Aplikace matematiky

Ján Buša

Об одном методе регуляризации решения систем линейных алгебраических уравнений

Aplikace matematiky, Vol. 35 (1990), No. 2, 129–139

Persistent URL: `http://dml.cz/dmlcz/104394`

35 (1990) APLIKACE MATEMATIKY No. 2, 129—139

ОБ ОДНОМ МЕТОДЕ РЕГУЛЯРИЗАЦИИ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ

Ján Buša

(Поступило в редакцию 7. IX. 1988 г.)

Summary. В работе предложен и обоснован новый метод стабилизации решения приближенно заданной системы линейных алгебраических уравнений. Доказана устойчивость метода и т.н. глобальная устойчивость метода. Коротко описана численная реализация метода.

Keywords: Solving of system of linear algebraic equations, ill-posed problems, regularization, approximator to pseudo-inverse matrix, singular value decomposition.

AMS Classification: 65F05 (15A09, 65F20).

1. На практике часто встречается задача решения системы линейных алгебраических уравнений (СЛАУ). Очень часто коэффциенты системы известны или представлены лишь приближенно. Решение такой задачи может оказаться неустойчивым. Поэтому важно разрабатывать устойчивые численные методы решения СЛАУ. В [5] автором предложен новый метод регуляризации решения СЛАУ, обладающий и свойством ,,глобальной устойчивости на множестве СЛАУ с правой стороной равномерно ограниченной". В настоящей работе дано полное доказательство теорем, опубликованных в [5] без доказательств. Теорема 2 обобщена на случай прямоугольных матриц.

Пусть матрица $\mathbf{A} \in \mathcal{M}_{m\times n}$, где $\mathcal{M}_{m\times n}$ — класс вещественных прямоугольных матриц размера $m \times n$. Согласно утверждениям 11.45–11.50 работы [2], для $\mathbf{A}$ существует сингулярное разложение

$$\mathbf{A} = \mathbf{U}\Lambda\mathbf{V}^T\,, \tag{1}$$

которое мы запишем в виде

$$\mathbf{A} = \sum_{i=1}^{\min(m,n)} \alpha_i \mathbf{y}_i \mathbf{x}_i^T\,, \tag{2}$$

где $\alpha_i \geqq 0$ суть сингулярные числа матрицы $\mathbf{A}$, $\{\mathbf{y}_i \in \mathbb{R}^m\}_{i=1}^m$ и $\{\mathbf{x}_i \in \mathbb{R}^n\}_{i=1}^n$ — ортономированные сингулярные базисы матрицы $\mathbf{A}$, Λ — диагональная прямоугольная матрица размера $m \times n$, содержащая на диагонали сингулярные числа матрицы $\mathbf{A}$, и столбцы матриц $\mathbf{U}$ и $\mathbf{V}$ образуют сингулярный базис матрицы $\mathbf{A}$.

Пусть задана СЛАУ ($\mathbf{A}$, $\mathbf{u}$)

$$(3) \qquad \mathbf{A}\mathbf{z} = \mathbf{u}\,,$$

где $\mathbf{A} \in \mathcal{M}_{m \times n}$, $\mathbf{u} \in \mathbb{R}^m$ и $\mathbf{z} \in \mathbb{R}^n$.

Под решением СЛАУ мы будем понимать *нормальное псевдорешение* СЛАУ (см. [1], с. 111). Тогда, согласно утверждению 6.54 работы [2], нормальным псевдорешением СЛАУ (3) является элемент $\mathbf{z}_{\text{НП}} \in \mathbb{R}^n$:

$$(4) \qquad \mathbf{z}_{\text{НП}} = \mathbf{A}^+ \mathbf{u}\,,$$

где $\mathbf{A}^+ \in \mathcal{M}_{n \times m}$ — псевдообратная к $\mathbf{A}$ матрица.

Матрица $\mathbf{A}^+$ имеет разложение

$$(5) \qquad \mathbf{A}^+ = \mathbf{V}\Lambda^+\mathbf{U}^T \quad \text{или} \quad \mathbf{A}^+ = \sum_{\substack{i=1 \\ \alpha_i \neq 0}}^{\min(m,n)} \alpha_i^{-1} \mathbf{x}_i \mathbf{y}_i^T\,,$$

причем матрица $\mathbf{A}^+$, псевдообратная к $\mathbf{A}$, всегда существует и единственна.

В дальнейшем мы будем использовать некоторые известные результаты, которые сформулируем в виде утверждений.

Утверждение 1. *Пусть* $\mathbf{A} \in \mathcal{M}_{m \times n}$ *имеет спектральное разложение* (1). *Тогда симметричная матрица*

$$\mathbf{S}_A = \begin{bmatrix} 0 & \mathbf{A} \\ \mathbf{A}^T & 0 \end{bmatrix} \quad \textit{имеет разложение} \quad \mathbf{S}_A = \mathbf{P} \begin{bmatrix} \Lambda & 0 & 0 \\ 0 & -\Lambda & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{P}^T\,,$$

где матрицы $\mathbf{U}$ *и* $\mathbf{V}$ *однозначно определеяют матрицу* $\mathbf{P}$ ([3], с. 22).

