ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Январь—июнь 2002. Серия 2. Том 9, № 1, 33–60

УДК 519.852

СВОЙСТВА ОПТИМАЛЬНОЙ ОДНОРАНГОВОЙ КОРРЕКЦИИ МАТРИЦ КОЭФФИЦИЕНТОВ НЕСОВМЕСТНЫХ НЕОДНОРОДНЫХ ЛИНЕЙНЫХ МОДЕЛЕЙ

В. И. Ерохин

Рассматриваются задачи построения оптимальных (в смысле минимума евклидовой или спектральной нормы) одноранговых матриц коррекции для матриц несовместных неоднородных линейных моделей вида $\mathbf{Ax} = \mathbf{b}$ и $\mathbf{Ax} = \mathbf{b}$, $\mathbf{x} \geqslant \mathbf{0}$. Исследуются необходимые и достаточные условия достижимости нижней грани нормы матрицы коррекции. Анализируются случаи неединственности как самих оптимальных матриц коррекции, так и решений скорректированных линейных систем, строятся соответствующие параметрические семейства матриц и векторов.

Введение

Пусть даны две несовместные линейные системы

$$\mathbf{Ax} = \mathbf{b} \tag{1}$$

и

$$\mathbf{Ax} = \mathbf{b}, \quad \mathbf{x} \geqslant \mathbf{0}, \tag{2}$$

где $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m\times n}$, причем $\operatorname{rank} \mathbf{A}$ произволен,

$$\mathbf{b} \neq \mathbf{0}. \tag{3}$$

Требуется построить матрицу $\mathbf{H} \in \mathbb{R}^{m\times n}$ такую, что системы (1), (3) и (2), (3) с матрицей $\widetilde{\mathbf{A}} = \mathbf{A} + \mathbf{H}$ совместны, $\|\mathbf{H}\| \to \inf$ и $\operatorname{rank} \mathbf{H} = 1$. При этом здесь и ниже символ $\|\cdot\|$ обозначает в зависимости от контекста евклидову норму вектора или евклидову (спектральную) норму [1)] одноранговой матрицы, т. е. $\|\mathbf{x}\| = \sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{x}}$, $\|\mathbf{H}\| = \max\limits_{\|\mathbf{x}\|=1} \|\mathbf{Hx}\| = \sqrt{\sum\limits_{i,j=1}^{m,n} h_{ij}^2}$.

[1)] Известно (см., например, [3]), что одноранговость матрицы эквивалентна совпадению ее евклидовой и спектральной норм.

Множество решений системы (1), (3), скорректированной с помощью некоторой (не обязательно оптимальной) матрицы $\mathbf{H}$, будем обозначать через $\mathscr{X}(\mathbf{A}, \mathbf{b}, \mathbf{H})$; для соответствующего множества решений скорректированной системы (2), (3) будем использовать обозначение $\mathscr{X}_+(\mathbf{A}, \mathbf{b}, \mathbf{H})$.

Очевидно, что коррекция системы (2), (3) может быть интерпретирована как коррекция модели допустимой области некоторой несобственной задачи линейного программирования, записанной в канонической форме. Проблема коррекции системы (1), (3) может представлять самостоятельный интерес или рассматриваться в качестве вспомогательной задачи при решении проблемы коррекции системы (2), (3).

В настоящей статье продолжаются исследования, начатые в основополагающих работах И. И. Еремина и А. А. Ватолина [2, 8, 14] и в более поздних работах В. А. Горелика и В. А. Кондратьевой [7, 11]. Основные результаты из [7, 11], которые мы будем использовать в настоящей статье, сформулированы в приводимых ниже теоремах 1–3: [2)]

Теорема 1.

$$\inf_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\infty} \|\mathbf{H}\|^2 = \mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}), \tag{4}$$

где $\mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A})$ *— минимальное собственное значение матрицы* $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$, $\mathbf{I}$ *— единичная матрица,*

$$\mathbf{b}^+ = (\mathbf{b}^{\mathbf{T}}\mathbf{b})^{-1} \cdot \mathbf{b}^{\mathbf{T}} \tag{5}$$

— вектор, псевдообратный [3)] *вектору* $\mathbf{b}$.

Если нижняя грань достигается, то

$$\mathbf{H}^* = -(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}\mathbf{y}\mathbf{y}^{\mathbf{T}} \in \operatorname*{Arginf}_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|, \tag{6}$$

где [4)] $\mathbf{y} \in \mathscr{M}_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A})$ *и* $\mathscr{M}_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A})$ *— множество единичных собственных векторов матрицы* $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$, *соответствующих ее минимальному собственному значению* $\mu_{\min}$. *При этом*

$$\mathbf{x}^* = \frac{\mathbf{b}^{\mathbf{T}}\mathbf{b}}{\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}} \cdot \mathbf{y} \in \mathscr{X}(\mathbf{A}, \mathbf{b}, \mathbf{H}^*). \tag{7}$$

[2)] В теоремах 1–3 предполагается несовместность систем вида (1), (3) и (2), (3).

[3)] В соответствии, например, с [4] вектор $\mathbf{b}^+$, псевдообратный некоторому ненулевому вектору-столбцу $\mathbf{b}$, может рассматриваться как вектор-строка, являющийся решением с минимальной евклидовой нормой уравнения $\mathbf{b}^+\mathbf{b} = 1$.

[4)] В дальнейшем символом $\mathscr{M}_{\min}(\cdot)$ обозначается множество единичных собственных векторов некоторой вещественной симметричной матрицы, соответствующих ее минимальному собственному значению.

Теорема 2. $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ *достигается тогда и только тогда, когда существует вектор* $\mathbf{y} \in \mathscr{M}_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A})$ *такой, что*

$$\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y} \neq 0. \tag{8}$$

Теорема 3.

$$\inf_{\mathscr{X}_{+}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|^2 = \min\{d_1, d_2\}, \tag{9}$$

где

$$d_1 = \min_{\mathbf{x}:\mathbf{x}\geqslant\mathbf{0},\|\mathbf{x}\|=1,\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{x}\geqslant 0} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x})\}, \tag{10}$$

$$d_2 = \min_{\mathbf{x}:\mathbf{x}\geqslant\mathbf{0},\|\mathbf{x}\|=1,\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{x}\leqslant 0} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{A}\mathbf{x}\}, \tag{11}$$

причем нижняя грань в (9) *достигается тогда и только тогда, когда* $d_1 \leqslant d_2$ *и существует вектор* $\mathbf{y}$, *являющийся решением задачи* (10), *такой, что*

$$\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y} > 0. \tag{12}$$

При этом

$$\mathbf{H}^* = -(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{y}\mathbf{y}^{\mathbf{T}} \in \operatorname*{Arginf}_{\mathscr{X}_{+}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|, \tag{13}$$

$$\mathbf{x}^* = \frac{\mathbf{b}^{\mathbf{T}}\mathbf{b}}{\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}} \cdot \mathbf{y} \in \mathscr{X}_{+}(\mathbf{A},\mathbf{b},\mathbf{H}^*). \tag{14}$$

В настоящей статье получены некоторые новые результаты, относящиеся к одноранговым оптимальным коррекциям систем (1), (3) и (2), (3).

1. Свойства одноранговой оптимальной коррекции матрицы коэффициентов несовместной линейной модели вида (1), (3)

Обсуждение свойств одноранговой оптимальной коррекции матрицы коэффициентов несовместной линейной модели вида (1), (3) начнем с альтернативной по отношению к теореме 2 формулировки необходимых и достаточных условий достижимости нижней грани нормы матрицы коррекции. Следуя теореме 2, для соответствующей проверки предварительно следует определить $\mu_{\min}$ — минимальное собственное значение матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}$ — и построить соответствующий этому

значению собственный вектор или соответствующее семейство собственных векторов в случае неединичной кратности [5] значения $\mu_{\min}(\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{bb}^{+})\mathbf{A})$. Ниже будет показано, что предлагаемый в теореме 4 альтернативный вариант при определенных условиях обладает меньшей трудоемкостью.

Теорема 4. *Для несовместной системы вида* (1), (3) $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|$ *достигается тогда и только тогда, когда* (a) *либо минимальное собственное значение* $\mu_{\min}$ *матрицы* $\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{bb}^{+})\mathbf{A}$ *не является собственным значением матрицы* $\mathbf{A}^{\mathrm{T}}\mathbf{A}$, (b) *либо кратность значения* $\mu_{\min}$ *среди собственных значений матрицы* $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ *меньше его кратности среди собственных значений матрицы* $\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{bb}^{+})\mathbf{A}$.

При доказательстве теоремы будут использоваться два вспомогательных утверждения.

Лемма 1. *Пусть для некоторой* (*не обязательно несовместной*) *системы вида* (1), (3) *существует некоторый вектор* $\mathbf{y}$ *такой, что*

$$\mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{y}=0, \tag{15}$$

$$\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{bb}^{+})\mathbf{A}\mathbf{y}=\mu\mathbf{y} \tag{16}$$

($\mathbf{y}$ — *собственный вектор матрицы* $\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{bb}^{+})\mathbf{A}$, μ — *соответствующее собственное значение*). *Тогда*

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{y}=\mu\mathbf{y} \tag{17}$$

($\mathbf{y}$ *и* μ — *собственный вектор и собственное значение матрицы* $\mathbf{A}^{\mathrm{T}}\mathbf{A}$).

Соотношение (17) следует из условий (15)–(16) и тождества

$$\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{bb}^{+})\mathbf{A}=\mathbf{A}^{\mathrm{T}}\mathbf{A}-\frac{\mathbf{A}^{\mathrm{T}}\mathbf{b}\mathbf{b}^{\mathrm{T}}\mathbf{A}}{\mathbf{b}^{\mathrm{T}}\mathbf{b}}, \tag{18}$$

которое, в свою очередь, вытекает из соотношения (5).

