Math-Net.Ru

Общероссийский математический портал

А. И. Жданов, Об одном численно устойчивом алгоритме решения систем линейных алгебраических уравнений неполного ранга, *Вестн. Сам. гос. техн. ун-та. Сер. Физ.-мат. науки*, 2008, выпуск 1(), 149–153

DOI: <https://doi.org/10.14498/vsgtu588>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 107.22.136.117

26 апреля 2023 г., 03:35:55

Вычислительная математика

УДК 519.612

А. И. Жданов

ОБ ОДНОМ ЧИСЛЕННО УСТОЙЧИВОМ АЛГОРИТМЕ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ НЕПОЛНОГО РАНГА

Рассматривается новый метод решения неустойчивых задач, сводящихся к решению произвольных (возможно, с неполным рангом и несовместных) систем линейных алгебраических уравнений. Данный метод основан на преобразовании регуляризованной нормальной системы уравнений к эквивалентной расширенной регуляризованной нормальной системе уравнений.

1. Постановка задачи. Рассмотрим произвольную (возможно неполного ранга и несовместную) систему линейных алгебраических уравнений (СЛАУ)

$$Au = f, \quad A \in \mathbb{R}^{m \times n}, \quad f \in \mathbb{R}^m. \quad (1)$$

Под «решением» СЛАУ (1) в общем случае будет пониматься нормальное псевдорешение $u_* = A^+ f$, где A^+ — псевдообратная матрица.

Пусть $\tilde{u}_*$ — вычисленные любым известным прямым или итерационным методом [1, 2] на компьютере (в арифметике с плавающей точкой) решения u_* . Эти решения различаются по точности. Оценки погрешностей $\|\tilde{u}_* - u_*\|_2$ компьютерных решений $\tilde{u}_*$ для этих методов зависят от чисел обусловленности матриц A . Спектральное число обусловленности СЛАУ (1) очевидно определяется как $\kappa_2(A) = \sigma_1/\sigma_\tau$, где $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_k$ — сингулярные числа матрицы A , $k = \min(m, n)$, $\tau = \text{rank}(A) \leq k$.

Если A — хорошо обусловленная матрица, т.е. её число обусловленности $\kappa_2(A) \ll 1/\varepsilon_{\text{mach}}$, где $\varepsilon_{\text{mach}}$ — машинное эпсилон [4], то компьютерные решения $\tilde{u}_*$, вычисленные любыми известными методами [4], по точности сравнимы между собой. В противном случае, известно [1, 2], что значительная часть имеющихся рекомендаций по решению плохо обусловленных ($\kappa_2(A) \approx 1/\varepsilon_{\text{mach}}$) и неполного ранга СЛАУ ($\tau < k$) требует нахождения сингулярного разложения (SVD-разложения) матрицы системы или действий, по существу, к этому сводящихся. Это связано с тем, что методы, основанные на SVD-разложении, являются непревзойдёнными (по точности [1]) методами решения этого класса задач. Однако методы, основанные на SVD-разложении, не позволяют эффективно решать СЛАУ большой размерности с разрежёнными матрицами, так как для их применения требуется большой объем оперативной памяти компьютера.

В настоящей работе предлагается новый метод, основанный на регуляризации СЛАУ (1) с использованием эквивалентных расширенных систем. Показано, что этот метод является обратно устойчивым [3] и конкурентно способным (по точности) методам, основанным на SVD-разложении. Кроме того, предлагаемый метод позволяет эффективно использовать математическое обеспечение, созданное для решения больших разрежённых систем на высокопроизводительных параллельных вычислительных системах.

2. Регуляризация на основе расширенных систем. Метод регуляризации А.Н. Тихонова может использоваться для приближенного нахождения нормального псевдорешения u_* СЛАУ (1). Именно, в качестве приближения к u_* берётся решение регуляризованной нормальной системы

$$(A^T A + \alpha I_n)u = A^T f \quad (2)$$

с подходящим значением параметра регуляризации $\alpha > 0$, где T — знак транспонирования, I_n — единичная матрица порядка n . Этот подход основан на известном факте [4], что при $\alpha \rightarrow 0$ решение u_α системы (2) стремится к нормальному псевдорешению u_* СЛАУ (1). К сожалению, в случае плохо обусловленных или неполного ранга матриц A , при достаточно малых значениях α ($\alpha \approx \varepsilon_{\text{mach}}$) СЛАУ (2) оказывается очень плохо обусловленной и, как следствие, не может быть решена никаким компьютерным алгоритмом.

В настоящей работе регуляризованная нормальная система (2) преобразуется к эквивалентной ей расширенной системе, которая имеет существенно меньшее число обусловленности и для неё может быть получено устойчивое компьютерное решение.