Утверждение 2. *Пусть* $\mathbf{B}$, $\mathbf{A}$, $\mathbf{E} \in \mathcal{M}_{m \times n}$ *и* $\mathbf{B} - \mathbf{A} = \mathbf{E}$. *Если сингулярные числа этих матриц, упорядоченные по невозрастанию, обозначить через* β_i, α_i, ε_i, $i = 1, 2, \ldots, \min(m, n)$, *тогда*

$$\sum_{i=1}^{\min(m,n)} (\beta_i - \alpha_i)^2 \leqq \sum_{i=1}^{\min(m,n)} \varepsilon_i^2 = \|\mathbf{E}\|_E^2$$

(см. [3], с. 23).

Утверждение 3. *Пусть* $\mathbf{A}$, $\mathbf{B} \in \mathcal{M}_{m \times n}$ *и* $\operatorname{rank} \mathbf{A} = \operatorname{rank} \mathbf{B}$. *Тогда*

$$\|\mathbf{B}^+ - \mathbf{A}^+\| \leqq L \|\mathbf{B}^+\|_2 \cdot \|\mathbf{A}^+\|_2 \cdot \|\mathbf{B} - \mathbf{A}\|\,,$$

где для евклидовой нормы $L = \sqrt{2}$, *если* $\operatorname{rank} \mathbf{A} < \min(m, n)$, *и* $L = 1$, *если* $\operatorname{rank} \mathbf{A} = \min(m, n)$ (см. [4]).

2. Пусть вместо точной СЛАУ $({}^{-}\mathbf{A}, {}^{-}\mathbf{u})$ нам известна СЛАУ $(\mathbf{A}, \mathbf{u})$, причем известны также оценки

$$(6) \qquad \|{}^{-}\mathbf{A} - \mathbf{A}\|_{E,2} \leqq \mu\,, \quad \|{}^{-}\mathbf{u} - \mathbf{u}\|_2 \leqq \delta\,,$$

где $\|\mathbf{A}\|_{E,2}$ означает евклидову или спектральную норму матрицы $\mathbf{A}$.

В случае $\mu \neq 0$ нормальное псевдорешение может быть неустойчивым, так как в общем случае из факта $\mathbf{A}_\mu \to {}^{-}\mathbf{A}$ не следует, что $\mathbf{A}_\mu^+ \to {}^{-}\mathbf{A}^+$.

Рассмотрим далее следующий метод определения „решения" СЛАУ, предложенный автором в [5]: при заданном параметре f разложим матрицу $\mathbf{A}$ на сумму двух матриц

$$(7) \qquad \mathbf{A} = \sum_{\alpha_i > f} \alpha_i \mathbf{y}_i \mathbf{x}_i^T + \sum_{\alpha_i \leqq f} \alpha_i \mathbf{y}_i \mathbf{x}_i^T = \mathbf{A}_1 + \mathbf{A}_2\,.$$

Определим далее матрицу $\mathbf{A}^0$, которую можно рассматривать как регуляризованную аппроксимацию псевдообратной матрицы. Положим

$$(8) \qquad \mathbf{A}^0 = \mathbf{A}_1^+ + \mathbf{A}_2^T \,.\, f^{-2}$$

и определим обобщенное решение $\mathbf{z}_f(\mathbf{A}, \mathbf{u})$ системы $(\mathbf{A}, \mathbf{u})$:

$$(9) \qquad \mathbf{z}_f(\mathbf{A}, \mathbf{u}) = \mathbf{A}^0 \mathbf{u}\,.$$

Справедлива

Теорема 1. *Метод определения обобщенного решения СЛАУ* $(\mathbf{A}, \mathbf{u})$, *заданной с точностью до* μ *и* δ, *по правилу* (9), *определяет регуляризирующий оператор, если* $f = f(\mu, \delta)$ *выбирать так, чтобы* $f(\mu, \delta) \to 0$, $\delta \,.\, f^{-1}(\mu, \delta) \to 0$ *и* $\mu \,.\, f^{-2}(\mu, \delta) \to 0$ *при* $\mu, \delta \to 0$.

Замечание 1. Понятие регуляризирующего оператора введено в [1]. Теорема 1 обозначает, что определенное в (9) „решение" является устойчивым, т.е. что оно сходится к точному нормальному псевдорешению СЛАУ $({}^{-}\mathbf{A}, {}^{-}\mathbf{u})$, когда μ и δ сходятся к нулю.

Доказательство. На основании утверждения 2 можно заключить, что $\mathbf{A}_1 \to {}^{-}\mathbf{A}$ и $\mathbf{A}_2 \to 0$ при $\mu, \delta \to 0$, причем $\operatorname{rank} \mathbf{A}_1 = \operatorname{rank} {}^{-}\mathbf{A}$ и $\|\mathbf{A}_2\|_{E,2} \leqq \mu$. Оценим разность ${}^{-}\mathbf{z}_{\text{НП}} - \mathbf{z}_f(\mathbf{A}, \mathbf{u})$:

$$(10) \qquad \begin{aligned} &\|{}^{-}\mathbf{z}_{\text{НП}} - \mathbf{z}_f(\mathbf{A}, \mathbf{u})\|_2 = \|{}^{-}\mathbf{A}^+\, {}^{-}\mathbf{u} - \mathbf{A}^0 \mathbf{u}\|_2 = \\ &= \|{}^{-}\mathbf{A}^+\, {}^{-}\mathbf{u} - \mathbf{A}_1^+\, {}^{-}\mathbf{u} + \mathbf{A}_1^+\, {}^{-}\mathbf{u} - \mathbf{A}^0\, {}^{-}\mathbf{u} + \mathbf{A}^0\, {}^{-}\mathbf{u} - \mathbf{A}^0 \mathbf{u}\|_2 \leqq \\ &\leqq \|{}^{-}\mathbf{A}^+ - \mathbf{A}_1^+\|_{E,2} \cdot \|{}^{-}\mathbf{u}\|_2 + \|\mathbf{A}_2^T\|_2 \cdot \|{}^{-}\mathbf{u}\|_2 \cdot f^{-2} + \\ &+ \|\mathbf{A}^0\|_2 \cdot \|{}^{-}\mathbf{u} - \mathbf{u}\|_2\,. \end{aligned}$$

Используя утверждение 3, мы получим

$$\|{}^{-}\mathbf{A}^+ - \mathbf{A}_1^+\|_{E,2} \leqq L \|{}^{-}\mathbf{A}^+\|_2 \cdot \|\mathbf{A}_1^+\|_2 \cdot \|{}^{-}\mathbf{A} - \mathbf{A}_1\|_{E,2}$$

Ясно, что $\|\mathbf{A}_1^+\|_2 \leqq ⨍^{-1}$ и также $\|\mathbf{A}^0\|_2 \leqq \max(⨍^{-1}, \mu \,.\, ⨍^{-2})$. Поэтому из (10) следует:

$$\|{}^{-}\mathbf{z}_{\text{НП}} - \mathbf{z}_{⨍}(\mathbf{A}, \mathbf{u})\|_2 \leqq C_1 \,.\, \mu \,.\, ⨍^{-1} + C_2 \,.\, \mu \,.\, ⨍^{-2} + $$
$$+ C_3 \,.\, \delta \,.\, \max(⨍^{-1}, \mu \,.\, ⨍^{-2}) \,.$$

Очевидно, что если выполнены условия теоремы, то $\mathbf{z}_{⨍}(\mathbf{A}, \mathbf{u}) \to {}^{-}\mathbf{z}_{\text{ПН}}$ при $\mu, \delta \to 0$. Теорема доказана.

Замечание 2. Если выбирать $⨍(\mu, \delta) = \max^{\alpha}(\mu, \delta)$ для некоторого $\alpha \in (0, \frac{1}{2})$, то условия теоремы выполнены и мы получим асимптотическую оценку

$$\|{}^{-}\mathbf{z}_{\text{ПН}} - \mathbf{z}_{⨍}(\mathbf{A}, \mathbf{u})\|_2 \lesssim C \,.\, \max^{1-2\alpha}(\mu, \delta) \,.$$

Далее мы сформулируем и докажем некоторое обобщение на случай прямоугольных матриц теоремы 2 работы [5], которое в [5] приводилось в качестве предположения.

Теорема 2. *Пусть* $\mathbf{A}$ *и* $\mathbf{B}$ – *любые две прямоугольные матрицы размера* $m \times n$. *Тогда для любого числа* $⨍ > 0$ *справедлива оценка*

$$\|\mathbf{A}^0 - \mathbf{B}^0\|_{E,2} \leqq \|\mathbf{A} - \mathbf{B}\|_{E,2} \,.\, K_{E,2} \,.\, ⨍^{-2} \,,$$

где $\mathbf{A}^0$ *и* $\mathbf{B}^0$ *определены согласно* (7) *и* (8) *и константа* $K_{E,2}$ *зависит лишь от выбора нормы и размера матриц.*

Замечание 3. Рассмотрим оценку

$$\|\mathbf{z}_{⨍}({}^{-}\mathbf{A}, {}^{-}\mathbf{u}) - \mathbf{z}_{⨍}(\mathbf{A}, \mathbf{u})\|_2 = \|{}^{-}\mathbf{A}^0\, {}^{-}\mathbf{u} - \mathbf{A}^0\mathbf{u}\|_2 =$$
$$= \|{}^{-}\mathbf{A}^0\, {}^{-}\mathbf{u} - \mathbf{A}^0\, {}^{-}\mathbf{u} + \mathbf{A}^0\, {}^{-}\mathbf{u} - \mathbf{A}^0\mathbf{u}\|_2 \leqq$$
$$\leqq \|{}^{-}\mathbf{A}^0 - \mathbf{A}^0\|_{E,2} \,.\, \|{}^{-}\mathbf{u}\|_2 + \|\mathbf{A}^0\|_2 \,.\, \|{}^{-}\mathbf{u} - \mathbf{u}\|_2 \,.$$

Видно, что если $⨍(\mu, \delta) = \max^{\alpha}(\mu, \delta)$, то

$$\|\mathbf{z}_{⨍}({}^{-}\mathbf{A}, {}^{-}\mathbf{u}) - \mathbf{z}_{⨍}(\mathbf{A}, \mathbf{u})\|_2 \leqq C \,.\, \max^{1-2\alpha}(\mu, \delta) \,, \tag{11}$$

где C – некоторая постоянная, если нормы ${}^{-}\mathbf{u}$ и $\mathbf{u}$ ограничены. Это свойство, которое мы будем называть *глобальной* устойчивостью метода на классе систем с ограниченной правой стороной, обозначает, что „решая" предложенным методом две близкие системы (при фиксированном уровне погрешности μ и δ), мы получим близкие „решения". Этим свойством в общем случае не обладают известные методы регуляризации (см. [5]).