Лемма 2. *Пусть для некоторой* (*не обязательно несовместной*) *системы вида* (1), (3) *минимальное собственное значение* $\mu_{\min}$ *матрицы* $\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{bb}^{+})\mathbf{A}$ *является собственным значением матрицы* $\mathbf{A}^{\mathrm{T}}\mathbf{A}$. *Тогда для любого собственного вектора* $\mathbf{z}$ *матрицы* $\mathbf{A}^{\mathrm{T}}\mathbf{A}$, *соответствующего*

[5] Поскольку матрицы $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ и $\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{bb}^{+})\mathbf{A}$ являются симметричными, для их собственных значений *алгебраическая* кратность (кратность корня характеристического многочлена) совпадает с *геометрической* кратностью (максимальным количеством линейно независимых собственных векторов, соответствующих данному собственному значению) [3]. Поэтому здесь и в дальнейшем будем говорить просто о *кратности* собственных значений соответствующих матриц.

собственному значению $\mu_{\min}$, *выполняется условие* $\mathbf{b^T A z} = 0$, *а вектор* $\mathbf{z}$ *принадлежит набору собственных векторов матрицы* $\mathbf{A^T(I - bb^+)A}$, *соответствующих собственному значению* $\mu_{\min}$.

Доказательство. Предположим противное. Пусть существует такой собственный вектор $\mathbf{z}$ матрицы $\mathbf{A^T A}$, соответствующий собственному значению $\mu_{\min}$, что $\mathbf{b^T A z} \neq 0$. Положим $\mathbf{\bar{z}} = \frac{\mathbf{z}}{\|\mathbf{z}\|}$ и рассмотрим значение квадратичной формы $\mathbf{\bar{z}^T A^T (I - bb^+) A \bar{z}}$. В силу (18) имеем $\mathbf{\bar{z}^T A^T (I - bb^+) A \bar{z}} = \mathbf{\bar{z}^T A^T A \bar{z}} - (\mathbf{b^T A \bar{z}})^2 \cdot (\mathbf{b^T b})^{-1}$. Так как $\mathbf{\bar{z}^T A^T A \bar{z}} = \mu_{\min}$, то $\mathbf{\bar{z}^T A^T (I - bb^+) A \bar{z}} < \mu_{\min}$. Противоречие. Таким образом, $\mathbf{b^T A \bar{z}} = 0$. Отсюда следует, что $\mathbf{b^T A z} = 0$, а также то, что вектор $\mathbf{z}$ принадлежит набору собственных векторов матрицы $\mathbf{A^T(I - bb^+)A}$, соответствующих собственному значению $\mu_{\min}$. Лемма 2 доказана.

Перейдем к доказательству теоремы 4. (Достаточность и необходимость обосновываются рассуждениями от противного.)

Доказательство. **Достаточность.**

Случай (a). Пусть $\mu_{\min}(\mathbf{A^T(I - bb^+)A})$ не является собственным значением матрицы $\mathbf{A^T A}$, но $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H}) \neq \varnothing} \|\mathbf{H}\|$ не достигается. В этом случае в силу теоремы 2 для любого $\mathbf{y}$, $\|\mathbf{y}\| = 1$, имеем

$$\mathbf{A^T(I - bb^+)Ay} = \mu_{\min}\mathbf{y} \Rightarrow \mathbf{b^T A y} = 0.$$

Отсюда и из леммы 1 следует, что $\mu_{\min}$ — собственное значение матрицы $\mathbf{A^T A}$. Противоречие.

Случай (b). Пусть $\mu_{\min}(\mathbf{A^T(I-bb^+)A})$ входит в набор собственных значений матрицы $\mathbf{A^T A}$; $\mathbf{z}_1, \dots, \mathbf{z}_k$ — соответствующий $\mu_{\min}$ набор линейно независимых собственных векторов матрицы $\mathbf{A^T A}$, содержащий максимально возможное число векторов k; $\mathbf{y}_i, \dots, \mathbf{y}_p$ — соответствующий $\mu_{\min}$ набор линейно независимых собственных векторов матрицы $\mathbf{A^T(I - bb^+)A}$, содержащий максимально возможное число векторов, равное p, $k < p$, но $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H}) \neq \varnothing} \|\mathbf{H}\|$ не достигается. Тогда по теореме 2 для любого вектора $\mathbf{y}_i$ из указанного выше набора векторов выполняется условие $\mathbf{b^T A y}_i = 0$. Отсюда с использованием леммы 1 k векторов $\mathbf{y}_i$ образуют альтернативный векторам $\mathbf{z}_i$ набор собственных векторов матрицы $\mathbf{A^T A}$, соответствующих ее собственному значению $\mu_{\min}$. Таким образом, кратность собственного значения $\mu_{\min}$ для указанных матриц оказалась одинаковой и равна k. Противоречие.

Необходимость. Пусть $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H}) \neq \varnothing} \|\mathbf{H}\|$ достигается, но число $\mu_{\min}(\mathbf{A^T(I - bb^+)A})$ является собственным значением матрицы $\mathbf{A^T A}$, причем его кратность для матриц $\mathbf{A^T(I - bb^+)A}$ и $\mathbf{A^T A}$ одинакова.

Пусть она равна $k \geqslant 1$. В силу теоремы 2 существует хотя бы один единичный вектор $\mathbf{y}$ матрицы $\mathbf{A}^\mathbf{T}(\mathbf{I}-\mathbf{b}\mathbf{b}^+)\mathbf{A}$, соответствующий ее минимальному собственному значению $\mu_{\min}$, такой, что $\mathbf{b}^\mathbf{T}\mathbf{A}\mathbf{y} \neq 0$. В то же время в силу леммы 2 все векторы из набора собственных векторов матрицы $\mathbf{A}^\mathbf{T}\mathbf{A}$, соответствующие собственному значению $\mu_{\min}$, также являются собственными векторами матрицы $\mathbf{A}^\mathbf{T}(\mathbf{I}-\mathbf{b}\mathbf{b}^+)\mathbf{A}$. Эти векторы обозначим через $\mathbf{z}_1, \ldots, \mathbf{z}_k$. В силу леммы 2 при любом $i = 1, \ldots, k$ имеем $\mathbf{b}^\mathbf{T}\mathbf{A}\mathbf{z}_i = 0$. По этой причине вектор $\mathbf{y}$ не может быть линейной комбинацией векторов $\mathbf{z}_1, \ldots, \mathbf{z}_k$. Это означает, что максимальное число линейно независимых векторов в наборе собственных векторов матрицы $\mathbf{A}^\mathbf{T}(\mathbf{I}-\mathbf{b}\mathbf{b}^+)\mathbf{A}$, соответствующих ее минимальному собственному значению $\mu_{\min}$, больше k. Противоречие. Теорема 4 доказана.

Замечание. Известно (см., например, [5]), что в настоящее время основным практическим вычислительным инструментом исследования собственных значений и собственных векторов вещественных матриц является *сингулярное разложение*, т. е. разложение вида $\mathbf{G} = \mathbf{V}\mathbf{\Sigma}\mathbf{W}^\mathbf{T}$, где $\mathbf{G} \in \mathbb{R}^{m\times n}$ — исследуемая матрица; $\mathbf{V} \in \mathbb{R}^{m\times m}$ — ортогональная матрица, составленная из собственных векторов матрицы $\mathbf{G}\mathbf{G}^\mathbf{T}$; $\mathbf{W} \in \mathbb{R}^{n\times n}$ — ортогональная матрица, составленная из собственных векторов матрицы $\mathbf{G}^\mathbf{T}\mathbf{G}$; $\mathbf{\Sigma} = (\sigma_{i,j}) \in \mathbb{R}^{m\times n}$ — такая матрица, что $\sigma_{i,j} = 0$ при $i \neq j$; числа $\sigma_{i,i}$, где $i = 1, 2, \ldots, \min\{m, n\}$, — *сингулярные числа* матрицы $\mathbf{G}$.

Ниже приводится таблица из [5], в которой даны оценки трудоемкости [6)] двенадцати вариантов сингулярного разложения $\mathbf{G} = \mathbf{V}\mathbf{\Sigma}\mathbf{W}^\mathbf{T}$ матрицы $\mathbf{G} \in \mathbb{R}^{m\times n}$, где $m \geqslant n$, различающихся особенностями реализации и числом элементов сингулярного разложения, вычисляемых и сохраняемых в явном виде.

Общей особенностью указанных алгоритмов является то, что они не позволяют вычислять отдельные сингулярные числа — может быть вычислен только весь набор сингулярных чисел одновременно. Невозможно также вычислить отдельный собственный вектор из набора векторов $\mathbf{V}$ или $\mathbf{W}$ — вычисляется (или не вычисляется) либо соответствующая матрица целиком, либо, в случае матрицы $\mathbf{V}$, ее подматрица $\mathbf{V}_1$, состоящая из первых n столбцов. Для проверки достижимости величины $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ в задаче оптимальной коррекции матрицы коэффициентов несовместной линейной системы (1), (3) в случае

[6)] В работе [5] мерой трудоемкости является количество «флопов» — операций с плавающей запятой (без детализации на операции сложения, умножения, издержки на индексацию и пр.).

$m \geqslant n$ [7] в соответствии с теоремой 2 с использованием сингулярного разложения необходимо определить сингулярные числа матрицы $(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$ [8] и n линейно независимых собственных векторов матрицы $\mathbf{A}^\mathrm{T}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$. В то же время для аналогичной проверки в соответствии с теоремой 4 с помощью сингулярного разложения необходимо определить сингулярные числа матриц $\mathbf{A}$ и $(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$.

Трудоемкость сингулярного разложения вещественной матрицы размера $m \times n$

Вычисляется	Метод Голуба — Рейнча	R-SVD
$\mathbf{\Sigma}$	$4mn^2 - 4n^3/3$	$2mn^2 + 2n^3$
$\mathbf{\Sigma}, \mathbf{W}$	$4mn^2 + 8n^3$	$2mn^2 + 11n^3$
$\mathbf{\Sigma}, \mathbf{V}$	$4m^2n - 8mn^2$	$4m^2n + 13n^3$
$\mathbf{\Sigma}, \mathbf{V}_1$	$14mn^2 - 2n^3$	$6mn^2 + 11n^3$
$\mathbf{\Sigma}, \mathbf{V}, \mathbf{W}$	$4m^2n + 8mn^2 + 9n^3$	$4m^2n + 22n^3$
$\mathbf{\Sigma}, \mathbf{V}_1, \mathbf{W}$	$14mn^2 + 8n^3$	$6mn^2 + 20n^3$

Таким образом, как следует из приведенных в таблице данных, при $n \leqslant m < 3{,}5n$ двукратное использование алгоритма R-SVD для определения сингулярных чисел матриц $\mathbf{A}$ и $(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$ имеет меньшую суммарную временную сложность, чем построение собственных векторов матрицы $\mathbf{A}^\mathrm{T}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$ и сингулярных чисел матрицы $(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$ любым из перечисленных в таблице способов сингулярного разложения.

