

Глава 10. ЛИНЕЙНЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

10.1. Сингулярные числа и сингулярные базисы матрицы

Пусть A – матрица размера $m \times n$. Как известно, она порождает линейный оператор $x \rightarrow Ax$, действующий из \mathbb{C}^n в \mathbb{C}^m .

Оператор $y \rightarrow A^*y$, порождаемый сопряженной к A матрицей A^* , действует из \mathbb{C}^m в \mathbb{C}^n .

Рассмотрим матрицы Грама $P = G_A = A^*A$ и $Q = G_{A^*} = (A^*)^*A^* = AA^*$. Как известно, они эрмитовы и неотрицательно определены.

Пусть $\lambda_1, \dots, \lambda_n$ – собственные числа матрицы P , $v^{(1)}, \dots, v^{(n)}$ – соответствующие им ортонормированные собственные векторы в пространстве \mathbb{C}^n . Рассмотрим свойства образов этих векторов в \mathbb{C}^m .

Теорема. 1. Векторы $Av^{(i)}$ и $Av^{(j)}$ ортогональны при $i \neq j$.

2. Если $\lambda_i = 0$, то $Av^{(i)} = \theta_m$.

3. Ненулевое собственное число матрицы P ($\lambda_i > 0$), является также собственным числом матрицы Q , а $Av^{(i)}$ – соответствующий ему собственный вектор этой матрицы.

Доказательство. Обозначим $h^{(i)} = Av^{(i)}$.

$$\begin{aligned} \langle h^{(i)}, h^{(j)} \rangle &= \langle Av^{(i)}, Av^{(j)} \rangle = \langle A^*Av^{(i)}, v^{(j)} \rangle = \\ &= \langle Pv^{(i)}, v^{(j)} \rangle = \lambda_i \langle v^{(i)}, v^{(j)} \rangle. \end{aligned} \quad (10.1.1)$$

Полагая в (10.1.1) $i \neq j$, получаем доказательство утверждения (1).

Полагая там же $i = j$, получаем

$$\|h^{(i)}\|^2 = \lambda_i \|v^{(i)}\|^2 = \lambda_i. \quad (10.1.2)$$

Следовательно, при $\lambda_i = 0$ $\|h^{(i)}\|^2 = 0$, т.е. $h^{(i)} = \theta_m$. Доказано утверждение (2).

Далее

$$\begin{aligned} Qh^{(i)} &= (AA^*)(Av^{(i)}) = A(A^*A)v^{(i)} = A(Pv^{(i)}) = \\ &= A(\lambda_i v^{(i)}) = \lambda_i (Av^{(i)}) = \lambda_i h^{(i)}. \end{aligned} \quad (10.1.3)$$

Из (10.1.2) следует утверждение (3) (при $\lambda_i \neq 0$ $h^{(i)} \neq \theta_m$).

Поменяв местами матрицы P и Q и повторив доказательство, получим

Следствие. Ненулевые собственные числа матриц P и Q попарно совпадают.

В дальнейшем будем считать, что r – общее количество ненулевых собственных чисел этих матриц. Тогда кратности нулевого собственного числа у матриц P и Q будут соответственно равны $n - r$ и $m - r$. При разных порядках матриц у "меньшей" может вообще не быть нулевых собственных чисел.

Упорядочим теперь положительные собственные числа по убыванию

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0.$$

Введем векторы $u^{(j)}$, $j = 1, \dots, m$:

$$\text{При } j \leq r \quad u^{(j)} = \frac{1}{\sqrt{\lambda_j}} h^{(j)} \quad (\text{в силу (10.1.2) } \|u^{(j)}\| = 1.)$$

При $r > j \leq m$ $u^{(j)}$ – ортонормированные собственные векторы матрицы Q , соответствующие нулевому собственному числу.

