+1} = (\alpha + \lambda\beta) z_i^k$; $z_j^{k+1} = (-\bar{\beta} + \lambda\alpha) z_i^k$. Опуская тривиальный нулевой случай, всюду далее считаем $z_i^k \neq 0$.

Из условия $I)$ имеем:

$$0 = (z_i^{k+1}, \overline{z_j^{k+1}}) = ((\alpha + \lambda\beta) z_i^k, \overline{(-\bar{\beta} + \lambda\alpha) z_i^k}) = (\alpha + \lambda\beta)(-\beta + \bar{\lambda}\alpha)\|z_i^k\|^2.$$

Так как $z_i^k \neq 0$, то $(\alpha + \lambda\beta)(-\beta + \bar{\lambda}\alpha) = 0$, а для обеспечения условия $II)$ необходимо выполнение неравенства $z_i^{k+1} \neq 0$ или $(\alpha + \lambda\beta) \neq 0$, следовательно, единственно возможный вариант - это $(-\beta + \bar{\lambda}\alpha) = 0$, откуда вытекает заключение $z_j^{k+1} = (-\bar{\beta} + \lambda\alpha) z_i^k = 0$, что завершает доказательство леммы.

Лемма 2 показывает, что в процессе итерационных вращений с соблюдением условий $I), II)$ обнуляются все линейно зависимые векторы или их комбинации, и в пределе количество ненулевых векторов определяется рангом исходной матрицы.

2. Итерационная ортогонализация. В имеющихся реализациях данного алгоритма (для вещественного случая см.[3, 4]) используется следующая итерационная схема: на каждом шаге увеличивается норма первого столбца изменяющейся матрицы, а в пределе имеем матрицу, первый вектор-столбец которой ортогональный всем остальным столбцам. При численной реализации условие ортогональности выполняется с точностью до некоторого заданного ε.

Возможны различные варианты:

1) $A^{k+1} = A^k \cdot U_{12}^k(\varphi_k, \psi_k)$; $\quad A^{k+2} = A^{k+1} \cdot U_{13}^{k+1}(\varphi_{k+1}, \psi_{k+1}); \ldots A^{k+m-1} = A^{k+m-2} \cdot U_{1m}^{k+m-2}(\varphi_{k+m-2}, \psi_{k+m-2})$; критерием останова является условие: $U_{12}^k = U_{13}^{k+1} = \ldots = U_{1m}^{k+m-2} = E$, то есть за полный "цикл" не изменился ни один из столбцов матрицы A;

2) $A^{k+1} = A^k \cdot U_{m-1,m}^k(\varphi_k, \psi_k)$; $\quad A^{k+2} = A^{k+1} \cdot U_{m-2,m-1}^{k+1}(\varphi_{k+1}, \psi_{k+1}); \ldots$ $A^{k+m-1} = A^{k+m-2} \cdot U_{12}^{k+m-2}(\varphi_{k+m-2}, \psi_{k+m-2})$; затем проводится цикл по варианту 1) и проверяется условие отсутствия изменений первого столбца, что одновременно означает его ортогональность всем остальным столбцам.

Предельную матрицу, отмечая ортогональность первого столбца всем остальным, обозначим $A_{[1]}$, а композицию унитарных матриц поворота обозначим $U_{[1]} = \prod_k U_{1s}^k$, т.е. $A_{[1]} = A \cdot U_{[1]}$.

Продолжим циклическую процедуру "вращений" матрицы $A_{[1]}$ относительно второго столбца, не затрагивая при этом первый. Например: $A_{[1]}^{k+1} = A_{[1]}^k \cdot U_{23}^k(\varphi_k, \psi_k)$; $\quad A_{[1]}^{k+2} = A_{[1]}^{k+1} \cdot U_{24}^{k+1}(\varphi_{k+1}, \psi_{k+1}); \ldots A_{[1]}^{k+m-2} = A_{[1]}^{k+m-3} \cdot U_{2m}^{k+m-3}(\varphi_{k+m-3}, \psi_{k+m-3}); \ldots$ с критерием останова $U_{23}^k = U_{24}^{k+1} = \ldots = U_{2m}^{k+m-3} = E$.