Пример 1. Рассмотрим три несовместные системы линейных уравнений с общей матрицей коэффициентов $\mathbf{A}$: $\mathbf{A}[\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3] = [\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3]$, где

$$\mathbf{A} = \begin{bmatrix} 3 & -7 & 5 & -1 & 5 \\ -4 & -5 & 7 & -3 & 0 \\ 0 & -5 & 2 & -9 & 2 \\ 0 & 1 & 10 & -3 & 2 \\ 6 & -3 & 5 & -5 & 2 \\ 1 & -5 & 7 & -3 & -5 \end{bmatrix},$$

[7] Исследование случая $m < n$ с помощью сингулярного разложения не имеет смысла, поскольку, как будет показано ниже, условие $m \geqslant n$ является *необходимым* для достижимости $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H}) \neq \varnothing} \|\mathbf{H}\|$.

[8] В силу (5) имеем $(\mathbf{I} - \mathbf{b}\mathbf{b}^+)^\mathrm{T}(\mathbf{I} - \mathbf{b}\mathbf{b}^+) = (\mathbf{I} - \mathbf{b}\mathbf{b}^+)$. Поэтому матрица вида $\mathbf{A}^\mathrm{T}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$ допускает представление $\mathbf{A}^\mathrm{T}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A} = ((\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A})^\mathrm{T} \cdot (\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$.

$$[\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3] = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -2 \\ 1 & -2 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & -2 \\ -1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 16\sqrt{\frac{7}{2}} & \sqrt{70} & 2\sqrt{70} \\ 56/\sqrt{5} & 7 & 14 \\ 12\sqrt{\frac{133}{15}} & 6\sqrt{\frac{119}{33}} & 12\sqrt{14} \end{bmatrix}$$

$$\approx \begin{bmatrix} 90,70956 & 26,76038 & 75,63309 \\ -126,44190 & -38,15416 & -120,53298 \\ 15,57767 & 5,76038 & 33,63309 \\ 55,88700 & 17,02718 & 56,16669 \\ -16,48746 & -7,42096 & -59,06658 \\ -19,24488 & -3,97282 & 14,16669 \end{bmatrix}.$$

Значения следующих величин можно указать *точно* (они являются целочисленными): $\lambda_{\min}(\mathbf{A}^{\mathbf{T}}\mathbf{A}) = 40$, $\mu_{\min 1}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_1\mathbf{b}_1^+)\mathbf{A}) = 36$, $\mu_{\min 2}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_2\mathbf{b}_2^+)\mathbf{A}) = \mu_{\min 3}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_3\mathbf{b}_3^+)\mathbf{A}) = 40$. При этом значение $\lambda_{\min}(\mathbf{A}^{\mathbf{T}}\mathbf{A})$ имеет кратность, равную двум,

$$\zeta = \begin{bmatrix} \frac{2}{5} + \frac{2}{\sqrt{70}} \\ \frac{2}{5} - \frac{1}{\sqrt{70}} \\ \frac{2}{5} - \frac{2}{\sqrt{70}} \\ \frac{2}{5} - \frac{1}{\sqrt{70}} \\ \frac{2}{5} + \frac{2}{\sqrt{70}} \end{bmatrix} \approx \begin{bmatrix} 0,6390457 \\ 0,2804771 \\ 0,1609543 \\ 0,2804771 \\ 0,6390457 \end{bmatrix}$$

— определенный с точностью до знака единичный собственный вектор матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_1\mathbf{b}_1^+)\mathbf{A}$, соответствующий ее единственному минимальному собственному значению $\mu_{\min 1}$, т. е. $\zeta \in \mathscr{M}_{min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_1\mathbf{b}_1^+)\mathbf{A})$, причем $\mathbf{b}_1^{\mathbf{T}}\mathbf{A}\zeta = \frac{2016}{\sqrt{5}} \neq 0$; векторы

$$[\gamma\ \eta\ \theta] = \begin{bmatrix} \frac{1}{2}\cdot\left(\frac{5}{\sqrt{511}}+\frac{11}{\sqrt{730}}\right) - \sqrt{\frac{3}{10}} & \frac{\sqrt{3}}{2}\cdot\left(\frac{5}{\sqrt{511}}+\frac{11}{\sqrt{730}}\right) + \frac{1}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \\ \frac{1}{4}\cdot\left(\frac{-5}{\sqrt{511}}+\frac{22}{\sqrt{730}}\right) - \sqrt{\frac{3}{40}} & \frac{\sqrt{3}}{4}\cdot\left(\frac{-5}{\sqrt{511}}+\frac{22}{\sqrt{730}}\right) + \frac{1}{\sqrt{40}} & \frac{2}{\sqrt{10}} \\ \frac{1}{2}\cdot\left(\frac{-5}{\sqrt{511}}+\frac{11}{\sqrt{730}}\right) & \frac{\sqrt{3}}{2}\cdot\left(\frac{-5}{\sqrt{511}}+\frac{11}{\sqrt{730}}\right) & 0 \\ \frac{1}{4}\cdot\left(\frac{-5}{\sqrt{511}}+\frac{22}{\sqrt{730}}\right) + \sqrt{\frac{3}{40}} & \frac{\sqrt{3}}{4}\cdot\left(\frac{-5}{\sqrt{511}}+\frac{22}{\sqrt{730}}\right) - \frac{1}{\sqrt{40}} & \frac{-2}{\sqrt{10}} \\ \frac{1}{2}\cdot\left(\frac{5}{\sqrt{511}}+\frac{11}{\sqrt{730}}\right) + \sqrt{\frac{3}{10}} & \frac{\sqrt{3}}{2}\cdot\left(\frac{5}{\sqrt{511}}+\frac{11}{\sqrt{730}}\right) - \frac{1}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{bmatrix}$$

$$\approx \begin{bmatrix} -0,2335649 & 0,8603647 & -0,3162278 \\ -0,1255939 & 0,4149205 & 0,6324555 \\ 0,0929706 & 0,1610299 & 0 \\ 0,4221287 & 0,0986928 & -0,6324555 \\ 0,8618802 & 0,2279092 & 0,3162278 \end{bmatrix}$$

образуют ортогональный базис линейного подпространства собственных векторов матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}_2\mathbf{b}_2{}^{+})\mathbf{A}$, соответствующих ее минимальному собственному значению $\mu_{\min 2}$, имеющему кратность три, причем $\mathbf{b}_2^{\mathbf{T}}\mathbf{A}[\boldsymbol{\gamma}\ \boldsymbol{\eta}\ \boldsymbol{\theta}] = \left[336\sqrt{\frac{10}{73}}\ 336\sqrt{\frac{30}{73}}\ 0\right]$, т. е. существует $\mathbf{y} \in \mathscr{M}_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_2\mathbf{b}_2{}^{+})\mathbf{A})$ такой, что $\mathbf{b}_2^{\mathbf{T}}\mathbf{A}\mathbf{y} \neq 0$; векторы

$$[\nu\ \xi] = \begin{bmatrix} \frac{-1}{\sqrt{10}} & \frac{-2}{\sqrt{10}} \\ \frac{2}{\sqrt{10}} & \frac{-1}{\sqrt{10}} \\ 0 & 0 \\ \frac{-2}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{2}{\sqrt{10}} \end{bmatrix} \approx \begin{bmatrix} -0,3162278 & -0,6324555 \\ 0,6324555 & -0,3162278 \\ 0 & 0 \\ -0,6324555 & 0,3162278 \\ 0,3162278 & 0,6324555 \end{bmatrix}$$

образуют ортогональный базис линейного подпространства собственных векторов матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}_3\mathbf{b}_3{}^{+})\mathbf{A}$, соответствующих ее минимальному собственному значению $\mu_{\min 3}$, имеющему кратность два, причем $\mathbf{b}_3^{\mathbf{T}}\mathbf{A}[\nu\ \xi] = \mathbf{0}$. Поэтому не существует такой вектор $\mathbf{y} \in \mathscr{M}_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_3\mathbf{b}_3{}^{+})\mathbf{A})$, что $\mathbf{b}_3^{\mathbf{T}}\mathbf{A}\mathbf{y} \neq 0$.

Таким образом, из теоремы 2 заключаем, что для систем с правыми частями $\mathbf{b}_1$ и $\mathbf{b}_2$ нижняя грань нормы матрицы коррекции достигается, а для системы с правой частью $\mathbf{b}_3$ — не достигается. В то же время из сопоставления собственных значений матриц $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ и $\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}_i\mathbf{b}_i{}^{+})\mathbf{A}$ следует, что система с вектором $\mathbf{b}_1$ в правой части соответствует случаю, когда минимальное собственное значение $\mu_{\min 1}$ матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}_1\mathbf{b}_1{}^{+})\mathbf{A}$ не является собственным значением матрицы $\mathbf{A}^{\mathbf{T}}\mathbf{A}$; система с вектором $\mathbf{b}_2$ в правой части соответствует случаю, когда кратность значения $\mu_{\min 2}$ среди собственных значений матрицы $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ меньше его кратности среди собственных значений матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_2\mathbf{b}_2{}^{+})\mathbf{A}$; система с вектором $\mathbf{b}_3$ в правой части соответствует случаю, когда минимальное собственное значение $\mu_{\min 3}$ матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_3\mathbf{b}_3{}^{+})\mathbf{A}$ является также собственным значением матрицы $\mathbf{A}^{\mathbf{T}}\mathbf{A}$, причем его кратность для обеих матриц одинакова.

Следовательно, в рассматриваемом примере выводы теоремы 4 совпали с выводами теоремы 2.

Теперь покажем, что условие $m \geqslant n$ является необходимым для достижимости $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ в задаче коррекции матрицы коэффициентов несовместной системы вида (1), (3). Для этого рассмотрим следующее вспомогательное утверждение.