Из формулы (10.1.1) видно, что векторы $u^{(j)}$, $j = 1, \dots, m$ образуют ортонормированный собственный базис матрицы Q . Далее, по построению

$$\begin{aligned} Av^{(j)} &= \sqrt{\lambda_j} u^{(j)} \quad (j \leq r); \\ Av^{(j)} &= \theta_m \quad (r < j \leq n). \end{aligned} \quad (10.1.4)$$

Аналогично

$$\begin{aligned} A^*u^{(j)} &= \sqrt{\lambda_j} v^{(j)} \quad (j \leq r); \\ A^*u^{(j)} &= \theta_n \quad (r < j \leq m). \end{aligned} \quad (10.1.5)$$

Определение. Квадратные корни из общих ненулевых собственных чисел матриц $P = A^*A$ и $Q = AA^*$ ($\sigma_j = \sqrt{\lambda_j}$, $j = 1, \dots, r$) называются *сингулярными числами* матрицы A . Ортонормированные базисы, состоящие из собственных векторов матрицы P (в пространстве \mathbb{C}^n) и из собственных векторов матрицы Q (в пространстве \mathbb{C}^m) называются *сингулярными базисами* матрицы A (соответственно *правым* и *левым*).

Сведем векторы сингулярных базисов в унитарные матрицы

$$U = [u^{(1)}, \dots, u^{(m)}] \text{ и } V = [v^{(1)}, \dots, v^{(n)}].$$

Теорема.

$$\Sigma = U^*AV = \begin{bmatrix} \Sigma_{r \times r} & \vdots & \mathbb{O}_{r \times (n-r)} \\ \dots\dots\dots & & \\ \mathbb{O}_{(m-r) \times r} & \vdots & \mathbb{O}_{(m-r) \times (n-r)} \end{bmatrix}, \quad (10.1.6)$$

где $\Sigma_{r \times r} = \text{diag}[\sigma_1, \dots, \sigma_r]$, а \mathbb{O} - нулевые матрицы. Размеры матриц обозначены в виде индексов.

Доказательство. Из (10.1.4) имеем

$$\begin{aligned} AV &= A \cdot [v^{(1)}, \dots, v^{(r)}, v^{(r+1)}, \dots, v^{(m)}] = [Av^{(1)}, \dots, Av^{(r)}, Av^{(r+1)}, \dots, Av^{(m)}] = \\ &= [\sigma_1 u^{(1)}, \dots, \sigma_r u^{(r)}, \underbrace{\theta_m, \dots, \theta_m}_{(n-r)}]. \end{aligned}$$

Поэтому

$$\begin{aligned} U^*AV &= U^* \cdot [\sigma_1 u^{(1)}, \dots, \sigma_r u^{(r)}, \underbrace{\theta_m, \dots, \theta_m}_{(n-r)}] = \\ &= [u^{(1)}, \dots, u^{(r)}, u^{(r+1)}, \dots, u^{(m)}]^* \cdot [\sigma_1 u^{(1)}, \dots, \sigma_r u^{(r)}, \underbrace{\theta_m, \dots, \theta_m}_{(n-r)}] = \Sigma. \end{aligned}$$

Равенство $U^*AV = \Sigma$ можно переписать в виде $A = U\Sigma V^*$. Такое представление матрицы называется ее *сингулярным разложением*.

Серьезное предупреждение. Сингулярное разложение матрицы несколько напоминает спектральное разложение эрмитовой матрицы. Однако в общем случае правый и левый сингулярные базисы не совпадают. Даже если матрица эрмитова, можно утверждать лишь, что ее сингулярные числа равны *модулям* ненулевых собственных чисел. Только для неотрицательно определенных эрмитовых матриц сингулярные числа совпадают с ненулевыми собственными числами, правый и левый сингулярные базисы одинаковы и совпадают с собственным базисом матрицы.

Пример.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad A^* = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}, \quad A^*A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}, \quad AA^* = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix};$$

$$\det(A^*A - \lambda I) = \lambda^2 - 6\lambda + 8; \quad \lambda_1 = 4, \quad \lambda_2 = 2.$$

$$\sigma_1 = \sqrt{\lambda_1} = 2, \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{2}; \quad v^{(1)} = \frac{1}{\sqrt{2}}[1, 1]^T, \quad v^{(2)} = \frac{1}{\sqrt{2}}[1, -1]^T.$$

Два собственных числа матрицы AA^* совпадают с ненулевыми собственными числами матрицы A^*A , а третье обязано быть нулем.

Итак, $\mu_1 = \lambda_1 = 4$, $\mu_2 = \lambda_2 = 2$, $\mu_3 = 0$.