Регуляризованную нормальную систему уравнений (2) можно записать в виде

$$A^T r - \alpha u = 0,$$

где $r = f - Au$, или

$$\alpha^{-1/2} A^T r - \alpha^{1/2} u = 0. \quad (3)$$

Объединяя $r + Au = f$ и (3), получаем систему:

$$\begin{aligned} r + Au &= f, \\ \alpha^{-1/2} A^T r - \alpha^{1/2} u &= 0 \end{aligned}$$

или

$$\begin{pmatrix} \alpha^{1/2} I_m & A \\ A^T & -\alpha^{1/2} I_n \end{pmatrix} \begin{pmatrix} \alpha^{-1/2} r \\ u \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}. \quad (4)$$

Если обозначить $y = \alpha^{-1/2} r$ и $\omega = \alpha^{1/2}$, то (4) можно записать в виде

$$\begin{pmatrix} \omega I_m & A \\ A^T & -\omega I_n \end{pmatrix} \begin{pmatrix} y \\ u \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix} \iff \tilde{A}_\omega x = \tilde{f}, \quad (5)$$

где

$$\tilde{A}_\omega = \begin{pmatrix} \omega I_m & A \\ A^T & -\omega I_n \end{pmatrix}, \quad x = \begin{pmatrix} y \\ u \end{pmatrix} \quad \text{и} \quad \tilde{f} = \begin{pmatrix} f \\ 0 \end{pmatrix}.$$

СЛАУ (5) является расширенной системой с квадратной матрицей $\tilde{A}_\omega$ порядка $m + n$. Так как $\alpha > 0$, то и $\omega > 0$. Покажем, что матрица $\tilde{A}_\omega$ – невырожденная и, следовательно, расширенная СЛАУ всегда имеет единственное решение $z_\alpha = (y_\alpha^T, u_\alpha^T)^T$, совпадающее с решением регуляризованной нормальной системы (2)

$$u_\alpha = (A^T A + \alpha I_n)^{-1} A^T f,$$

и $y_\alpha = \alpha^{-1/2} r_\alpha$, а $r_\alpha = f - Au_\alpha$.

Для этого сформулируем следующую теорему, которая также дает возможность оценить число обусловленности вычислительной задачи (5).

ТЕОРЕМА. Если $A = O_{m \times n}$ – нулевая $m \times n$ -матрица, то собственными значениями матрицы $\tilde{A}_\omega$ являются числа ω и $-\omega$ кратностей m и n соответственно.

Если $A \neq O_{m \times n}$, то собственными значениями матрицы $\tilde{A}_\omega$ являются числа $\pm \sqrt{\sigma_i^2 + \omega^2}$ ($i = 1, 2, \dots, \tau$), а также ω и $-\omega$ кратностей $m - \tau$ и $n - \tau$ соответственно, где $\tau = \text{rank}(A) \geq 1$.

Доказательство. Так как $\tilde{A}_\omega = \tilde{A}_\omega^T$, то все её собственные значения $\lambda_i(\tilde{A}_\omega) \in \mathbb{R}$ ($i = 1, 2, \dots, m + n$).

Если $A = O_{m \times n}$, тогда матрица $\tilde{A}_\omega$ – диагональная, т. е.

$$\tilde{A}_\omega = \text{diag}(\underbrace{\omega, \dots, \omega}_m, \underbrace{-\omega, \dots, -\omega}_n),$$

откуда непосредственно следует первое утверждение теоремы.

Пусть $A \neq O_{m \times n}$ и

$$\begin{pmatrix} \omega I_m & A \\ A^T & -\omega I_n \end{pmatrix} \begin{pmatrix} z \\ v \end{pmatrix} = \lambda \begin{pmatrix} z \\ v \end{pmatrix},$$

где $z \in \mathbb{R}^m$, $v \in \mathbb{R}^n$. Тогда

$$\begin{aligned} \omega z + Av &= \lambda z, \\ A^T z - \omega v &= \lambda v. \end{aligned} \quad (6)$$

Если $(z^T, v^T)^T \neq 0$ и $\lambda \neq \omega$, то, исключая в (6) $z = (\lambda - \omega)^{-1} Av$, получаем

$$(\lambda - \omega)^{-1} A^T A v - \omega v = \lambda v$$

или

$$A^T A v = (\lambda - \omega)(\lambda + \omega)v \iff A^T A v = (\lambda^2 - \omega^2)v \iff (A^T A + \omega^2 I_n)v = \lambda^2 v. \quad (7)$$

Из (7) видно, что если $v \neq 0$, тогда v – собственный вектор и λ^2 соответствующее ему собственное значение матрицы $A^T A + \omega^2 I_n$, т.е. $\lambda_i^2 = \sigma_i^2 + \omega^2$ ($i = 1, 2, \dots, \tau$). Следовательно, числа $\pm\sqrt{\sigma_i^2 + \omega^2}$ ($i = 1, 2, \dots, \tau$) являются собственными значениями матрицы $\tilde{A}_\omega$.