Лемма 3. *Условие* $\operatorname{rank}\ \mathbf{A} = n$ *является необходимым для того, чтобы достигалась нижняя грань нормы корректирующей матрицы в задаче коррекции несовместной системы* (1), (3).

Доказательство. Предположим противное. Пусть rank $\mathbf{A} < n$, но нижняя грань нормы корректирующей матрицы достигается. В силу условия rank $\mathbf{A} < n$ существует непустое линейное подпространство векторов, ортогональных строкам матрицы $\mathbf{A}$. Каждый вектор из этого подпространства является собственным вектором матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}$, соответствующий ее минимальному собственному значению $\mu_{\min} = 0$. Очевидно, что ни для одного из указанных векторов условие (8) не выполняется, что в силу теоремы 2 означает недостижимость нижней грани нормы корректирующей матрицы. Противоречие. Лемма 3 доказана.

Следствие 1. *Пусть $\mathbf{A} \in \mathbb{R}^{m\times n}$ и система* (1), (3) *несовместна. Тогда условие $m \geqslant n$ является необходимым для достижимости*

$$\inf_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|.$$

Доказательство. Пусть $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ достигается, но условие $m \geqslant n$ не выполняется, т. е. $m < n$. Тогда rank $\mathbf{A} < n$ и в силу леммы 3 $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ не достигается. Противоречие.

Заметим, что в общем случае решение скорректированной системы вида (1), (3) неединственно, в силу чего правомерно задать вопрос: какова общая структура множества решений $\mathscr{X}(\mathbf{A}, \mathbf{b}, \mathbf{H}^*)$ и какое место в ней занимает решение вида (7)?

Насколько известно автору, данные вопросы ранее не рассматривались. Прежде чем перейти к их обсуждению, обратимся к исследованию множества $\mathscr{X}(\mathbf{A}, \mathbf{b}, \widehat{\mathbf{H}})$ допустимых решений системы

$$(\mathbf{A} + \widehat{\mathbf{H}})\mathbf{x} = \mathbf{b}, \quad \mathbf{b} \neq \mathbf{0}, \tag{19}$$

возникающей в результате коррекции матрицы коэффициентов несовместной линейной системы вида (1), (3) с помощью одноранговой матрицы

$$\widehat{\mathbf{H}} = (\mathbf{b} - \mathbf{A}\mathbf{y})\mathbf{y}^{+}, \tag{20}$$

где $\mathbf{y} \neq \mathbf{0}$ — некоторый вектор,

$$\mathbf{y}^{+} = (\mathbf{y}^{\mathbf{T}}\mathbf{y})^{-1}\mathbf{y}^{\mathbf{T}} \tag{21}$$

— вектор, псевдообратный вектору $\mathbf{y}$. Формула (20) интересует нас в качестве более общего, чем (6), способа коррекции матриц коэффициентов несовместных систем вида (1), (3) и (2), (3), поскольку она применима и при невыполнении условий теоремы 1.

Теорема 5. *Пусть матрица коэффициентов несовместной системы (1), (3) скорректирована с помощью некоторой одноранговой матрицы вида (20). Тогда если*

$$\operatorname{rank}\begin{bmatrix}\mathbf{A}\\ \mathbf{y}^{\mathbf{T}}\end{bmatrix} < n, \tag{22}$$

то

$$\mathbf{x} \in \mathscr{X}(\mathbf{A}, \mathbf{b}, \widehat{\mathbf{H}}) \Leftrightarrow \mathbf{x} = \mathbf{y} + \Delta\mathbf{y}, \tag{23}$$

где

$$\Delta\mathbf{y} = \mathbf{P}_{\perp[\mathbf{y},\mathbf{A}^{\mathrm{T}}]} \cdot \mathbf{u}, \tag{24}$$

$\mathbf{P}_{\perp[\mathbf{y},\mathbf{A}^{\mathrm{T}}]}$ — *ортогональный проектор на подпространство векторов, ортогональных вектору* $\mathbf{y}$ *и строкам матрицы* $\mathbf{A}$, $\mathbf{u} \in \mathbb{R}^n$ — *произвольный вектор, в противном случае* $\mathscr{X}(\mathbf{A}, \mathbf{b}, \widehat{\mathbf{H}}) \equiv \{\mathbf{y}\}$.

Доказательство. Сначала покажем, что формула из правой части соотношения (23) вместе с формулой (24) действительно описывает решение системы (19) при *любом* векторе $\mathbf{u}$.

Вначале убедимся, что равенства

$$\mathbf{A} \cdot \Delta\mathbf{y} = \mathbf{0}, \tag{25}$$

$$\mathbf{y}^{\mathbf{T}}\Delta\mathbf{y} = 0 \tag{26}$$

выполняются при любом векторе $\mathbf{u}$.

Действительно, в силу (24) имеем

$$\begin{bmatrix}\mathbf{A}\\ \mathbf{y}^{\mathbf{T}}\end{bmatrix} \cdot \Delta\mathbf{y} = \begin{bmatrix}\mathbf{A}\\ \mathbf{y}^{\mathbf{T}}\end{bmatrix} \cdot \mathbf{P}_{\perp[\mathbf{y},\mathbf{A}^{\mathrm{T}}]} \cdot \mathbf{u} = \mathbf{0}.$$

Теперь выполним элементарные преобразования

$$(\mathbf{A} + (\mathbf{b} - \mathbf{A}\mathbf{y})\mathbf{y}^+)\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{b}\mathbf{y}^+\mathbf{x} - \mathbf{A}\mathbf{y}\mathbf{y}^+\mathbf{x}$$

и рассмотрим каждое слагаемое отдельно. В силу правой части из (23) и формулы (25) имеем

$$\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{y} + \mathbf{A} \cdot \Delta\mathbf{y} = \mathbf{A}\mathbf{y}.$$

В то же время в силу правой части из (23), формул (21) и (26) имеем

$$\mathbf{b}\mathbf{y}^+\mathbf{x} = \mathbf{b}\mathbf{y}^+\mathbf{y} + \mathbf{b}\mathbf{y}^+\Delta\mathbf{y} = \mathbf{b}.$$

Осталось упростить выражение $\mathbf{A}\mathbf{y}\mathbf{y}^+\mathbf{x}$. Пользуясь правой частью соотношения (23), формулами (21) и (26), получаем

$$\mathbf{A}\mathbf{y}\mathbf{y}^+\mathbf{x} = \mathbf{A}\mathbf{y}\mathbf{y}^+\mathbf{y} + \mathbf{A}\mathbf{y}\mathbf{y}^+\Delta\mathbf{y} = \mathbf{A}\mathbf{y}.$$

Таким образом,

$$(\mathbf{A} + (\mathbf{b} - \mathbf{A}\mathbf{y})\mathbf{y}^+)\mathbf{x} = \mathbf{A}\mathbf{y} + \mathbf{b} - \mathbf{A}\mathbf{y} \equiv \mathbf{b}.$$

Теперь покажем, что *любое* решение системы (19) имеет вид, задаваемый правой частью соотношения (23) и формулой (24). Действительно, если $\mathbf{x}$ — решение системы (19), то должно выполняться тождество

$$(\mathbf{A} + (\mathbf{b} - \mathbf{A}\mathbf{y})\mathbf{y}^{+})\mathbf{x} = \mathbf{b}.$$

Выполним некоторые преобразования:

$$(\mathbf{A} + (\mathbf{b} - \mathbf{A}\mathbf{y})\mathbf{y}^{+})\mathbf{x} = \mathbf{b} \Leftrightarrow \mathbf{A}\mathbf{x} + \mathbf{b}\mathbf{y}^{+}\mathbf{x} - \mathbf{A}\mathbf{y}\mathbf{y}^{+}\mathbf{x} = \mathbf{b} \Leftrightarrow$$
$$\mathbf{A}(\mathbf{I} - \mathbf{y}\mathbf{y}^{+})\mathbf{x} = (1 - \mathbf{y}^{+}\mathbf{x}) \cdot \mathbf{b} \Rightarrow \mathbf{y}^{+}\mathbf{x} = 1. \quad (27)$$

Последнее утверждение доказывается от противного. Пусть $\mathbf{y}^{+}\mathbf{x} \neq 1$. Тогда вектор $(1 - \mathbf{y}^{+}\mathbf{x})^{-1} \cdot (\mathbf{I} - \mathbf{y}\mathbf{y}^{+})\mathbf{x}$ является решением системы (1), (3), которая по предположению несовместна. Противоречие. Несложно видеть, что повторное использование цепочки (27) вместе с соотношением $\mathbf{y}^{+}\mathbf{x} = 1$ и формулой (21) приводит к однородной системе линейных уравнений

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{y}^{\mathrm{T}} \end{bmatrix} \cdot (\mathbf{x} - \mathbf{y}) = \mathbf{0}. \quad (28)$$

Но решение системы (28) имеет вид (см., например, [1]) $\mathbf{x} - \mathbf{y} = \Delta\mathbf{y}$, где $\Delta\mathbf{y}$ — вектор, определяемый соотношением (24).

Наконец, покажем, что решение системы (19) единственно *тогда и только тогда*, когда $\operatorname{rank}\begin{bmatrix} \mathbf{A} \\ \mathbf{y}^{\mathrm{T}} \end{bmatrix} = n$. Действительно, в силу правой части из (23) и формулы (24) решение системы (19) единственно тогда и только тогда, когда выполняется условие $\mathbf{P}_{\perp[\mathbf{y},\mathbf{A}^{\mathrm{T}}]} \equiv \mathbf{0}$, которое (см., например, [12]) выполняется тогда и только тогда, когда $\operatorname{rank}\begin{bmatrix} \mathbf{A} \\ \mathbf{y}^{\mathrm{T}} \end{bmatrix} = n$. Теорема 5 доказана.