Собственные векторы матрицы AA^* , соответствующие ее ненулевым собственным числам, найдем по формуле (10.1.4):

$$u^{(1)} = \frac{1}{\sigma_1}Av^{(1)} = \frac{1}{\sqrt{2}}[1, 1, 0]^T, \quad v^{(2)} = \frac{1}{\sigma_2}Av^{(2)} = \frac{1}{\sqrt{2}}[0, 0, 1]^T.$$

Третий собственный вектор найдем из соответствующей однородной системы и условия нормировки:

$$AA^*x = \theta \bigwedge \|x\| = 1 \iff u^{(3)} = \frac{1}{\sqrt{2}}[1, -1, 0]^T.$$

Сведем собственные векторы в матрицы

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad U = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix}.$$

Выпишем сингулярное разложение $A = U\Sigma V^*$:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

10.2. Псевдорешение системы линейных алгебраических уравнений. Метод наименьших квадратов

Рассмотрим систему уравнений

$$Ax = b \quad (10.1)$$

с матрицей A размера $m \times n$, столбцом свободных членов $b \in \mathbb{C}^m$ и переменным вектором $x \in \mathbb{C}^n$.

Назовем *вектором невязок* системы (10.2.1) вектор

$$d(x) = b - Ax.$$

Тогда, очевидно, решением системы (10.2.1) будет такой вектор $x^{(0)} \in \mathbb{C}^n$, что $d(x^{(0)}) = \theta_m$.

При решении содержательных прикладных задач часто встречаются ситуации, в которых система линейных уравнений несовместна, хотя по "физическому смыслу" решение должно существовать. Объясняется это, как правило, тем, что коэффициенты и свободные члены системы, полученные из эксперимента, содержат погрешности.

В таких ситуациях разумно выбирать переменный вектор в системе (10.2.1) так, чтобы норма невязки, которую мы не можем сделать равной нулю, оказалась бы минимально возможной.

Определение. Псевдорешением системы линейных алгебраических уравнений (10.2.1) называется такой вектор $\tilde{x} \in \mathbb{C}^n$, что $\|d(\tilde{x})\| \leq \|d(x)\|$ для всех $x \in \mathbb{C}^n$, т.е. вектор, минимизирующий *евклидову* норму невязки.

Отметим, что в случае совместной системы ее псевдорешение \tilde{x} совпадает с решением $x^{(0)}$, так как $d(x^{(0)}) = \theta_m$. и, следовательно, $0 = \|d(x^{(0)})\| \leq \|d(x)\|$ для всех $x \in \mathbb{C}^n$.

На вопрос о существовании псевдорешения дает ответ следующая

Теорема. Всякая система линейных алгебраических уравнений имеет псевдорешение.

Доказательство. Пусть $A = U\Sigma V^*$ – сингулярное разложение матрицы коэффициентов системы (напомним, что U и V – унитарные матрицы сингулярных базисов – имеют порядки m и n соответственно, а $(m \times n)$ -матрица Σ имеет вид (10.1.6)).

Разложим вектор свободных членов b по сингулярному базису в \mathbb{C}^m :

$$b = c_1 u^{(1)} + \dots + c_m u^{(m)} \quad \text{или} \quad b = Uc. \quad (10.2.2)$$

Вектор-псевдорешение \tilde{x} будем искать в виде разложения по сингулярному базису в \mathbb{C}^n :

$$\tilde{x} = \alpha_1 v^{(1)} + \dots + \alpha_n v^{(n)} \quad \text{или} \quad \tilde{x} = V\alpha, \quad (10.2.3)$$

где α – новый искомый вектор.

Подставив (10.2.2) и (10.2.3) в уравнение (10.2.1), получим

$$AV\alpha = Uc.$$

Умножая это уравнение на $U^* = U^{-1}$ слева и учитывая, что $U^*AV = \Sigma$, имеем:

$$\Sigma\alpha = c. \quad (10.2.4)$$

Покажем, что нормы невязок систем (10.2.1) и (10.2.4) равны, т.е. задача свелась к отысканию псевдорешения системы (10.2.4). Действительно, из (10.2.3) имеем $\alpha = V^* \tilde{x}$ и, следовательно,

$$b - A\tilde{x} = Uc - U\Sigma V^* \tilde{x} = U \cdot (c - \Sigma\alpha).$$

Но умножение на унитарную матрицу сохраняет норму вектора и

$$\|b - A\tilde{x}\| = \|U \cdot (c - \Sigma\alpha)\| = \|c - \Sigma\alpha\|.$$

Выпишем подробно вектор $\Sigma\alpha$:

$$[\sigma_1\alpha_1, \dots, \sigma_r\alpha_r, \underbrace{0, \dots, 0}_{(m-r)}]^T$$

и квадрат нормы невязки для системы (10.2.4)

$$\|c - \Sigma\alpha\|^2 = |c_1 - \sigma_1\alpha_1|^2 + \dots + |c_r - \sigma_r\alpha_r|^2 + |c_{r+1}|^2 + \dots + |c_m|^2 \quad (10.2.5)$$

(по-прежнему, r – количество сингулярных чисел).