3. Для доказательства теоремы 2 нам необходимо доказать некоторые вспомогательные утверждения.

Рассмотрим вещественные симметричные матрицы $\mathbf{A}$ и $\mathbf{B}$ размера n и их спектральные разложения

$$\mathbf{A} = \sum_{i=1}^{n} \alpha_i \mathbf{a}_i \mathbf{a}_i^T , \quad \mathbf{B} = \sum_{i=1}^{n} \beta_i \mathbf{b}_i \mathbf{b}_i^T ,$$

где $\{\alpha_i, \beta_i \in \mathbb{R}\}$ – собственные значения и $\{\mathbf{a}_i, \mathbf{b}_i \in \mathbb{R}^n\}$ – соотвествующие им собственные векторы матриц $\mathbf{A}$ и $\mathbf{B}$.

Обозначим через $\mathbf{A}_\ell$ „нижнюю срезку" матрицы $\mathbf{A}$, т.е.

$$\mathbf{A}_\ell = \sum_{i=1}^{n} \max(\alpha_i, \ell)\, \mathbf{a}_i \mathbf{a}_i^T$$

и через $\mathbf{A}^\ell$ – „верхнюю срезку", т.е.

$$\mathbf{A}^\ell = \sum_{i=1}^{n} \min(\alpha_i, \ell)\, \mathbf{a}_i \mathbf{a}_i^T$$

где ℓ – заданное число.

Замечание 4. „Верхняя срезка" матрицы обозначает, что собственные значения матрицы, которые больше заданного числа ℓ, уменьшаются (срезаются) до значения ℓ. Следующая лемма выражает тот интутивно ясный факт, что если ту же операцию срезки проделать с двумя матрицами $\mathbf{A}$ и $\mathbf{B}$, то „срезанные" матрицы остаются такими же близкими, как и $\mathbf{A}$ и $\mathbf{B}$.

Лемма 1. *Пусть $\mathbf{A}$ и $\mathbf{B}$ – вещественные симметричные матрицы размера n. Тогда имеют место неравенства:*

$$\|\mathbf{A}^\ell - \mathbf{B}^\ell\|_E \leqq \|\mathbf{A} - \mathbf{B}\|_E \tag{12}$$

$$\|\mathbf{A}_\ell - \mathbf{B}_\ell\|_E \leqq \|\mathbf{A} - \mathbf{B}\|_E \tag{13}$$

Доказательство. Докажем сначала справедливость (12). Из нее прямо следует справедливость (13), так как $\mathbf{A}_\ell = -(-\mathbf{A})^{(-\ell)}$ и поэтому, если справедливо (12),

$$\|\mathbf{A}_\ell - \mathbf{B}_\ell\|_E = \|-(-\mathbf{A})^{-\ell} + (-\mathbf{B})^{-\ell}\|_E \leqq \|(-\mathbf{B}) - (-\mathbf{A})\|_E =$$
$$= \|\mathbf{A} - \mathbf{B}\|_E .$$

Нетрудно проверить, что евклидова норма матриц размера $m \times n$, соответствует скалярному произведению матриц размера $m \times n$, введенному в $\mathcal{M}_{m \times n}$ следующим образом:

$$(\mathbf{A}, \mathbf{B}) = \operatorname{tr}(\mathbf{A}^T\mathbf{B}) , \quad \text{где } \operatorname{tr}(\mathbf{A}) \text{ обозначает след матрицы } \mathbf{A} .$$

Тогда $\|\mathbf{A}\|_E^2 = \operatorname{tr}(\mathbf{A}^T\mathbf{A})$.

Отметим следующие свойства следа матриц, которые мы используем дальше:

а) $\operatorname{tr}(\mathbf{A}+\mathbf{B})=\operatorname{tr}\mathbf{A}+\operatorname{tr}\mathbf{B}$

б) $\operatorname{tr}(\mathbf{A}^T\mathbf{B})=\operatorname{tr}(\mathbf{B}^T\mathbf{A})=\operatorname{tr}(\mathbf{B}\mathbf{A}^T)=\operatorname{tr}(\mathbf{A}\mathbf{B}^T)$

в) $\operatorname{tr}(\mathbf{a}\mathbf{b}^T)=\operatorname{tr}(\mathbf{b}^T\mathbf{a})=(\mathbf{a},\mathbf{b})\,,\quad \mathbf{a},\mathbf{b}\in\mathbb{R}^n\,.$