Следствие 2.

$$\mathbf{y} = \operatorname*{argmin}_{\mathscr{X}(\mathbf{A},\mathbf{b},\widehat{\mathbf{H}})\neq\varnothing} \{\|\mathbf{x}\|\},$$

т. е. среди всевозможных решений системы (19) *вектор* $\mathbf{y}$ (*и только он*) *имеет минимальную евклидову норму.*

Доказательство. Из (26) следует, что $\Delta\mathbf{y}\perp\mathbf{y}$. Вместе с правой частью соотношения (23) это дает возможность использовать теорему Пифагора $\|\mathbf{x}\|^2 = \|\mathbf{y}\|^2 + \|\Delta\mathbf{y}\|^2$. Следовательно, $\|\mathbf{x}\| \geqslant \|\mathbf{y}\|$, причем равенство достигается тогда и только тогда, когда $\Delta\mathbf{y} = \mathbf{0}$. Следствие 2 доказано.

Следствие 3. *Если* $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ *достигается и*

$$\mathbf{H}^{*} \in \operatorname*{Arginf}_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|,$$

то множество $\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H}^*)$ *состоит из единственного элемента:*

$$\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H}^*) = \Big\{\mathbf{x}^* = \frac{\mathbf{b^T b}}{\mathbf{b^T A y}}\cdot \mathbf{y}\Big\},$$

где $\mathbf{y}\in\mathscr{M}_{\min}(\mathbf{A^T}(\mathbf{I}-\mathbf{bb}^+)\mathbf{A})$.

Доказательство. Согласно лемме 3 условие $\operatorname{rank}\mathbf{A}=n$ является необходимым для того, чтобы нижняя грань нормы корректирующей матрицы достигалась. Но для любых $\mathbf{A}\in\mathbb{R}^{m\times n}$ и $\mathbf{y}\in\mathbb{R}^n$ справедливо соотношение $\operatorname{rank}\mathbf{A}=n \Rightarrow \operatorname{rank}\begin{bmatrix}\mathbf{A}\\ \mathbf{y^T}\end{bmatrix}=n$. Отсюда и из теоремы 5 следует, что множество $\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H}^*)$ состоит из единственного вектора. Вместе с тем вид этого вектора определяется теоремой 1. Следствие 3 доказано.

Замечание. Теорема 5 допускает естественное обобщение на случай коррекции несовместной системы вида (2), (3). Это обобщение будет рассмотрено в следующем разделе вместе с соответствующим численным примером.

Теперь рассмотрим случай, когда в задаче построения оптимальной одноранговой матрицы коррекции матрицы коэффициентов несовместной системы (1), (3) $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|$ достигается, но решение задачи $\|\mathbf{H}\|\to\min$, $\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing$ *не единственно.* В этом случае возникает вопрос о структуре множества $\operatorname*{Arginf}\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|$. Для того чтобы ответить на данный вопрос, предварительно убедимся в справедливости следующего утверждения.

Лемма 4. *Если* $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|$ *достигается, то существует такой собственный вектор* $\tilde{\mathbf{y}}$ *матрицы* $\mathbf{A^T}(\mathbf{I}-\mathbf{bb}^+)\mathbf{A}$, *соответствующий ее минимальному собственному значению* $\mu_{\min}$, *что выполняется условие*

$$(\mathbf{b^T b})^{-1}\cdot\mathbf{b^T A}\tilde{\mathbf{y}} = 1. \tag{29}$$

При этом

$$\Big\{\widetilde{\mathbf{H}} = (\mathbf{b}-\mathbf{A}\tilde{\mathbf{y}})\tilde{\mathbf{y}}^+ \equiv \mathbf{H}^* = -(\mathbf{I}-\mathbf{bb}^+)\mathbf{Ayy^T}\Big\} \in \operatorname{Arg}\inf_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|, \tag{30}$$

$$\mathscr{X}(\mathbf{A},\mathbf{b},\widetilde{\mathbf{H}}) = \mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H}^*) = \Big\{\tilde{\mathbf{y}}\equiv\mathbf{x}^* = \frac{\mathbf{b^T b}}{\mathbf{b^T A y}}\cdot\mathbf{y}\Big\}, \tag{31}$$

где

$$\tilde{\mathbf{y}}^+ = (\tilde{\mathbf{y}}^{\mathbf{T}}\tilde{\mathbf{y}})^{-1}\tilde{\mathbf{y}}^{\mathbf{T}} \tag{32}$$

— *вектор, псевдообратный вектору* $\tilde{\mathbf{y}}$.

Доказательство. Пусть $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ достигается. Тогда в силу теоремы 2 $\exists \mathbf{y} \in \mathscr{M}_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}) \mid \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y} \neq 0$. Пусть

$$\tilde{\mathbf{y}} \equiv \mathbf{x}^* = \frac{\mathbf{b}^{\mathbf{T}}\mathbf{b}}{\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}} \cdot \mathbf{y}. \tag{33}$$

Поскольку в этом случае вектор $\tilde{\mathbf{y}}$ отличается от вектора $\mathbf{y}$ только скалярным множителем, $\tilde{\mathbf{y}}$ также является собственным вектором матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}$, соответствующим ее минимальному собственному значению. Непосредственной подстановкой легко убедиться в том, что этот вектор удовлетворяет условию (29).

Теперь убедимся в том, что матрица

$$\widetilde{\mathbf{H}} = (\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}})\tilde{\mathbf{y}}^{+} \tag{34}$$

корректирует систему (1), (3). Действительно, из (32)–(34) следует, что

$$(\mathbf{A} + \widetilde{\mathbf{H}})\tilde{\mathbf{y}} = \Big(\mathbf{A} + (\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}})\frac{\tilde{\mathbf{y}}^{\mathbf{T}}}{\tilde{\mathbf{y}}^{\mathbf{T}}\tilde{\mathbf{y}}}\Big)\tilde{\mathbf{y}} \equiv \mathbf{b}.$$

Покажем, что $\widetilde{\mathbf{H}} \in \operatorname*{Arginf}\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$. Согласно теореме 1 для этого достаточно убедиться в справедливости условия

$$\|\widetilde{\mathbf{H}}\|^2 = \mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}). \tag{35}$$

Действительно, в силу (34) имеем

$$\|\widetilde{\mathbf{H}}\|^2 = \|(\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}})\tilde{\mathbf{y}}^{+}\|^2.$$

Поскольку матрица $\widetilde{\mathbf{H}}$ является одноранговой, для ее евклидовой нормы справедливо (см., например, [10]) соотношение

$$\|(\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}})\tilde{\mathbf{y}}^{+}\|^2 = \|(\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}})\|^2 \cdot \|\tilde{\mathbf{y}}^{+}\|^2,$$

из которого с учетом (32) получаем

$$\|(\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}})\tilde{\mathbf{y}}^{+}\|^2 = \frac{\|\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}}\|^2}{\|\tilde{\mathbf{y}}\|^2}. \tag{36}$$

Рассмотрим числитель полученного выражения. Согласно (18) имеем

$$\|\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}}\|^2 = (\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}})^{\mathbf{T}}(\mathbf{b} - \mathbf{A}\tilde{\mathbf{y}}) = \mathbf{b}^{\mathbf{T}}\mathbf{b} - 2\mathbf{b}^{\mathbf{T}}\mathbf{A}\tilde{\mathbf{y}} + \tilde{\mathbf{y}}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{A}\tilde{\mathbf{y}} =$$
$$\Big(\mathbf{b}^{\mathbf{T}}\mathbf{b} - 2\mathbf{b}^{\mathbf{T}}\mathbf{A}\tilde{\mathbf{y}} + \frac{\tilde{\mathbf{y}}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{b}\mathbf{b}^{\mathbf{T}}\mathbf{A}\tilde{\mathbf{y}}}{\mathbf{b}^{\mathbf{T}}\mathbf{b}}\Big) + \tilde{\mathbf{y}}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}\tilde{\mathbf{y}}.$$

Но в силу (29)

$$\mathbf{b}^{\mathbf{T}}\mathbf{b} - 2\mathbf{b}^{\mathbf{T}}\mathbf{A}\tilde{\mathbf{y}} + \frac{\tilde{\mathbf{y}}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{b}\mathbf{b}^{\mathbf{T}}\mathbf{A}\tilde{\mathbf{y}}}{\mathbf{b}^{\mathbf{T}}\mathbf{b}} \equiv 0,$$

а в силу (33)

$$\tilde{\mathbf{y}}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\tilde{\mathbf{y}} = \mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A})\cdot\|\tilde{\mathbf{y}}\|^2.$$

Отсюда и из (36) следует выполнение условия (35), т. е.

$$\widetilde{\mathbf{H}} \in \operatorname{Arg}\inf_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|.$$

Покажем, что $\widehat{\mathbf{H}} \equiv \mathbf{H}^*$. Для этого убедимся в справедливости тождества

$$-(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{y}\mathbf{y}^{\mathbf{T}} \equiv (\mathbf{b}-\mathbf{A}\tilde{\mathbf{y}})\tilde{\mathbf{y}}^{+}.$$

Действительно, в силу (5), (32) и (33)

$$-(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{y}\mathbf{y}^{\mathbf{T}} = \mathbf{b}\cdot\mathbf{b}^{+}\mathbf{A}\mathbf{y}\cdot\mathbf{y}^{\mathbf{T}} - \mathbf{A}\mathbf{y}\mathbf{y}^{\mathbf{T}} = \mathbf{b}\tilde{\mathbf{y}}^{+} - \mathbf{A}\tilde{\mathbf{y}}\tilde{\mathbf{y}}^{+} = (\mathbf{b}-\mathbf{A}\tilde{\mathbf{y}})\tilde{\mathbf{y}}^{+}.$$

Теперь несложно заметить, что истинность утверждения (31) следует из следствия 3. Лемма 4 доказана.

Лемма 4 открывает возможность установления основного в данном разделе результата — общего вида множества оптимальных одноранговых матриц коррекции для матриц коэффициентов несовместных систем (1), (3) и соответствующего множества решений скорректированных систем в случае, когда решение $\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing$ задачи $\|\mathbf{H}\|\to\min$ не единственно.

