Доклады Академии наук СССР
1965. Том 163, № 3

МАТЕМАТИКА

Член-корреспондент АН СССР А. Н. ТИХОНОВ

О НЕКОРРЕКТНЫХ ЗАДАЧАХ ЛИНЕЙНОЙ АЛГЕБРЫ И УСТОЙЧИВОМ МЕТОДЕ ИХ РЕШЕНИЯ

1. Рассмотрим систему линейных алгебраических уравнений

$$Az = \bar{u}, \quad A = \{a_{ij}\}, \quad z = \{z_j\}, \quad u = \{u_i\},$$
$$i = 1, \ldots, m; \; j = 1, \ldots, n. \tag{1}$$

Подобная система не всегда разрешима. Обозначим $u_A = \{u: u = Az, z \in R_n\} \subset U = R_m$, $N_A = \{z: Az = 0\}$ и Z_A — ортогональное дополнение к N_A в пространстве $Z = R_n$. Очевидно, что система (1) разрешима в том и только в том случае, если правая часть уравнения $\bar{u} \in U_A$. Если $N_A \equiv 0$, то решение системы определенно однозначно. Если же $N_A \not\equiv 0$, то решение z системы (1) неоднозначно: если $z^{(0)}$ — какое-либо ее решение, то $\{\bar{z}\} = \{\bar{z}^0 + z: z \in N_A\}$ представляет полную совокупность решений системы (1). Совокупность условий

$$A\bar{z} \doteq \bar{u}, \quad (\bar{z} \cdot e_k) = 0,$$

где $\{e_k\}$ — базис N_A, однозначно определяет элемент $\bar{z}^{(0)}$. Будем называть $\bar{z}^{(0)}$ **нормальным решением системы** (1). Если система невырождена, то нормальное решение совпадает с единственным решением этой системы.

Отметим для дальнейшего, что нормальное решение может быть определено также из условий

$$A\bar{z}^{(0)} = \bar{u}, \quad \| \bar{z}^{(0)} \| = \min \| \bar{z} \|, \text{ если } A\bar{z} = \bar{u}.$$

Целью настоящей статьи является построение устойчивого алгоритма для определения нормального решения системы (1).

2. Пусть исходные данные системы (1), т. е. матрица A и вектор $\bar{u}$, задаются с некоторым приближением, причем меру погрешности A, u и z будем определять при помощи норм

$$\| u \| = \Big(\sum_{i=1}^{m} u_i^2\Big)^{1/2}, \quad \| z \| = \Big(\sum_{j=1}^{n} z_j^2\Big)^{1/2}, \quad \| A \| = \Big(\sum (a_{ij}^2)\Big)^{1/2}. \tag{2}$$

Нетрудно убедиться, что задача определения нормального решения системы (1) некорректна в смысле Адамара.

Рассмотрим пример:

$$(x - y) = u_1, \quad (y - z) = u_2, \quad \xi x + \eta y + \zeta z = u_3.$$

Детерминант этой системы A и значение неизвестной x равны

$$\Delta = (\xi + \eta + \zeta), \quad x = \frac{u_3 + \zeta u_2 + (\zeta + \eta) u_3}{\xi + \eta + \zeta} = \frac{X}{\Delta}.$$

Если система вырождена, то $\xi + \eta + \zeta = 0$, и обращение X в нуль является одним из условий разрешимости. Пусть, однако, ξ, η, ζ иррациональны и при вычислениях с некоторым числом десятичных знаков задаются как ξ_n, η_n, ζ_n. Имеем, что Δ_n и X_n вообще отличны от нуля и значение x_n, как отношение малых чисел, может принимать любое значение. Естественно, что безнадежны попытки «уточнить» значение x за счет увеличения числа десятичных знаков.

Аналогично обстоит дело для любой вырожденной системы.

Таким образом, решения систем, представляющих как угодно близкие аппроксимации вырожденных систем, дают большой разброс, а задача об определении нормальных решений вырожденных систем некорректна в смысле Адамара.

3. Если в системе (1) $m > n$, но система может допускать лишь единственное решение ($N_A \equiv 0$), то метод наименьших квадратов решает эту задачу, причем он приведет к определенному значению $\bar{z}$ для любого $\bar{u}$ независимо от условия разрешимости. Рассмотрим, что представляет $\bar{z}$, если $\bar{u}$ не удовлетворяет условиям разрешимости. Множество u_A образует линейное пространство. Обозначим U_A проекцию $\bar{u}$ на U_A. Очевидно, что решение, определяемое методом наименьших квадратов, удовлетворяет уравнению

$$A\bar{z} = \hat{\bar{u}}_A.$$

В самом деле, метод наименьших квадратов состоит в том, что $\bar{z}$ определяется как элемент, реализующий минимум квадратической формы

$$N[z, \bar{u}] = \|Az - \bar{u}\|^2.$$

В нашем случае

$$N[z, \bar{u}] = \|Az - \hat{\bar{u}}_A\|^2 + \|\hat{\bar{u}}_A - \bar{u}\|^2,$$

так как $\hat{\bar{u}}_A$ — ортогональная проекция $\bar{u}$ на U_A и, очевидно, минимум этой системы реализует элемент $\bar{z}$, удовлетворяющий уравнению $A\bar{z} = \hat{\bar{u}}_A$. Однако, если матрица A вырожденная ($N_A \neq 0$), то $\bar{z}$ определен неоднозначно.