Из (10.2.5) видно, что минимум нормы невязки достигается при

$$\alpha_j = \frac{c_j}{\sigma_j}, \quad j = 1, \dots, r \quad (10.2.6)$$

и равен $(|c_{r+1}|^2 + \dots + |c_m|^2)^{1/2}$. Теорема доказана.

Рассмотрим частные случаи.

1. $r = n$. Равенства (10.2.6) однозначно определяют все компоненты вектора α . По формуле (10.2.3) получаем $\tilde{x} = V\alpha$ – единственное псевдорешение системы (10.2.1).

2. $r < n$. Из (10.2.6) определяются только первые r компонент вектора α . Однако рассмотрение (10.2.5) показывает, что оставшиеся компоненты не влияют на величину невязки и могут быть выбраны произвольно. В этом случае псевдорешение *не единственно*. Обычно приравнивают эти компоненты нулю и называют полученное псевдорешение *нормальным псевдорешением*. Нормальное псевдорешение имеет наименьшую среди всех псевдорешений норму.

3. $r = m$. В этом случае любое псевдорешение обеспечивает нулевую норму невязки, т.е. является решением в обычном смысле этого слова.

Серьезное предупреждение. Не следует путать понятия "псевдорешение" и "приближенное решение" системы линейных алгебраических уравнений. О приближенном решении можно говорить только в случае совместной системы (вектор приближенного решения в каком-то смысле близок к существующему "точному" решению). Но в этом случае псевдорешение совпадает с решением.

В отличие от решения псевдорешение существует и у несовместной системы, когда говорить о приближенном решении бессмысленно, так как решение отсутствует и приближаться не к чему!

Псевдорешение часто называют *решением по методу наименьших квадратов*. Это историческое название объясняется тем, что минимизируется квадрат евклидовой нормы невязки, который сводится к сумме квадратов модулей невязок всех уравнений системы.

Почему из всевозможных мер близости выбрана именно *евклидова* норма? А лишь потому, что такой выбор приводит к простому алгоритму построения псевдорешения. Никаких "более глубоких" обоснований у этого метода нет. Заметим, что иногда пользуются и другими нормами.

Замечание. Геометрическая интерпретация псевдорешения системы $Ax = b$ весьма проста: это такой вектор \tilde{x} , который при умножении на матрицу A переходит в вектор пространства \mathbb{C}^m , "ближайший" к вектору b . При этом расстояние между векторами определяется по теореме Пифагора.

Пример. Найдём нормальное псевдорешение симтемы линейных уравнений

$$\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \\ x_1 - x_3 = 3 \end{cases}$$

или, в матричном виде,

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Сингулярное разложение матрицы коэффициентов этой системы было получено в п.10.1. Используя его, перепишем систему

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Умножив равенство слева на $U^* = U^{-1}$, получим

$$\begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ 3 \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{или} \quad \begin{cases} 2\alpha_1 & = & \frac{3}{\sqrt{2}} \\ \sqrt{2}\alpha_2 & = & 3 \\ 0\alpha_3 & = & -\frac{1}{\sqrt{2}} \end{cases}.$$

Из первых двух уравнений имеем

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{2\sqrt{2}} \\ \frac{3}{\sqrt{2}} \end{bmatrix}.$$

Умножив обе части равенства на $V = (V^*)^{-1}$, получим нормальное псевдорешение системы

$$\tilde{x} = \begin{bmatrix} \frac{9}{4} \\ -\frac{3}{4} \end{bmatrix}.$$

Рассмотрим геометрическую интерпретацию этого примера.