Теорема 6. *Пусть в задаче коррекции матрицы коэффициентов несовместной системы* (1), (3) $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|$ *достигается, причем кратность минимального собственного значения* $\mu_{\min}$ *матрицы* $\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}$ *равна* $p>1$. *Пусть при этом* $\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_p$ *— некоторый набор линейно независимых векторов таких, что*

$$\mathbf{y}_i \in \mathscr{M}_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}),\quad \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}_i\neq 0,\quad i=1,2,\ldots,r, \tag{37}$$

$$\mathbf{y}_i \in \mathscr{M}_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}),\quad \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}_i = 0,\quad i=r{+}1,r{+}2,\ldots,p, \tag{38}$$

где $1<r\leqslant p$ *и запись* $i=r{+}1,r{+}2,\ldots,p$ *при* $r=p$ *означает пустое множество индексов. Тогда*

1. *Существует параметрическое семейство оптимальных одноранговых матриц коррекции вида*

$$\left\{\mathbf{H}^{**}=(\mathbf{b}-\mathbf{A}\tilde{\mathbf{y}}^{*})\tilde{\mathbf{y}}^{*+}\right\} \in \operatorname*{Arginf}_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|, \tag{39}$$

где

$$\left\{\tilde{\mathbf{y}}^{*}=\sum_{i=1}^{r} d_i\cdot\frac{\mathbf{b}^{\mathbf{T}}\mathbf{b}}{\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}_i}\cdot\mathbf{y}_i+\sum_{i=r+1}^{p} d_i\cdot\mathbf{y}_i\right\}=\{\mathbf{x}^{**}\}=\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H}^{**}), \tag{40}$$

т. е. $\tilde{\mathbf{y}}^*$ *и* $\mathbf{x}^{**}$ *— синонимы для обозначения единственного решения скорректированной системы* (1), (3) *при фиксированной матрице* $\mathbf{H}^{**}$, *вектор*

$$\tilde{\mathbf{y}}^{*+} = \left(\tilde{\mathbf{y}^*}^{\mathbf{T}}\tilde{\mathbf{y}}^*\right)^{-1}\tilde{\mathbf{y}^*}^{\mathbf{T}} \tag{41}$$

псевдообратен вектору $\tilde{\mathbf{y}}^*$, *и*

$$\sum_{i=1}^{r} d_i = 1, \tag{42}$$

$$|d_i| < \infty, \quad i = r{+}1, r{+}2, \ldots, p. \tag{43}$$

2. *Любая другая одноранговая матрица, не принадлежащая семейству матриц вида* (37)–(43), *не принадлежит множеству* $\operatorname*{Arginf}_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|$.

Доказательство. 1. В силу линейной независимости векторов $\mathbf{y}_1, \ldots, \mathbf{y}_r$ и формул (37), (38), (40), (42), (43) выполняется условие $\tilde{\mathbf{y}}^* \neq \mathbf{0}$. Непосредственная проверка показывает, что вектор $\tilde{\mathbf{y}}^*$ является собственным вектором матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}$, соответствующим ее минимальному собственному значению $\mu_{\min}$, и удовлетворяет условию (29). При этом несложно заметить, что формулы (39) и (41), определяющие матрицу $\mathbf{H}^{**}$, аналогичны формулам (34) и (32), что позволяет использовать лемму 4.

2. Пусть

$$\mathscr{H} \in \operatorname*{Arginf}_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|, \tag{44}$$

$$\operatorname{rank} \mathscr{H} = 1 \tag{45}$$

и матрица $\mathscr{H}$ не представима в виде (37)–(42).

Для дальнейших рассуждений рассмотрим представление

$$\mathscr{H} = \mathbf{H}^{**} + \Delta\mathbf{H}, \tag{46}$$

где матрица $\mathbf{H}^{**}$ вместе с соответствующим ей вектором $\tilde{\mathbf{y}}^*$ определяются формулами (37)–(43), причем

$$\Delta\mathbf{H} \neq \mathbf{0}. \tag{47}$$

Пусть

$$\mathbf{z} = \mathbf{b} - \mathbf{A}\tilde{\mathbf{y}}^*, \tag{48}$$

$$\mathbf{z}^+ = \frac{\mathbf{z}^{\mathbf{T}}}{\mathbf{z}^{\mathbf{T}}\mathbf{z}} \tag{49}$$

— вектор, псевдообратный вектору $\mathbf{z}$ (заметим, что $\mathbf{z} \neq \mathbf{0}$ в силу несовместности системы (1), (3)),

$$\Delta\mathbf{H}_1 = \mathbf{z}\mathbf{z}^+ \cdot \Delta\mathbf{H}, \tag{50}$$

$$\Delta\mathbf{H}_2 = (\mathbf{I} - \mathbf{z}\mathbf{z}^+) \cdot \Delta\mathbf{H}. \tag{51}$$

Кроме того, из (39) и (48) имеем

$$\mathbf{H}^{**} = \mathbf{z} \cdot \tilde{\mathbf{y}}^{*+}. \tag{52}$$

С точностью до скалярных множителей столбцы матрицы $\Delta\mathbf{H}_1$ совпадают со столбцами матрицы $\mathbf{H}^{**}$, т. е.

$$\Delta\mathbf{H}_1 = \mathbf{z} \cdot \mathbf{c}^{\mathbf{T}}, \tag{53}$$

где $\mathbf{c} \in \mathbb{R}^n$ и $\mathbf{c}^{\mathbf{T}} = \mathbf{z}^+\Delta\mathbf{H}$. В свою очередь, из формул (48), (49), (51) следует, что столбцы матрицы $\Delta\mathbf{H}_2$ ортогональны вектору $\mathbf{z}$ и, как следствие, столбцам матрицы $\mathbf{H}^{**}$.

В силу (44), (45), совпадения спектральной и евклидовой норм для одноранговой матрицы, известного (см., например, [13]) соотношения для спектральной нормы матрицы, а также формул (50), (51) можно записать

$$\|\mathscr{H}\| = \mu_{\min}^{\frac{1}{2}}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^+)\mathbf{A}) = \max_{\|\mathbf{p}\|=\|\mathbf{q}\|=1} \left|\mathbf{p}^{\mathbf{T}}(\mathbf{H}^{**} + \Delta\mathbf{H}_1 + \Delta\mathbf{H}_2)\mathbf{q}\right|. \tag{54}$$

Как уже отмечалось выше, $\mathbf{z} \neq \mathbf{0}$ и $\tilde{\mathbf{y}}^* \neq \mathbf{0}$. Поэтому можно положить

$$\mathbf{p} = \frac{\mathbf{z}}{\|\mathbf{z}\|}, \quad \mathbf{q} = \frac{\tilde{\mathbf{y}}^*}{\|\tilde{\mathbf{y}}^*\|}.$$

Тогда из (4), (39), (48)–(51) и (53) имеем

$$\mathbf{p}^{\mathbf{T}}(\mathbf{H}^{**} + \Delta\mathbf{H}_1 + \Delta\mathbf{H}_2)\mathbf{q} = \mu_{\min}^{\frac{1}{2}}\Big(1 + \frac{\|\mathbf{z}\|}{\|\tilde{\mathbf{y}}^*\|} \cdot \mathbf{c}^{\mathbf{T}}\tilde{\mathbf{y}}^*\Big).$$

Отсюда в силу (54) следует, что

$$\mathbf{c}^{\mathbf{T}}\tilde{\mathbf{y}}^* = 0, \tag{55}$$

т. е. векторы $\mathbf{c}$ и $\tilde{\mathbf{y}}^*$ ортогональны.

Теперь воспользуемся формулами (48)–(53) для того, чтобы формулу (46) переписать в виде

$$\mathscr{H} = \mathbf{z}\,(\tilde{\mathbf{y}}^{*+} + \mathbf{c}^{\mathbf{T}}) + \Delta\mathbf{H}_2. \tag{56}$$

Но в силу (45), (52) и (56) имеем

$$\|\mathscr{H}\|^2 = \|\mathbf{z}\|^2\left(\|\tilde{\mathbf{y}}^*\|^2 + \|\mathbf{c}\|^2\right) + \|\Delta\mathbf{H}_2\|^2 = \|\mathbf{H}^{**}\|^2 + \|\mathbf{z}\|^2 \cdot \|\mathbf{c}\|^2 + \|\Delta\mathbf{H}_2\|^2. \tag{57}$$

Здесь величина $\|\mathscr{H}\|^2$ вычисляется как квадрат евклидовой нормы с использованием соотношения $\|\mathbf{z}\,(\tilde{\mathbf{y}}^{*+} + \mathbf{c}^{\mathbf{T}})\|^2 = \|\mathbf{z}\|^2 \cdot \|\tilde{\mathbf{y}}^{*+} + \mathbf{c}^{\mathbf{T}}\|^2$. Оно справедливо для евклидовой нормы одноранговой матрицы с учетом ортогональности векторов $\tilde{\mathbf{y}}^*$ и $\mathbf{c}$ (соотношение (55)) и ортогональности столбцов матрицы $\Delta\mathbf{H}_2$ столбцам матрицы $\mathbf{z}\,(\tilde{\mathbf{y}}^{*+} + \mathbf{c}^{\mathbf{T}})$ (соотношение (51)).

С учетом (54) и (57) получаем

$$\mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}) = \mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}) + \|\mathbf{z}\|^2 \cdot \|\mathbf{c}\|^2 + \|\Delta\mathbf{H}_2\|^2$$
$$\Leftrightarrow \|\mathbf{z}\|^2 \cdot \|\mathbf{c}\|^2 + \|\Delta\mathbf{H}_2\|^2 = 0.$$

Но $\mathbf{z} \neq \mathbf{0}$ в силу несовместности системы (1), (3). Следовательно, $\|\mathbf{z}\|^2 > 0$ и мы приходим к условиям $\mathbf{c} \equiv \mathbf{0}$ и $\Delta\mathbf{H}_2 \equiv \mathbf{0}$, откуда в силу (50), (51) и (53) следует, что $\Delta\mathbf{H} \equiv \mathbf{0}$. Противоречие.