Представим матрицы $\mathbf{A}^{f}$ и $\mathbf{B}^{f}$ как результат следующего процесса: Если значение f больше всех собственных значений матриц $\mathbf{A}$ и $\mathbf{B}$, то $\mathbf{A}^{f}=\mathbf{A}$ и $\mathbf{B}^{f}=\mathbf{B}$ и справедливость (12) очевидна. С другой стороны, если значение f меньше всех собственных значений матриц $\mathbf{A}$ и $\mathbf{B}$, то $\mathbf{A}^{f}=\mathbf{B}^{f}=f\,.\,\mathbf{E}_n$, где $\mathbf{E}_n$ – единичная матрица, и справедливость (12) также очевидна. Эти случаи мы дальше не будем рассматривать. Положим $\mathbf{A}^{(0)}=\mathbf{A}$ и $\mathbf{B}^{(0)}=\mathbf{B}$. Если на k-ом шагу все собственные значения матриц $\mathbf{A}^{(k)}$ и $\mathbf{B}^{(k)}$ не превосходят значение f, то мы уже получили срезку матриц $\mathbf{A}^{f}=\mathbf{A}^{(k)}$ и $\mathbf{B}^{f}=\mathbf{B}^{(k)}$. Процесс останавливается. Если же на k-ом шагу существуют собственные значения, которые больше значения f, тогда возмем наибольшие из собственных значений матриц $\mathbf{A}^{(k)}$ и $\mathbf{B}^{(k)}$ и уменьшим их либо до ближайщего значения матриц $\mathbf{A}^{(k)}$ и $\mathbf{B}^{(k)}$, либо до значения f (смотря которое из них больше), получая таким образом матрицы $\mathbf{A}^{(k+1)}$ и $\mathbf{B}^{(k+1)}$. Сделав не более чем $2n-1$ таких преобразований, мы получим матрицы $\mathbf{A}^{f}=\mathbf{A}^{(k_p)}$ и $\mathbf{B}^{f}=\mathbf{B}^{(k_p)}$. Покажем, что

$$\left\|\mathbf{A}^{(k+1)}-\mathbf{B}^{(k+1)}\right\|_E\leqq\left\|\mathbf{A}^{(k)}-\mathbf{B}^{(k)}\right\|_E\,,$$

откуда следует

$$\left\|\mathbf{A}^{f}-\mathbf{B}^{f}\right\|_E=\left\|\mathbf{A}^{(k_p)}-\mathbf{B}^{(k_p)}\right\|_E\leqq\ldots\leqq\left\|\mathbf{A}^{(0)}-\mathbf{B}^{(0)}\right\|_E=\left\|\mathbf{A}-\mathbf{B}\right\|_B.$$

Пусть $\mathbf{A}^{(k)}$ и $\mathbf{B}^{(k)}$ имеют вид

$$\mathbf{A}^{(k)}=\alpha\,.\sum_{i=1}^{k_A}\mathbf{a}_i\mathbf{a}_i^T+\sum_{i=k_A+1}^{n}\alpha_i\mathbf{a}_i\mathbf{a}_i^T\,,\quad\alpha>\alpha_i\,,\tag{14}$$

$$\mathbf{B}^{(k)}=\alpha\,.\sum_{i=1}^{k_B}\mathbf{b}_i\mathbf{b}_i^T+\sum_{i=k_B+1}^{n}\beta_i\mathbf{b}_i\mathbf{b}_i^T\,,\quad\alpha>\beta_i\,,$$

где α – наибольшее собственное значение матриц $\mathbf{A}^{(k)}$ и $\mathbf{B}^{(k)}$, которое мы уменьшим до значения α', т.е. $\alpha>\alpha'\geqq\alpha_i$ и $\alpha'\geqq\beta_i$. Случай $k_A=0$ или $k_B=0$ не исключается. Тогда матрицы $\mathbf{A}^{(k+1)}$ и $\mathbf{B}^{(k+1)}$ также имеют вид (14), только вместо α стоит α'. Требуется доказать, что

$$\left\|\mathbf{A}^{(k)}-\mathbf{B}^{(k)}\right\|_E^2-\left\|\mathbf{A}^{(k+1)}-\mathbf{B}^{(k+1)}\right\|_E^2\geqq0\,.$$

Используя скалярное произведение матриц, мы получаем

$$\left\|\mathbf{A}^{(k)}-\mathbf{B}^{(k)}\right\|_E^2=\left(\mathbf{A}^{(k)}-\mathbf{B}^{(k)},\mathbf{A}^{(k)}-\mathbf{B}^{(k)}\right)=$$
$$=\operatorname{tr}\left(\mathbf{A}^{(k)^2}\right)+\operatorname{tr}\left(\mathbf{B}^{(k)^2}\right)-2\operatorname{tr}\left(\mathbf{A}^{(k)}\mathbf{B}^{(k)}\right)=$$

$$= (k_A + k_B)\,\alpha^2 + \sum_{i=k_A+1}^{n} \alpha_i^2 + \sum_{i=k_B+1}^{n} \beta_i^2 - 2\alpha^2 \,.\sum_{i=1}^{k_A} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2 -$$

$$- 2\alpha \,.\sum_{i=1}^{k_A} \sum_{j=k_B+1}^{n} \beta_j(\mathbf{a}_i, \mathbf{b}_j)^2 - 2\alpha \,.\sum_{i=k_A+1}^{n} \sum_{j=1}^{k_B} \alpha_i(\mathbf{a}_i, \mathbf{b}_j)^2 -$$

$$- 2 \,.\sum_{i=k_A+1}^{n} \sum_{j=k_B+1}^{n} \alpha_i\beta_j(\mathbf{a}_i, \mathbf{b}_j)^2 \,.$$