Если $v = 0$, а $z \neq 0$, то из (7) следует, что $\omega z = \lambda z$ и $A^T z = 0$. Так как $\dim \ker A^T = m - \tau$, то ω является собственным значением кратности $m - \tau$ матрицы $\tilde{A}_\omega$.

Если $z = 0$, а $v \neq 0$, то из (7) следует, что $-\omega v = \lambda v$ и $Av = 0$. Так как $\dim \ker A = n - \tau$, то $-\omega$ является собственным значением кратности $n - \tau$ матрицы $\tilde{A}_\omega$. $\square$

Так как матрица $\tilde{A}_\omega$ – симметричная, то из теоремы непосредственно получаем

СЛЕДСТВИЕ. Если $m \neq n$ или $m = n > \tau$, то максимальное сингулярное число матрицы $\tilde{A}_\omega$ равно

$$\sigma_{\max}(\tilde{A}_\omega) = \sqrt{\sigma_1^2 + \omega^2},$$

а минимальное сингулярное число $\sigma_{\min}(\tilde{A}_\omega) = \omega$ и спектральное число обусловленности матрицы $\tilde{A}_\omega$:

$$\kappa_2(\tilde{A}_\omega) = \frac{\sigma_{\max}(\tilde{A}_\omega)}{\sigma_{\min}(\tilde{A}_\omega)} = \frac{\sqrt{\sigma_1^2 + \omega^2}}{\omega}.$$

Если $m = n = \tau$, то $\sigma_{\max}(\tilde{A}_\omega) = \sqrt{\sigma_1^2 + \omega^2}$, $\sigma_{\min}(\tilde{A}_\omega) = \sqrt{\sigma_n^2 + \omega^2}$, где $\sigma_n = \sigma_\tau$, а

$$\kappa_2(\tilde{A}_\omega) = \sqrt{\frac{\sigma_1^2 + \omega^2}{\sigma_n^2 + \omega^2}}.$$

Для регуляризованной нормальной системы (2) параметр регуляризации практически всегда должен удовлетворять условию $\alpha > \sigma_k$. Это условие необходимо, чтобы регуляризованное решение u_α было устойчивым решением. Таким образом, на основании следствия из теоремы можно получить оценку для числа обусловленности СЛАУ (5), независимую от ранга матрицы A ,

$$\kappa_2(\tilde{A}_\omega) \leq \frac{\sqrt{\sigma_1^2 + \omega^2}}{\omega}. \quad (8)$$

В силу трудоёмкости вычисления σ_1 , учитывая, что $\sigma_1 = \|A\|_2 \leq \|A\|_E$, где $\|A\|_E$ – евклидова матричная норма, вместо неравенства (8) на практике можно воспользоваться неравенством

$$\kappa_2(\tilde{A}_\omega) \leq \hat{\kappa}_2(\tilde{A}_\omega) = \frac{\sqrt{\|A\|_E^2 + \omega^2}}{\omega}.$$

3. Решение систем неполного ранга. Если выбрать параметр регуляризации α достаточно малым, то расширенную систему (5) можно использовать для приближенного вычисления численно устойчивых нормальных псевдорешений систем (1) неполного ранга или систем с матрицами A , имеющими неполный численный ранг [2, стр. 75]. В действительности, на практике чаще интересует ε -ранг матрицы, который для некоторого малого ε определяется по формуле

$$\text{rank}(A, \varepsilon) = \min_{\|A-B\|_2 \leq \varepsilon} \text{rank}(B).$$

Если $A \in \mathbb{R}^{m \times n}$ – матрица из чисел с плавающей точкой $A = \text{fl}(A)$, то разумно считать A численно неполноранговой, если $\text{rank}(A, \varepsilon) < \min(m, n)$, где $\varepsilon = \varepsilon_{\text{mach}} \|A\|_2$.