Пример 2. Обратимся к несовместной системе линейных уравнений $\mathbf{A}\mathbf{x} = \mathbf{b}_2$, рассмотренной в примере 1. Выше было показано, что для указанной системы $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b}_2,\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ достигается, при этом $\boldsymbol{\gamma}, \boldsymbol{\eta}, \boldsymbol{\theta} \in \mathscr{M}_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_2\mathbf{b}_2{}^{+})\mathbf{A})$, $\mathbf{b}_2^{\mathbf{T}}\mathbf{A}\boldsymbol{\gamma} \neq 0$, $\mathbf{b}_2^{\mathbf{T}}\mathbf{A}\boldsymbol{\eta} \neq 0$, $\mathbf{b}_2\mathbf{A}\boldsymbol{\theta} = 0$. В силу теоремы 1 векторы $\boldsymbol{\gamma}$ и $\boldsymbol{\eta}$ могут служить основой для построения оптимальных корректирующих матриц $\mathbf{H}^*(\boldsymbol{\gamma})$ и $\mathbf{H}^*(\boldsymbol{\eta})$ и соответствующих им решений скорректированных систем $\boldsymbol{\gamma}^*$ и $\boldsymbol{\eta}^*$:

$$\boldsymbol{\gamma}^* = \frac{\mathbf{b}_2^{\mathbf{T}}\mathbf{b}_2}{\mathbf{b}_2^{\mathbf{T}}\mathbf{A}\boldsymbol{\gamma}} \cdot \boldsymbol{\gamma} = \frac{42}{55} \cdot \sqrt{730} \cdot \boldsymbol{\gamma} \approx \begin{bmatrix} -4,818986 \\ -2,591293 \\ 1,918200 \\ 8,709493 \\ 17,782586 \end{bmatrix},$$

$$\boldsymbol{\eta}^* = \frac{\mathbf{b}_2^{\mathbf{T}}\mathbf{b}_2}{\mathbf{b}_2^{\mathbf{T}}\mathbf{A}\boldsymbol{\eta}} \cdot \boldsymbol{\eta} = \frac{14}{55} \cdot \sqrt{2190} \cdot \boldsymbol{\eta} \approx \begin{bmatrix} 10,248729 \\ 4,942564 \\ 1,918200 \\ 1,175636 \\ 2,714871 \end{bmatrix}.$$

В то же время в силу теоремы 6 любой вектор $\mathbf{x}^{**}$ вида

$$\mathbf{x}^{**} = d_1 \cdot \boldsymbol{\gamma}^* + d_2 \cdot \boldsymbol{\eta}^* + d_3 \cdot \boldsymbol{\theta},$$

где $d_1 + d_2 = 1$, $|d_3| < \infty$, может служить как основой для построения оптимальной корректирующей матрицы, так и решением соответствующей скорректированной системы. Пусть, например,

$$d_1 = \frac{1}{4} + \sqrt{\frac{3}{1168}} \approx 0,3006803; \quad d_2 = \frac{3}{4} - \sqrt{\frac{3}{1168}} \approx 0,6993197;$$
$$d_3 = \frac{21}{55} \cdot \sqrt{10} \approx 1,207415.$$

Тогда

$$\mathbf{x}^{**} = \begin{bmatrix} \frac{168}{55} + \sqrt{\frac{630}{121}} \\ \frac{252}{55} - \sqrt{\frac{315}{242}} \\ \frac{21}{5} - \sqrt{\frac{630}{121}} \\ \frac{42}{11} - \sqrt{\frac{315}{242}} \\ \frac{294}{55} + \sqrt{\frac{630}{121}} \end{bmatrix} \approx \begin{bmatrix} 5,336346 \\ 3,440918 \\ 1,918200 \\ 2,677282 \\ 7,627255 \end{bmatrix}.$$

Средствами компьютерной алгебры [9] можно показать, что

$$(\mathbf{b}_2 - \mathbf{A}\mathbf{x}^{**}) = \frac{1}{11} \cdot \begin{bmatrix} -154 - 10\sqrt{70} + 2\sqrt{3927} \\ 112 + 10\sqrt{70} - 4\sqrt{3927} \\ 266 - 10\sqrt{70} + 2\sqrt{3927} \\ -350 + 10\sqrt{70} + 2\sqrt{3927} \\ -112 - 10\sqrt{70} - 4\sqrt{3927} \\ 238 + 10\sqrt{70} + 2\sqrt{3927} \end{bmatrix}, \ \|\mathbf{b}_2 - \mathbf{A}\mathbf{x}^{**}\|^2 = \frac{529200}{121},$$

$$\|\mathbf{x}^{**}\|^2 = \frac{13230}{121}, \ \|\mathbf{H}^{**}\|^2 = \frac{\|\mathbf{b}_2 - \mathbf{A}\mathbf{x}^{**}\|^2}{\|\mathbf{x}^{**}\|^2} = \mu_{\min 2}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}_2\mathbf{b}_2{}^{+})\mathbf{A}) = 40,$$

причем матрица $\mathbf{H}^{**} = (\mathbf{b}_2 - \mathbf{A}\mathbf{x}^{**})\mathbf{x}^{**+}$ может быть также получена средствами компьютерной алгебры. Но поскольку ее элементы не являются рациональными числами, запись матрицы $\mathbf{H}^{**}$ оказывается достаточно громоздкой. По этой причине приведем только приближенные значения ее элементов, что вполне достаточно для численной проверки ее оптимальности и совместности соответствующей скорректированной системы:

$$\mathbf{H}^{**} \approx \begin{bmatrix} -0,4984134 & -0,3213809 & -0,1791594 & -0,2500575 & -0,7123838 \\ -0,2440151 & -0,1573429 & -0,0877135 & -0,1224241 & -0,3487715 \\ 1,3650720 & 0,8802095 & 0,4906882 & 0,6848663 & 1,9511020 \\ -0,6256096 & -0,4033981 & -0,2248813 & -0,3138727 & -0,8941857 \\ -1,9803050 & -1,2769160 & -0,7118393 & -0,9935327 & -2,8304550 \\ 1,9832700 & 1,2788290 & 0,7129053 & 0,9950206 & 2,8346940 \end{bmatrix}.$$

Следующий результат, возможно, будет востребованным при рассмотрении задачи об оптимальной коррекции матрицы коэффициентов несовместной системы вида (2), (3). Мы рассматриваем его пока в разделе, в котором изучаются системы вида (1), (3). Исследуется нижняя оценка величины $\mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A})$, а точнее, условия, при которых $\mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}) = 0$. При этом допускается, что система вида (1),

[9] Автор использовал Mathcad 7.0 Pro для вычислений без преобразования рациональных чисел и квадратных корней в формат чисел с плавающей запятой.

(3) может быть совместна. Очевидно, что данная ситуация может возникнуть при несовместности системы (2), (3), вызванной тем, что среди решений системы (1), (3) нет неотрицательных решений.

Случай $\mu_{\min}(\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A})=0$ еще интересен тем, что, как нетрудно показать, при этом условии $\inf\limits_{\mathscr{X}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|$ не достигается.

Лемма 5. *Для выполнения условия* $\mu_{\min}(\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A})=0$ *необходимо и достаточно, чтобы выполнялось хотя бы одно из следующих условий*: 1) *система* (1), (3) *совместна;* 2) $\operatorname{rank}\mathbf{A}<n$.

Доказательство. **Необходимость.** Пусть $\mu_{\min}(\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A})=0$, но система (1), (3) несовместна, $\operatorname{rank}\mathbf{A}=n$ и $\mathbf{y}$ — собственный вектор матрицы $\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}$, соответствующий $\mu_{\min}$. Рассмотрим две возможности.

1. Пусть $\mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{y}\neq 0$. Ясно, что

$$\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}=((\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A})^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}=\mathbf{V}^{\mathrm{T}}\mathbf{V},$$

т. е. $\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}$ — матрица Грама столбцов матрицы $\mathbf{V}=(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}$. Разложим вектор $\mathbf{b}$ на две ортогональные составляющие:

$$\mathbf{b}=\mathbf{A}\mathbf{A}^{+}\mathbf{b}+(\mathbf{I}-\mathbf{A}\mathbf{A}^{+})\mathbf{b}=\mathbf{b}_1+\mathbf{b}_2.$$

Заметим, что $\mathbf{b}_1\neq\mathbf{0}$ в силу условия $\mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{y}\neq 0$, а $\mathbf{b}_2\neq\mathbf{0}$ в силу несовместности системы (1). Следовательно, матрица $\mathbf{V}$ представима в виде

$$\mathbf{V}=\Big(\mathbf{A}-\mathbf{b}_1\cdot\frac{\mathbf{b}_1^{\mathrm{T}}\mathbf{A}}{\mathbf{b}^{\mathrm{T}}\mathbf{b}}\Big)-\Big(\mathbf{b}_2\cdot\frac{\mathbf{b}_1^{\mathrm{T}}\mathbf{A}}{\mathbf{b}^{\mathrm{T}}\mathbf{b}}\Big)=\mathbf{V}_1-\mathbf{V}_2.$$

Вследствие условия $\operatorname{rank}\mathbf{A}=n$ и результатов [9] матрица $\mathbf{V}_1$ невырождена, причем ее псевдообратная матрица может быть построена с помощью обобщенной формулы Шермана — Моррисона:

$$\mathbf{V}_1^{+}=\mathbf{A}^{+}+\frac{\mathbf{A}^{+}\mathbf{b}_1\mathbf{b}_1^{\mathrm{T}}}{\mathbf{b}_2^{\mathrm{T}}\mathbf{b}_2}. \tag{58}$$

Учитывая (58), можно записать

$$\mathbf{V}_1^{+}\mathbf{V}=\mathbf{V}_1^{+}\mathbf{V}_1=\mathbf{I}.$$

Отсюда с использованием неравенства Сильвестра следует, что

$$\operatorname{rank}\mathbf{V}=\operatorname{rank}\mathbf{V}_1=n,$$

т. е. матрица $\mathbf{V}$ имеет полный ранг, что является необходимым и достаточным условием (см., например, [13]) для того, чтобы матрица $\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}$ (как матрица Грама столбцов матрицы $\mathbf{V}$) была положительно определена, в силу чего $\mu_{\min}>0$. Противоречие.

2. Пусть $\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y} = 0$. В этом случае по лемме 1 вектор $\mathbf{y}$ является собственным вектором матрицы $\mathbf{A}^{\mathbf{T}}\mathbf{A}$, а число $\mu_{\min}$ — ее собственным значением. Но матрица $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ является матрицей Грама столбцов матрицы $\mathbf{A}$. Поэтому, как уже упоминалось выше, условие $\operatorname{rank}\,\mathbf{A} = n$ является необходимым и достаточным для того, чтобы матрица $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ была положительно определена. Следовательно, она не может иметь нулевых собственных значений. Противоречие.