Аналогично и для матриц $\mathbf{A}^{(k+1)}$ и $\mathbf{B}^{(k+1)}$. Далее мы получаем

(15)
$$\|\mathbf{A}^{(k)} - \mathbf{B}^{(k)}\|_E^2 - \|\mathbf{A}^{(k+1)} - \mathbf{B}^{(k+1)}\|_E^2 =$$

$$= (k_A + k_B)(\alpha^2 - \alpha'^2) - 2(\alpha^2 - \alpha'^2) \,.\sum_{i=1}^{k_A} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2 -$$

$$- 2(\alpha - \alpha')\,\Big[\sum_{i=1}^{k_A} \sum_{j=k_B+1}^{n} \beta_j(\mathbf{a}_i, \mathbf{b}_j)^2 + \sum_{i=k_A+1}^{n} \sum_{j=1}^{k_B} \alpha_i(\mathbf{a}_i, \mathbf{b}_j)^2\Big] \geqq$$

$\geqq$ (используя неравенства $\alpha' \geqq \alpha_i$ и $\alpha' \geqq \beta_j$) $\geqq$

$$\geqq (\alpha - \alpha')\,\{(k_A + k_B)(\alpha + \alpha') - 2(\alpha + \alpha')\sum_{i=1}^{k_A} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2 -$$

$$- 2\alpha'\Big[\sum_{i=1}^{k_A} \sum_{j=k_B+1}^{n} (\mathbf{a}_i, \mathbf{b}_j)^2 + \sum_{i=k_A+1}^{n} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2\Big]\} =$$

$$= (\alpha - \alpha')\,\{(k_A + k_B)(\alpha + \alpha') - 2(\alpha - \alpha')\sum_{i=1}^{k_A} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2 -$$

$$- 4\alpha' \sum_{i=1}^{k_A} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2 -$$

$$- 2\alpha'\Big[\sum_{i=1}^{k_A} \sum_{j=k_B+1}^{n} (\mathbf{a}_i, \mathbf{b}_j)^2 + \sum_{i=k_A+1}^{n} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2\Big]\} =$$

$$= (\alpha - \alpha')\,\{(k_A + k_B)(\alpha + \alpha') - 2(\alpha - \alpha')\sum_{i=1}^{k_A} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2 -$$

$$- 2\alpha'\Big[\sum_{i=1}^{k_A} \sum_{j=1}^{n} (\mathbf{a}_i, \mathbf{b}_j)^2 + \sum_{i=1}^{n} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2\Big]\} =$$

$$= \Big/ \sum_{i=1}^{n} (\mathbf{a}_i, \mathbf{b}_j)^2 = \sum_{j=1}^{n} (\mathbf{a}_i, \mathbf{b}_j)^2 = 1 \Big/ =$$

$$= (\alpha - \alpha') \,.\, \Big[(k_A + k_B) \,.\, (\alpha + \alpha' - 2\alpha') - 2(\alpha - \alpha')\sum_{i=1}^{k_A} \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2\Big] \geqq$$

$$\geqq \Big/\text{так как } \sum_{i=1}^{k_A} (\mathbf{a}_i, \mathbf{b}_j)^2 \leqq 1 \text{ и } \sum_{j=1}^{k_B} (\mathbf{a}_i, \mathbf{b}_j)^2 \leqq 1 \Big/ \geqq$$

$$\geqq (\alpha - \alpha')^2\,\Big[(k_A + k_B) - 2\min(k_A, k_B)\Big] \geqq 0 \,.$$

Доказанное неравенство (15) вместе с проведенными выше рассуждениями установливает справедливость леммы 1.

Следствие. *Пусть* $\mathbf{A}$ *и* $\mathbf{B}$ *— вещественные симметричные матрицы размера* n*. Тогда их положительные (отрицательные) части отличаются в евклидовой норме друг от друга не более чем сами матрицы* $\mathbf{A}$ *и* $\mathbf{B}$.

Доказательство. Если через $\mathbf{A}_-$ и $\mathbf{A}_+$ обозначить соответственно отрицательную и положительную части матрицы $\mathbf{A}$, то достаточно заметить, что $\mathbf{A}_- = \mathbf{A}^{(0)}$ (но не в смысле (7) и (8)!) и $\mathbf{A}_+ = \mathbf{A}_0$.

Далее предварительно докажем, что теорема 2 справедлива для неотрицательно определенных матриц, как утверждается в [5].

Лемма 2. *Пусть* $\mathbf{A}$ *и* $\mathbf{B}$ *— вещественные неотрицательно определенные симметричные матрицы размера* n*. Тогда*

$$\|\mathbf{A}^0 - \mathbf{B}^0\|_E \leqq 2\|\mathbf{A} - \mathbf{B}\|_E \cdot \ell^{-2},$$

где $\mathbf{A}^0$ *и* $\mathbf{B}^0$ *определены соотношениями* (7) *и* (8).