Если $y = [y_1, y_2, y_3]^T$ – образ вектора $x = [x_1, x_2]^T$ при отображении $y = Ax$, порожденном матрицей коэффициентов системы, то $y_1 = x_1 + x_2 = y_2$, т.е. множество значений этого отображения есть плоскость $y_1 = y_2$ в \mathbb{R}^3 . Точка с координатами $(1, 2, 3)$ – конец направленного отрезка \vec{b} в этой плоскости не лежит, т.е. система уравнений не имеет решения. Ближайшая к этой точке точка плоскости есть образ псевдорешения (см. замечание):

$$A\tilde{x} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} \frac{9}{4} \\ -\frac{3}{4} \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ \frac{3}{2} \\ 3 \end{bmatrix} \quad (.10.1).$$

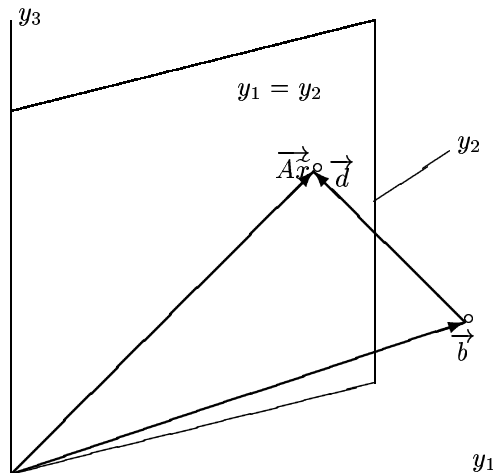


Рис.10.1

Вектор невязки $d(\tilde{x}) = [\frac{1}{2}, \frac{1}{2}, 0]^T$, а его норма $\|d(\tilde{x})\| = \frac{1}{\sqrt{2}}$.

10.3 Псевдообратная матрица. Нормальные уравнения

В предыдущем пункте было построено псевдорешение системы (10.2.1). Запишем формулы для вычисления псевдорешения в матричной форме. Для этого введем матрицу

$$\Sigma^+ = \begin{bmatrix} \Sigma_{r \times r}^+ & \vdots & \mathbb{O}_{r \times (n-r)} \\ \dots\dots\dots & & \\ \mathbb{O}_{(m-r) \times r} & \vdots & \mathbb{O}_{(m-r) \times (n-r)} \end{bmatrix},$$

где $\Sigma_{r \times r}^+ = \text{diag}[\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}]$, а \mathbb{O} - нулевые матрицы. Размеры матриц обозначены в виде индексов.

Нетрудно убедиться в том, что формулы (10.2.6) и (10.2.7) можно переписать в виде $\alpha = \Sigma^+ c$ (проделайте это полезное упражнение). Тогда нормальное псевдорешение системы (10.2.1) запишется в виде

$$\tilde{x} = V\alpha = V\Sigma^+c = V\Sigma^+U^*b.$$

Если матрица коэффициентов системы A *квадратная и обратимая*, то решение этой системы получается, как известно, умножением свободного члена слева на обратную матрицу: $x = A^{-1}b$.

Назовем матрицу $A^+ = V\Sigma^+U^*$ *псевдообратной*. Тогда псевдорешение будет получаться умножением свободного члена слева на псевдообратную матрицу:

$$\tilde{x} = A^+b. \quad (10.3.1)$$

Замечание. Если $r = m = n$, матрица Σ , очевидно, обратима и $\Sigma^+ = \Sigma^{-1}$. Поэтому

$$A^+ = V\Sigma^+U^* = V\Sigma^{-1}U^{-1} = (U\Sigma V^{-1})^{-1} = (U\Sigma V^*)^{-1} = A^{-1},$$

т.е. если матрица обратима, то ее псевдообратная совпадает с ее обратной. Это согласуется с утверждением (см.п.10.2), что в этом случае псевдорешение единственно ($r = n$) и является решением системы ($r = m$).

Серьезное предупреждение. Формула (10.3.1) есть не более, чем удобная форма *записи результата*. Так же как для нахождения решения системы не следует вычислять обратную матрицу и пользоваться затем формулой $x = A^{-1}b$, для нахождения нормального псевдорешения не следует вычислять псевдообратную матрицу и пользоваться затем формулой $\tilde{x} = A^+b$.