Для СЛАУ (1) неполного ранга ($\text{rank}(A) < k$) число обусловленности регуляризованной нормальной системы (2) равно

$$\kappa_2(A^T A + \omega^2) = \frac{\lambda_{\max}(A^T A + \omega^2)}{\lambda_{\min}(A^T A + \omega^2)} = \frac{\sigma_1^2 + \omega^2}{\omega^2},$$

а расширенной регуляризованной нормальной системы (РРНС) (5) —

$$\kappa_2(\tilde{A}_\omega) = \frac{\lambda_{\max}(\tilde{A}_\omega)}{\lambda_{\min}(\tilde{A}_\omega)} = \frac{\sqrt{\sigma_1^2 + \omega^2}}{\omega}.$$

Таким образом,

$$\kappa_2(A^T A + \omega^2 I_n) = \kappa_2^2(\tilde{A}_\omega).$$

Если выбрать параметр ω из условий: $\omega^2 < \varepsilon_{\text{mach}}$ и $\omega > \varepsilon_{\text{mach}}$, то очевидно, что

$$\Pi(A^T A + \omega^2 I_n) = \Pi(A^T A).$$

Следовательно, при таком значении ω компьютерное решение $\tilde{u}_\omega$ РРНС (5) будет достаточно точным приближением к u_* .

Как показано в [4, стр. 210], имеет место оценка

$$\|u_\omega - u_*\|_2 \leq \frac{\omega^2 \|f\|_2}{\sigma_\tau(\sigma_\tau + \omega^2)} \leq \frac{\omega^2 \|f\|_2}{\sigma_\tau^3} = \omega^2 q, \quad q = \text{const}.$$

Если для решения РРНС (5) применить любой обратно устойчивый алгоритм [5], например, основанный на QR-разложении Хаусхолдера [2], то как показано в [3]:

$$\frac{\|\tilde{x}_\omega - x_*\|_2}{\|x_*\|} = O(\kappa_2(\tilde{A}_\omega)\varepsilon_{\text{mach}}), \quad (9)$$

где $\tilde{x}_\omega$ — компьютерное решение РРНС (5), $x_* = (y_*^T, u_*^T)^T$, $y_* = \omega^{-1}(f - Au_*)$.

4. Тестовые исследования и выводы. В качестве иллюстрации возможностей предлагаемого метода рассмотрим следующий пример. В этом примере задача решалась двумя методами: SVD-методом из пакета Matlab и предлагаемым методом на основе РРНС. Для решения РРНС (5) (учитывая, что матрица $\tilde{A}_\omega$ является симметричной) был использован LDL^T -метод [3], также из пакета Matlab. В РРНС (5) параметр $\omega = 10^{-15}$.

Рассмотрим несовместную СЛАУ

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1.00000001 \\ 1 & 1.00000002 & 1 \end{pmatrix} \in \mathbb{R}^{4 \times 3}, \quad f = \begin{pmatrix} -94 \\ 106 \\ 6.00000003 \\ 6.00000004 \end{pmatrix} \in \mathbb{R}^4. \quad (10)$$

Точное псевдорешение СЛАУ (10) равно $u_* = (1, 2, 3)^T$. Решение вычисленное SVD-методом из пакета Matlab — $\tilde{u} = (1.431, 3.668, 1.400)^T$, а с использованием РРНС (5) — $\tilde{u}_\omega = (1.0000000, 1.9999999, 2.9999999)^T$.

Таким образом, предлагаемый метод на основе РРНС (при соответствующем выборе параметра ω) может даже превосходить по точности SVD-метод. При этом для СЛАУ с $m \approx n$ РРНС превосходит по быстродействию SVD-метод.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Деммель, Дж. Вычислительная линейная алгебра. Теория и алгоритмы [Текст] / Дж. Деммель. — М.: Мир, 2001. — 430 с.
2. Голуб, Дж. Матричные вычисления [Текст] / Дж. Голуб, Ч. Ван Лоун. — М.: Мир, 1999. — 548 с.
3. Trefethen, L. N. Numerical Linear Algebra [Text] / L. N. Trefethen, D. Bau. — SIAM, Philadelphia, 1997. — 373 p.
4. Беклемишев, Д. В. Дополнительные главы линейной алгебры [Текст] / Д. В. Беклемишев. — М.: Наука, 1983. — 336 с.
5. Жданов, А. И. Введение в методы решения некорректных задач. Часть 2: учеб. пособие [Текст] / А. И. Жданов. — Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2007. — 79 с.

Самарский государственный аэрокосмический университет
им. академика С.П. Королева, г. Самара
zhdanov@smr.ru

Поступила 25.09.2007

A. I. Zhdanov

ABOUT ONE NUMERICAL STABLE ALGORITHM FOR SOLVING SYSTEM LINEAR ALGEBRAIC EQUATIONS OF DEFECT RANK

A new method for solving unstable problems that can be reduced to arbitrary systems of linear algebraic equations (which may not be of full rank or may be inconsistent) is examined. This method is based on the reduction of regularization of normal system equations to an equivalent augmented regularization of normal system equations.

Samara State Aerospace University, Samara, Russia
zhdanov@smr.ru

Received 25.09.2007