1 Введение: задачи вычислительной линейной алгебры

1.1 Три фактора вычислений

Пусть $A \in Mat_{\mathbb{R}}(n)$ - вещественная $n \times n$ - матрица. Рассмотрим систему линейных алгебраических уравнений:

$$Ax = b, \quad x, b \in \mathbb{R}^n.$$

где вводятся следующие обозначения:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Далее мы будем исследовать свойства решения системы уравнений с учетом трех факторов:

1. конечность арифметики (u) - машинная точность;
2. проблема алгоритма (κ) - устойчивость;
3. проблема данных (ΔA) - обусловленность.

NB. 1. *С учетом перечисленных факторов систему можно записать в виде:*

$$(A + \Delta A)(x_0 + \Delta x) = b + \Delta b, \quad Ax_0 = b,$$

тогда оценка "погрешности" результата:

$$\Delta x = (A + \Delta A)^{-1}(\Delta b - \Delta A x_0 + \kappa).$$

1.2 Погрешность и машинная арифметика

Def 1. **Моделью арифметики с плавающей точкой** называется тройка $\langle F, G, fl \rangle_{\mathbb{R}}$, где

$$F = F(\beta, t, [L, U]) = \{x = \pm d_1 d_2 \dots d_t \cdot \beta^e \quad | \quad 0 \leqslant d_i < \beta, \quad d_1 \neq 0, \quad L \leqslant e \leqslant U\},$$
$$G = \{y \in \mathbb{R} : \quad m \leqslant |y| \leqslant M\} \cup \{0\}, \quad m = \beta^{L-1}, \quad M = \beta^U \cdot (1 - \beta^{-t}),$$
$$fl : G \to F, \quad fl(x) = x(1 + \delta) : \quad |\delta| \leqslant u.$$

Числа β, t и $[L, U]$ называются, соответственно, **базой**, **точностью** и **интервалом показателя**. Число u, равное расстоянию от 1 до ближайшего к ней большего числа, называется **единичной ошибкой округления** (или машинной точностью).

NB. 2. *(Аксиома машинной арифметики) Для любых двух чисел x и yс плавающей точкой и любой арифметической операции $\circ$ имеет место сравнение:*

$$fl(x \circ y) = (x \circ y)(1 + \delta), \quad |\delta| \leqslant u.$$

NB. 3. *Операция* fl *обладает свойством* $\mathrm{fl} \circ \mathrm{fl} = \mathrm{fl}$, *но не обладает другими свойствами, связанными со структурами множеств* $\mathbb{F}$ *и* $\mathbb{G}$*:*

$$\mathrm{fl}\left(\sum y_i\right) \neq \sum \mathrm{fl}(y_i).$$

Ex 1. *(Двойная точность в стандарте* IEEE*) Арифметике* $\mathbb{F}(2, 52, -1024, 1024)$

$$x: \quad \pm \Big[a_1 \quad a_2 \quad \dots a_{11} \quad | \quad b_1 \quad b_2 \quad \dots \quad b_{52} \Big] \quad \Rightarrow \quad x = \pm(1.b_1 b_2 \dots b_{52})_2 \times 2^{(a_1 a_2 \dots a_{11})_2 - 1023}.$$

Def 2. Договоримся число $\mathrm{fl}(x)$ называть **аппроксимацией** числа x и обозначать $\widehat{x}$.

Def 3. **Абсолютной и относительной погрешностями** величины $x \neq 0$ называются

$$\Delta = |\widehat{x} - x|, \quad \delta = |\widehat{x} - x|/|x|.$$

NB. 4. *Имея абсолютную и относительную погрешность величины* x, *обычно используют следующую запись:*

$$\widehat{x} = x \pm \Delta = x \pm \delta x = x(1 \pm \delta).$$

1.3 Проблема данных

В качестве примера, демонстрирующего специфику численного решения систем, рассмотрим следующую задачу:

Ex 2. *Решить СЛАУ, заданную матрицей* A *и правой частью* b_1 *и* b_2*:*

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1.0001 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 2 \\ 2.0001 \end{bmatrix}, \quad b_2 = \begin{bmatrix} 2 \\ 2.0000 \end{bmatrix}.$$

Применяя один из аналитических методов, легко получить решение

$$x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Видно, как незначительное изменение в правой части приводит к весьма ощутимому изменению решения.

Ex 3. *Другим примером системы, решение которой сильно зависит от правой части, является система, определяемая матрицей Гильберта:*

$$H_2 = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}, \quad \dots .$$

Исследуйте поведение решения для случая $n = 5$.

1.4 Проблема алгоритма

Рассмотрим другой пример, демонстрирующий роль алгоритма при численном решении:

Ex 4. *Решить СЛАУ методом Гаусса:*

$$A = \begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \varepsilon \approx 10^{-16}.$$

Прямое исключение приводит к потере точности и ошибке:

$$\left[\begin{array}{cc|c} \varepsilon & 1 & 1 \\ 1 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{cc|c} \varepsilon & 1 & 1 \\ 0 & 1 - 1/\varepsilon & 2 - 1/\varepsilon \end{array}\right]$$

Правильный выбор ведущего элемента позволяет получить точное решение:

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ \varepsilon & 1 & 1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 - \varepsilon & 1 - 2\varepsilon \end{array}\right]$$