4. Рассмотрим параметрический функционал

$$M^\alpha[z, \tilde{A}, \tilde{u}] = \|\tilde{A}z - \tilde{u}\|^2 + \alpha\Omega[z], \qquad \Omega[z) = \|z\|^2,$$

где $\tilde{A}$, $\tilde{u}$ — произвольные матрица и вектор, $\alpha > 0$ — параметр. Нетрудно видеть, что при любых $\tilde{A}$, $\tilde{u}$ и $\alpha > 0$ существует единственный элемент z^α, реализующий минимум этого функционала.

Т е о р е м а. Пусть A — *матрица;* $\bar{u}$ — *вектор, удовлетворяющий условию разрешимости уравнения* $Az = \bar{u}$; $\bar{z}^{(0)}$ — *нормальное решение. Пусть* $\tilde{A}$, $\tilde{u}$ — *какие-либо* δ*-приближения* A; $\bar{u}$, $\varepsilon(\delta)$, $\alpha(\delta)$ — *какие-либо убывающие функции* δ, *стремящиеся к нулю при* $\delta \to 0$ *и такие, что* $\delta^2 \leqslant \varepsilon(\delta)\alpha(\delta)$.

Каково бы ни было $\varepsilon > 0$, *существует такое* $\delta_0 = (\varepsilon, \|\bar{z}^0\|)$, *что вектор* z^α, *реализующий минимум функционала*

$$M^\alpha[z, \tilde{A}, \tilde{u}] = \|\tilde{A}z - \tilde{u}\|^2 + \alpha\Omega[z],$$

где α — *любое число такое, что*

$$\frac{1}{\varepsilon(\delta)}\delta^2 \leqslant \alpha \leqslant \alpha_0(\delta), \tag{α^0}$$

удовлетворяет неравенству

$$\|z^\alpha - \bar{z}^{(0)}\| \leqslant \varepsilon,$$

если только $\delta \leqslant \alpha_0(\varepsilon, \|\bar{z}^{(0)}\|)$.

Эта теорема имеет место как для вырожденных, так и для невырожденных матриц.

Обозначим $\hat{\tilde{u}}_{\tilde{A}}$ проекцию $\tilde{u}$ на линейное пространство

$$U_A = \{u = \tilde{A}z, z \in Z = R_n\}.$$

В этом случае

$$\|\tilde{u} - \hat{\tilde{u}}_{\tilde{A}}\| \leqslant \|\tilde{u} - \tilde{A}z\| \quad (z \in Z = R_n),$$

$$\|\tilde{A}z - \tilde{u}\|^2 \leqslant \|\tilde{A}z - \hat{\tilde{u}}_{\tilde{A}}\|^2 + \|\hat{\tilde{u}}_{\tilde{A}} - \tilde{u}\|^2.$$

Таким образом,

$$M^\alpha[z, \tilde{A}, \tilde{u}] = \|\tilde{u} - \hat{\tilde{u}}_{\tilde{A}}\|^2 + M^\alpha[z, \tilde{A}, \hat{\tilde{u}}_{\tilde{A}}],$$

и функционалы в правой и левой частях этого равенства имеют общий элемент z^α, их минимизирующий.

Воспользуемся неравенством

$$\alpha\Omega[z^\alpha] \leqslant M^\alpha[z^\alpha, \tilde{A}, \hat{\tilde{u}}_{\tilde{A}}] \leqslant M^\alpha[\bar{z}^{(0)}, \tilde{A}, \hat{\tilde{u}}_{\tilde{A}}] = \|\tilde{A}\bar{z}^{(0)} - \hat{\tilde{u}}_{\tilde{A}}\|^2 + \alpha\Omega[\bar{z}^{(0)}],$$

а также неравенством

$$\|\tilde{A}\bar{z}^{(0)} - \hat{\tilde{u}}_{\tilde{A}}\| \leqslant \|\tilde{A}\bar{z}^{(0)} - A\bar{z}^{(0)}\| + \|\bar{u} - \hat{\tilde{u}}_{\tilde{A}}\| \leqslant$$

$$\leqslant \delta\|\bar{z}^{(0)}\| + \|\bar{u} - \tilde{u}\| + \|\tilde{u} - \hat{\tilde{u}}_{\tilde{A}}\| \leqslant \delta(\|\bar{z}^{(0)}\| + 1) + \|\tilde{u} - \tilde{A}[\bar{z}^{(0)}]\| \leqslant$$

$$\leqslant \delta(\|\bar{z}^{(0)}\| + 1) + \{\|\tilde{u} - \bar{u}\| + \|A\bar{z}^{(0)} - \tilde{A}\bar{z}^{(0)}\|\} \leqslant C\delta, \quad C = 2(1 + \|\bar{z}^{(0)}\|).$$