Докажем теперь, что $A^*AA^+ = A^*$.

$$\begin{aligned} 1) \Sigma \cdot \Sigma^+ &= \begin{bmatrix} \Sigma_{r \times r} & \vdots & \mathbb{O}_{r \times (n-r)} \\ \dots\dots\dots & & \\ \mathbb{O}_{(m-r) \times r} & \vdots & \mathbb{O}_{(m-r) \times (n-r)} \end{bmatrix} \cdot \begin{bmatrix} \Sigma_{r \times r}^+ & \vdots & \mathbb{O}_{r \times (m-r)} \\ \dots\dots\dots & & \\ \mathbb{O}_{(n-r) \times r} & \vdots & \mathbb{O}_{(n-r) \times (m-r)} \end{bmatrix} = \\ &= \begin{bmatrix} I_r & \vdots & \mathbb{O}_{r \times (m-r)} \\ \dots\dots\dots & & \\ \mathbb{O}_{(m-r) \times r} & \vdots & \mathbb{O}_{(m-r) \times (m-r)} \end{bmatrix}, \\ 2) \Sigma^* \cdot \Sigma \cdot \Sigma^+ &= \begin{bmatrix} \Sigma_{r \times r} & \vdots & \mathbb{O}_{r \times (n-r)} \\ \dots\dots\dots & & \\ \mathbb{O}_{(m-r) \times r} & \vdots & \mathbb{O}_{(m-r) \times (n-r)} \end{bmatrix} \cdot \begin{bmatrix} I_r & \vdots & \mathbb{O}_{r \times (m-r)} \\ \dots\dots\dots & & \\ \mathbb{O}_{(m-r) \times r} & \vdots & \mathbb{O}_{(m-r) \times (m-r)} \end{bmatrix} = \Sigma^*. \end{aligned}$$

Поэтому

$$\begin{aligned} A^*AA^+ &= (U\Sigma V^*)^* \cdot (U\Sigma V^*) \cdot (V\Sigma^+U^*) = \\ &= V\Sigma^*(U^*U)\Sigma(V^*V)\Sigma^+U^* = V \cdot (\Sigma^*\Sigma\Sigma^+) \cdot U^* = V\Sigma^*U^* = A^*. \end{aligned}$$

Умножив теперь обе части (10.3.1) слева на A^*A , получим

$$A^*A\tilde{x} = (A^*AA^+)b = A^*b.$$

Отсюда *хочется сделать вывод*, что *нормальное псевдорешение* системы (10.2.1) будет *решением* системы

$$(A^*A)x = A^*b, \quad (10.3.2)$$

получающейся из (10.2.1) умножением обеих частей слева на матрицу A^* . Уравнения (10.3.2) называются *нормальными уравнениями* метода наименьших квадратов.

Серьезное предупреждение. Возникает вопрос: для чего же было мучиться столько времени, вводить новые понятия "сингулярное" разложение и "псевдорешение"? Не проще ли построить систему нормальных уравнений и решить ее?

Оказывается, все не так просто.

Во-первых, никто не гарантирует нам невырожденности матрицы $G_A = A^* A$, (которая равносильна, как известно, линейной независимости столбцов матрицы A). Если матрица Грама окажется вырожденной, то нам все равно придется искать псевдорешение, но уже системы (10.3.2)!

Во-вторых, если даже матрица системы (10.3.2) не вырождена, то при переходе к нормальным уравнениям резко возрастают вычислительные трудности (см. главу 13).

Поэтому мы настоятельно рекомендуем *не пользоваться нормальными уравнениями* (тем более, что численно устойчивые алгоритмы сингулярного разложения и построения нормального псевдорешения реализованы и в средах конечного пользователя и в библиотеках Фортрана).

Замечания. 1. Отметим один случай, когда все-таки можно переходить к нормальным уравнениям. Если столбцы матрицы попарно ортогональны, то матрица Грама окажется диагональной, и решение системы (10.3.2) – нормальное псевдорешение системы (10.2.1) – выписывается в явном виде:

$$\tilde{x}_k = \frac{(A^* b)_k}{\|a^{(k)}\|^2} = \frac{\langle b, a^{(k)} \rangle}{\|a^{(k)}\|^2}, \quad k = 1, \dots, n. \quad (10.3.3)$$

2. Численно устойчивые алгоритмы сингулярного разложения прямоугольной матрицы и построения нормального псевдорешения системы линейных уравнений реализованы в библиотеках ФОРТРАНА и в средах конечного пользователя.