Доказательство. Заметим сначала, что

$$\mathbf{A}^0 + \ell^{-1} \,.\, \mathbf{E}_n = (\mathbf{A}_\ell)^+ + (\mathbf{A}^\ell)^T \,.\, \ell^{-2};$$

$$\mathbf{B}^0 + \ell^{-1} \,.\, \mathbf{E}_n = (\mathbf{B}_\ell)^+ + (\mathbf{B}^\ell)^T \,.\, \ell^{-2}.$$

Это легко показать с использованием равенства $\mathbf{E}_n = \sum_{i=1}^{n} \mathbf{a}_i\mathbf{a}_i^T = \sum_{i=1}^{n} \mathbf{b}_i\mathbf{b}_i^T$. Матрицы $\mathbf{A}^\ell$, $\mathbf{A}_\ell$, $\mathbf{B}^\ell$, $\mathbf{B}_\ell$ — это срезки матриц $\mathbf{A}$ и $\mathbf{B}$. Важно что $\operatorname{rank} \mathbf{A}_\ell = \operatorname{rank} \mathbf{B}_\ell = n$ и поэтому значение L в утверждении 3 равно 1. Отметим также, что $\|\mathbf{A}_\ell^+\|_2 \leqq \ell^{-1}$ и $\|\mathbf{B}_\ell^+\|_2 \leqq \ell^{-1}$. Итак, мы имеем

$$\begin{aligned}
&\|\mathbf{A}^0 - \mathbf{B}^0\|_E = \|\mathbf{A}^0 + \ell^{-1} \,.\, \mathbf{E}_n - \mathbf{B}^0 - \ell^{-1} \,.\, \mathbf{E}_n\|_E = \\
&= \|(\mathbf{A}_\ell)^+ + (\mathbf{A}^\ell)^T \,.\, \ell^{-2} - (\mathbf{B}_\ell)^+ - (\mathbf{B}^\ell)^T \,.\, \ell^{-2}\|_E \leqq \\
&\leqq \|(\mathbf{A}_\ell)^+ - (\mathbf{B}_\ell)^+\|_E + \|(\mathbf{A}^\ell - \mathbf{B}^\ell)^T\|_E \cdot \ell^{-2} \leqq \\
&\leqq (\text{используя утверждение 3}) \leqq \\
&\leqq \|\mathbf{A}_\ell - \mathbf{B}_\ell\|_E \cdot \|\mathbf{A}_\ell^+\|_2 \cdot \|\mathbf{B}_\ell^+\|_2 + \|\mathbf{A}^\ell - \mathbf{B}^\ell\|_E \cdot \ell^{-2} \leqq \\
&\leqq (\text{используя лемму 1}) \leqq 2\|\mathbf{A} - \mathbf{B}\|_E \cdot \ell^{-2}.
\end{aligned}$$

Лемма 2 доказана.

Пусть $\mathbf{A}$ — симметричная вещественная матрица размера n. Матрица $\mathbf{A}$ имеет разложение

$$\mathbf{A} = \sum_{i=1}^{n} \alpha_i \mathbf{a}_i \mathbf{a}_i^T = \sum_{i=1}^{n} |\alpha_i| \,.\, (\operatorname{sgn} \alpha_i \,.\, \mathbf{a}_i) \,.\, \mathbf{a}_i^T.$$

Видно, что это есть сингулярное разложение матрицы размера $n \times n$. Поэтому, если для $\mathbf{A}$ определить на основе соотношений (7) и (8) матрицу $\mathbf{A}^0$, мы получим для нее следующее разложение:

$$\mathbf{A}^0 = \sum_{|\alpha_i| > \ell} \alpha_i^{-1} \mathbf{a}_i \mathbf{a}_i^T + \sum_{|\alpha_i| \leqq \ell} \alpha_i \,.\, \ell^{-2} \mathbf{a}_i \mathbf{a}_i^T \,.$$

Очевидно, что справедливы следующие равенства:

$$\mathbf{A}^0 = (\mathbf{A}^0)_- + (\mathbf{A}^0)_+ = (\mathbf{A}_-)^0 + (\mathbf{A}_+)^0 \,.$$

Перейдем к доказательству теоремы 2.

Рассмотрим далее матрицы $\mathbf{A}$ и $\mathbf{B} \in \mathcal{M}_{m \times n}$ и положим

$$\mathbf{S}_A = \begin{bmatrix} 0 & \mathbf{A} \\ \mathbf{A}^T & 0 \end{bmatrix}$$

$\mathbf{S}_A$ – симметричная матрица размера $(m + n)$ и из утверждения 1 с помощью несложных преобразований легко следует, что $(\mathbf{S}_A)^0 = \mathbf{S}_{(A^0)^T}$. Потом справедлива следующая цепочка неравенств:

$$\|\mathbf{A}^0 - \mathbf{B}^0\|_E^2 = \|(\mathbf{A}^0)^T - (\mathbf{B}^0)^T\|_E^2 =$$

$$\frac{1}{2} \left\| \begin{bmatrix} 0 & (\mathbf{A}^0)^T \\ \mathbf{A}^0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & (\mathbf{B}^0)^T \\ \mathbf{B}^0 & 0 \end{bmatrix} \right\|_E^2 =$$

$$= \tfrac{1}{2} \|\mathbf{S}_{(A^0)^T} - \mathbf{S}_{(B^0)^T}\|_E^2 = \tfrac{1}{2} \|(\mathbf{S}_A)^0 - (\mathbf{S}_B)^0\|_E^2 =$$