В пределе получим матрицу $A_{[2]}$, в которой первые два столбца ортогональны всем остальным и друг другу. Соответствующую унитарную композицию обозначим $U_{[2]} = \prod_k U_{2s}^k$. Таким образом $A_{[2]} = A_{[1]} \cdot U_{[2]}$ или $A_{[2]} = A \cdot U_{[1]} \cdot U_{[2]}$. Продолжая далее, получим матрицы $A_{[3]}, A_{[4]}, \ldots$

На последнем этапе имеем матрицу, в которой первые $m-2$ вектор-столбца попарно ортогональны всем остальным столбцам матрицы $A_{[m-2]}$, получающейся через композиционное произведение $A_{[m-2]} =$

$A_{[m-3]} \cdot U_{[m-3]} = A \cdot U_{[1]} \cdot U_{[2]} \cdot \ldots \cdot U_{[m-3]}$, и достаточно провести в точности одно вращение $A_{[m-1]} = A_{[m-2]} \cdot U_{m-1,m}(\varphi, \psi)$, чтобы ортогонализовать векторы a_{m-1} и a_m. Таким образом, результирующая матрица с попарно-ортогональными столбцами, получается из первоначальной через серию унитарных преобразований: $A_{[m-1]} = A \cdot U_{[1]} \cdot \ldots \cdot U_{[m-3]} \cdot U_{m-1,m}(\varphi, \psi)$, или $A_{[m-1]} = A \cdot V$, где матрица $V = U_{[1]} \cdot \ldots \cdot U_{m-1,m}(\varphi, \psi)$ является, очевидно, унитарной.

3. Сингулярное разложение. В полученной матрице $A_{[1]}$ первый столбец ортогонален всем остальным. Поэтому произведение матриц $A^*_{[1]}A_{[1]}$ имеет структуру

$$A^*_{[1]}A_{[1]} = \begin{pmatrix} (z_1^{[1]}, \bar{z}_1^{[1]}) & 0 & 0 & \ldots & 0 \\ 0 & (z_2^{[1]}, \bar{z}_2^{[1]}) & (z_2^{[1]}, \bar{z}_3^{[1]}) & \ldots & (z_2^{[1]}, \bar{z}_m^{[1]}) \\ . & . & . & & \\ 0 & (z_m^{[1]}, \bar{z}_2^{[1]}) & (z_m^{[1]}, \bar{z}_3^{[1]}) & \ldots & (z_m^{[1]}, \bar{z}_m^{[1]}) \end{pmatrix},$$

и очевидно, что квадрат нормы первого столбца $\|z_1^{[1]}\|^2$ является для $A^*_{[1]}A_{[1]}$ собственным числом: $\det(A^*_{[1]}A_{[1]} - \|z_1^{[1]}\|^2 \cdot E) = 0$, а сама норма $\|z_1^{[1]}\|$ – сингулярным числом исходной матрицы $A_{[1]}$ – см. [1].

Соответственно матрица $A^*_{[2]} \cdot A_{[2]}$ имеет в первых двух строках и столбцах ненулевыми только два диагональных элемента – квадраты норм соответствующих столбцов матрицы $A_{[2]}$.

Произведение $A^*_{[m-1]} \cdot A_{[m-1]}$ – диагональная матрица, образуемая квадратами сингулярных чисел.

Нормируем столбцы матрицы $A_{[m-1]}$ и дополним их, быть может до количества n необходимым числом нулевых столбцов, тем самым определяя матрицу W. Обозначим также диагональную матрицу соответствующих сингулярных чисел (т.е. норм столбцов $A_{[m-1]}$) размерности $n \times m$ через S, тогда $A_{[m-1]} = W \cdot S$. С другой стороны в результате вращений получено разложение $A_{[m-1]} = A \cdot V$. В итоге, учитывая унитарность V, получаем $A = W \cdot S \cdot V^*$. Данное разложение матрицы A может отличаться (в некоторых случаях) от сингулярного лишь возможной неполнотой левого сингулярного базиса – нулевые столбцы матрицы W, обусловленной размерностью исходной матрицы и ее рангом.

Литература

[1] **Воеводин В.В., Кузнецов Ю.А.** Матрицы и вычисления. М.: Наука, 1984.

[2] **Калиткин Н.Н.** Численные методы. М.: Наука, 1978.

[3] **Лежнев В.Г.** Метод решения алгебраической спектральной задачи // Численные методы анализа. М.: Изд-во Моск. ун-та, 1995.

[4] **Лежнев В.Г., Нестеренко А.Г.** Метод сингулярного разложения матрицы // Численные методы анализа. М.: Изд-во Моск. ун-та, 1995.