Таким образом,

$$\alpha\Omega[z^\alpha] \leqslant M^\alpha[z^\alpha, \tilde{A}, \hat{\tilde{u}}_{\tilde{A}}] \leqslant \alpha\left(\frac{C^2\delta^2}{\alpha} + \Omega[\bar{z}^{(0)}]\right) \leqslant \alpha(C^2\varepsilon(\delta) + \Omega[\bar{z}^{(0)}]).$$

Из последнего неравенства следует, что

$$\|z^\alpha\| \leqslant \|\bar{z}^{(0)}\| + \varepsilon_1(\delta) \quad (\varepsilon_1(\delta) \to 0, \delta \to 0),$$

т. е. что $\{z^\alpha\}$ образует компактное множество.

Убедимся теперь в том, что для любого $\varepsilon > 0$ существует такое $\delta_0(\varepsilon, \|\bar{z}^0\|)$, что если $\|\tilde{A} - A\| < \delta$ и $\|\tilde{u} - \bar{u}\| < \delta$, то $\|z^\alpha - \bar{z}^{(0)}\| < \varepsilon$, если $\delta < \delta_0$ и α удовлетворяет условию (α^0).

Предположим, что это неверно и что существуют $\varepsilon_0 > 0$ и A_n, $\tilde{u}_n$ и $\delta_n \to 0$ такие, что $\|z^{\alpha_n} - \bar{z}^{(0)}\| \geqslant \varepsilon_0$. В силу компактности z^α можно без ограничения общности считать, что последовательность z^{α_n} сходится к некоторому элементу $z^{(0)}$.

Убедимся в том, что $z^{(0)} = \bar{z}^{(0)}$, что будет противоречить предположению. Оценим

$$\|Az^{\alpha_n} - A\bar{z}^{(0)}\| \leqslant \|Az^{\alpha_n} - \tilde{A}z^{\alpha_n}\| + \|\tilde{A}z^{\alpha_n} - \hat{\tilde{u}}_{\tilde{A}}\| + \|\hat{\tilde{u}}_{\tilde{A}} - A\bar{z}^{(0)}\| \leqslant$$

$$\leqslant \delta_n\|z^{\alpha_n}\| + \sqrt{M^{\alpha_n}[z^{\alpha_n}, \hat{\tilde{u}}_{\tilde{A}}, \tilde{A}]} + \delta_n(2 + \|z^{(0)}\|) \leqslant$$

$$\leqslant \delta_n\{2(1 + \|\bar{z}^{(0)}\|) + \varepsilon_1(\delta_n)\} + \sqrt{C^2\delta_n^2 + \alpha_0(\delta_n)\Omega[\bar{z}^{(0)}]} \underset{(n\to\infty)}{=} \eta_n \to 0.$$

Таким образом, для $z^{(0)} = \lim\limits_{n\to\infty} z^{\alpha_n}$ получаем

$$Az^{(0)} = A\bar{z}^{(0)}, \quad \|z^{(0)}\| \leqslant \|\bar{z}^{(0)}\|,$$

откуда следует, что $z^{(0)} = \bar{z}^{(0)}$, так как эти условия определяют единственный элемент, что и доказывает теорему.

Замечание 1. Если матрица A плохо обусловлена и в δ-окрестности ее имеется вырожденная матрица $\tilde{A}$, где δ — точность задания A, то мы находимся в условиях рассматриваемой задачи. Без регуляризации мы можем получить сильно различающиеся решения, и применение регуляризации будет давать приближение к нормальному решению уравнения $Az = \bar{u}$.

Замечание 2. Определим обобщенное нормальное решение $z^{(0)}$ условиям

$$A\bar{\bar{z}}^{(0)} = \bar{u}, \quad \Omega(\bar{\bar{z}}^{(0)} - z_0) \leqslant \Omega(z - z_0) \quad \text{для всех } z\colon L(z) = \bar{u},$$

где z_0 — произвольный фиксированный элемент, Ω — положительно определенная квадратическая форма.

Регуляризация с функционалом $\Omega[z] = \|z - z_0\|^2$ проходит дословно подобно предшествующему и определяет обобщенное нормальное решение $\overline{\overline{z}}^{(0)}$.

Замечание 3. Проведенное исследование не связано с конечномерностью пространств z и u и дословно повторяется для произвольных непрерывных линейных операторов $A[z]$, если U — гильбертово пространство и Z — нормированное пространство, в которое s-компактно вложено гильбертово пространство $\overline{Z}$ ([2]). Это дает метод регуляризации решения линейных неоднородных уравнений на спектре ([1]).

Поступило
20 IV 1964

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] А. Н. Тихонов, ДАН, **151**, № 3 (1963); **153**, № 1 (1963). [2] А. Н. Тихонов, ДАН, **161**, № 5 (1965).