$$= \tfrac{1}{2} \|(\mathbf{S}_A)_+^0 + (\mathbf{S}_A)_-^0 - (\mathbf{S}_B)_+^0 - (\mathbf{S}_B)_-^0\|_E^2 =$$

$$= \tfrac{1}{2} \|((\mathbf{S}_A)_+)^0 - ((\mathbf{S}_B)_+)^0 + ((\mathbf{S}_A)_-)^0 - ((\mathbf{S}_B)_-)^0\|_E^2 \leqq$$

$$\leqq /(a + b)^2 \leqq 2(a^2 + b^2)/ \leqq$$

$$\leqq \|(\mathbf{S}_{A+})^0 - (\mathbf{S}_{B+})^0\|_E^2 + \|(\mathbf{S}_{A-})^0 - (\mathbf{S}_{B-})^0\|_E^2 \leqq$$

$$\leqq \text{(на основе леммы 2)} \leqq$$

$$\leqq 2^2 \|\mathbf{S}_{A+} - \mathbf{S}_{B+}\|_E^2 \,.\, \ell^{-4} + 2^2 \,.\, \|\mathbf{S}_{A-} - \mathbf{S}_{B-}\|_E^2 \,.\, \ell^{-4} \leqq$$

$$\leqq \text{(на основе леммы 1)} \leqq$$

$$\leqq 8 \,.\, \|\mathbf{S}_A - \mathbf{S}_B\|_E^2 \,.\, \ell^{-4} = 16 \,.\, \|\mathbf{A} - \mathbf{B}\|_E^2 \,.\, \ell^{-4} \,.$$

Итак, мы получили неравенство

$$\|\mathbf{A}^0 - \mathbf{B}^0\|_E \leqq 4 \,.\, \|\mathbf{A} - \mathbf{B}\|_E \,.\, \ell^{-2} \,.$$

Справедливость теоремы 2 вытекает из указанного неравенства и факта эквивалентности матричных норм.

4. В заключение заметим, что в силу (11) предложенный метод в общем случае обладает некоторыми преимуществами по сравнению с предложенными раньше методами. Они могут проявиться, например, при решении задач, для которых точная СЛАУ не разрешима. Однако метод может быть реализован лишь в случае небольшой (порядка ста) размерности, так как он использует сингулярное разложение матрицы. Для этого можно использовать библиотечную программу разложения, например известную подпрограмму **SVD**. По видимому, можно искать модификации уже предложенных методов, которые можно было бы использовать для больших размерностей и которые обладали бы свойством типа отношения (11).

Численная реализация предложенного метода: Шаг 1. Используя программу **SVD**, вычисляем сингулярное разложение

$$\mathbf{A} = \mathbf{U}\Lambda\mathbf{V}^T$$

Шаг 2. По заданным μ и δ, выбирая параметр $\alpha \in (0, \frac{1}{2})$, вычисляем значение $f(\mu, \delta) = \max^\alpha(\mu, \delta)$.

Шаг 3. Вычисляем матрицу Λ^0. Для этого обращаем те стоящие на диагонали Λ сингулярные числа, которые больше значения f. Для остальных положаем $\alpha_i^0 = \alpha_i \cdot f^{-2}$.

Шаг. 4 Вычисляем $\mathbf{z}_f(\mathbf{A}, \mathbf{u}) = \mathbf{V}\Lambda^0\mathbf{U}^T\mathbf{u}$, так как $\mathbf{A}^0 = \mathbf{V}\Lambda^0\mathbf{U}^T$ (см. (5)).

Литература

[1] *Тихонов А. Н., Арсенин В. Я.:* Методы решения некорректных задач. М.: Наука, 1979.

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

[3] *Лоусон Ч., Хенсон Р.:* Численное решение задач метода наименьших квадратов. М.: Наука, 1986 (перевод с английского:
Lawson Ch. L., Hanson R. I.: Solving least squares problems. New Jersey. Prentice-Hall: Englewood Cliffs, Inc., 1974, 340p.)

[4] *Wedin P.-A.:* Perturbation theory for psuedo-inverses. BIT, 13 (1973), pp. 217—232.

[5] *Буша Я.:* Об одном методе регуляризации решения систем линейных алгебраических уравнений. ДАН СССР, т. 295 (1987), Но 1, с. 11—14.

Súhrn

O JEDNEJ METÓDE REGULARIZÁCIE RIEŠENIA SÚSTAV LINEÁRNYCH ALGEBRAICKÝCH ROVNÍC

JÁN BUŠA

V práci je navrhnutá a zdôvodnená nová metóda stabilizácie riešenia sústavy lineárnych algebraických rovníc zadanej približne. Je dokázaná stabilita metódy a tzv. globálna stabilita metódy. Stručne je popísaná numerická realizácia metódy.

Summary

ON A REGULARIZATION METHOD FOR SOLVING OF SYSTEMS OF LINEAR ALGEBRAIC EQUATIONS

JÁN BUŠA

In this paper, a new regularization method for normal pseudo-solution of perturbed systems of linear algebraic equations is suggested. It is proved the stability and the global stability of this method. Short prescribe of the numerical realization of this method is given.

Адрес автора: RNDr. *Ján Buša,* CSc., Katedra matematiky EF VŠT, Zbrojnícka 3, 042 00 Košice.