Достаточность. 1. Пусть система (1), (3) совместна. Эта система может рассматриваться как система с матрицей $\mathbf{A} + \mathbf{H}$, где $\mathbf{H} \equiv \mathbf{0}$. В силу свойств евклидовой или спектральной нормы из $\mathbf{H} = \mathbf{0}$ следует $\|\mathbf{H}\| = 0$, откуда из теоремы 1 получаем $\mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}) = 0$.

2. Пусть $\operatorname{rank}\,\mathbf{A} < n$. В этом случае существует такой вектор $\mathbf{x}$, что $\mathbf{x} \neq \mathbf{0}$, $\mathbf{A}\mathbf{x} = \mathbf{0}$, т. е. $\mathbf{x}$ — собственный вектор матрицы $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}$, соответствующий ее минимальному собственному значению $\mu_{\min}(\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}) = 0$. Лемма 5 доказана.

2. Свойства одноранговой оптимальной коррекции матрицы коэффициентов несовместной линейной модели вида (2), (3)

Рассмотрим задачу квадратичного программирования вида

$$\varphi(\mathbf{y}) = \mathbf{y}^{\mathbf{T}}\mathbf{D}\mathbf{y} \to \min, \tag{59}$$

где $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{y} \geqslant \mathbf{0}$, $\mathbf{y}^{\mathbf{T}}\mathbf{y} = 1$ и $\mathbf{D} \in \mathbb{R}^{n\times n}$ — симметричная матрица.

Приводимая ниже теорема из [11], сформулированная в иной форме, определяет необходимые условия, которым должны отвечать решения задачи (59). Эти условия вместе с теоремой 3 позволяют исследовать достижимость нижней грани нормы матрицы коррекции несовместной системы (2), (3) и построение матрицы $\mathbf{H}^* \in \operatorname*{Arginf}\limits_{\mathscr{X}_+(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ и вектора $\mathbf{x}^* = \mathscr{X}_+(\mathbf{A}, \mathbf{b}, \mathbf{H}^*)$.

Теорема 7. *Для того чтобы единичный неотрицательный вектор* $\mathbf{y}$ *был решением задачи* (59), *необходимо выполнение условия*

$$\mathbf{D}\mathbf{y} - \lambda\mathbf{y} \geqslant \mathbf{0}, \tag{60}$$

где λ *— собственное значение матрицы* $\overline{\mathbf{D}}$, $\mathbf{y}$ *— соответствующий ему собственный вектор* $\mathbf{D}$, *а матрица* $\overline{\mathbf{D}}$ *либо совпадает с матрицей* $\mathbf{D}$, *либо получается из* $\mathbf{D}$ *обнулением некоторых столбцов и строк с одинаковыми номерами.*

Теорема 7 следует из условий Куна — Таккера задачи (59). Заметим, что в [11] также отмечается, что для проверки условия (59) необязательно проводить обнуление элементов матрицы $\mathbf{D}$. Вместо этого

можно рассматривать подматрицы $\widetilde{\mathbf{D}}$ матрицы $\mathbf{D}$, получаемые вычеркиванием столбцов и строк, и определять их собственные значения $\tilde{\lambda}$ и соответствующие собственные векторы $\tilde{\mathbf{y}}$. При этом $\tilde{\lambda} = \lambda$, а вектор $\mathbf{y}$ может быть получен дополнением компонент вектора $\tilde{\mathbf{y}}$ соответствующими нулевыми компонентами.

Положив $\mathbf{D} = \mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}$ для задачи (10) и $\mathbf{D} = \mathbf{A}^{\mathbf{T}}\mathbf{A}$ для задачи (11) и использовав теоремы 7 и 3, можно осуществить анализ задачи построения оптимальной матрицы коррекции для несовместной системы (2), (3).

Очевидно, что указанный способ сводится к комбинаторной процедуре поиска собственных значений и собственных векторов матриц $\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}$ и $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ и всех их квадратных подматриц порядка $n-1$, $n-2, \dots, 1$, получаемых вычеркиванием (или обнулением) строк и столбцов с одинаковыми номерами. Для решения задач (10)–(11) приходится решать $2^{n+1} - 2$ вспомогательных задач на собственные значения.

Поскольку практическая ценность данного способа для задач высокой размерности проблематична, дополнительная информация, характеризующая решения задач (10) и (11), представляет интерес. Так, теорема, полученная в [7], позволяет оценить верхнюю границу величины $\inf\limits_{\mathscr{X}_{+}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|^2$.

Теорема 8. *Для задачи коррекции матрицы коэффициентов несовместной линейной системы вида* (1), (3) *справедлива оценка*

$$\inf_{\mathscr{X}_{+}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|^2 \leqslant \frac{\lambda_{\min} + \lambda_{\max}}{2}, \tag{61}$$

где $\lambda_{\min}$ *и* $\lambda_{\max}$ — *минимальное и максимальное собственные значения матрицы* $\mathbf{A}^{\mathbf{T}}\mathbf{A}$.

Приведенные ниже теоремы, являющиеся *следствиями* теоремы 3, при решении задач (10) и (11) позволяют учитывать условия, связанные со знаком выражения $\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{x}$ a posteriori, и при определенных условиях делать выводы о достижимости (или недостижимости) нижней грани нормы корректирующей матрицы по результатам решения только *одной* задачи (10) или (11), что вдвое снижает трудоемкость решения основной задачи.

Теорема 9. *Если для несовместной системы вида* (2), (3) *существует*

$$\mathbf{y} \in \operatorname*{Argmin}_{\mathbf{x}:\mathbf{x}\geqslant 0,\|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} \text{ такое, что } \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y} > 0, \tag{62}$$

то $\inf\limits_{\mathscr{X}_{+}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ *достигается.*

Доказательство. Пусть

$$\widehat{\mathbf{x}} \in \operatorname*{Argmin}_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1, \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{x}\leqslant 0} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{A}\mathbf{x}\}, \quad \gamma = \widehat{\mathbf{x}}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{A}\widehat{\mathbf{x}}.$$

Очевидно, что из (11) $d_2 = \gamma$. Кроме того, в силу (18) имеем

$$\widehat{\mathbf{x}}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}\widehat{\mathbf{x}} = d_2 - (\widehat{\mathbf{x}}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{b})^2 \leqslant d_2.$$

В то же время согласно предположению (62) и формуле (10)

$$d_1 = \mathbf{y}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{y} \leqslant \widehat{\mathbf{x}}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}\widehat{\mathbf{x}}.$$

Следовательно, $d_1 \leqslant d_2$ и $\inf\limits_{\mathscr{X}_+(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ достигается в силу теоремы 3. Теорема 9 доказана.

Пример 3. Рассмотрим несовместную систему вида (2), (3) со следующими значениями $\mathbf{A}$ и $\mathbf{b}$:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 2 \\ 4 \\ 8 \end{bmatrix}.$$

Данная система несовместна, поскольку $\operatorname{rank} \mathbf{A} = 2$ и $\operatorname{rank} [\mathbf{A}\ \mathbf{b}] = 3$.

Пользуясь теоремой 7, можно показать, что

$$\min_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} = \frac{5}{3}.$$

При этом $\operatorname*{Argmin}\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I} - \mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} = \{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3\}$, где

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{y}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Нетрудно убедиться в том, что $\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}_1 = -14$, $\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}_2 = \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}_3 = 14$. Таким образом, $\inf\limits_{\mathscr{X}_+(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ достигается в силу теоремы 9.

В то же время, используя теорему 7, можно показать, что

$$\min_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{A}\mathbf{x}\} = 2.$$

При этом $\operatorname*{Argmin}\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{A}\mathbf{x}\} = \{\mathbf{x}_1, \mathbf{x}_2\}$, где

$$\mathbf{x}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix},$$

$$\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{x}_1 = \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{x}_2 = 0.$$

Таким образом, $\inf\limits_{\mathscr{X}_+(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ достигается в силу теоремы 3.

Теорема 10. *Если для несовместной системы вида* (2), (3) *существует*

$$\mathbf{z} \in \operatorname*{Argmin}_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x}\} \text{ \textit{такое, что} } \mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{z} > 0,$$

то $\inf\limits_{\mathscr{X}_+(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ *достигается.*

Как и в предыдущем случае, утверждение теоремы следует из теоремы 3 и условия $d_1 \leqslant \mathbf{z}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{z} \leqslant \mathbf{z}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{z} \leqslant d_2$, справедливого в силу соотношения (18).

Пример 4. Рассмотрим несовместную систему вида (2), (3) со следующими значениями $\mathbf{A}$ и $\mathbf{b}$:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ -8 \\ 12 \\ -1 \end{bmatrix}.$$

Эта система несовместна, поскольку все элементы матрицы $\mathbf{A}$ неотрицательны, а среди элементов вектора $\mathbf{b}$ есть отрицательные.

Используя теорему 7, можно показать, что $\min\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x}\} = 1$. При этом $\operatorname*{Argmin}\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{x}\} = \{\mathbf{x}_1, \mathbf{x}_2\}$, где

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

$$\mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{x}_1 = 4, \quad \mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{x}_2 = -1.$$

Таким образом, $\inf\limits_{\mathscr{X}_+(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ достигается в силу теоремы 10.

В то же время, используя теорему 7, можно показать, что

$$\min_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} = \frac{209}{225} < 1.$$

При этом

$$\operatorname*{argmin}_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} = \mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{y} = 4$$

и, таким образом, $\inf\limits_{\mathscr{X}_+(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ достигается в силу теоремы 3.

Теорема 11. *Если* $\forall \mathbf{y} \in \operatorname*{Argmin}\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} \Rightarrow \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y} \leqslant 0$, *причем существует* $\mathbf{y} \in \operatorname*{Argmin}\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\}$ *такое, что* $\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y} = 0$, *то* $\inf\limits_{\mathscr{X}_{+}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ *не достигается.*

Доказательство. Пусть $\mathbf{y} \in \operatorname*{Argmin}\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} \mid \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y} = 0$, $\tau = \mathbf{y}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{y}$. В силу (10), (18) и условия теоремы имеем $d_1 > \tau$. В то же время $\tau = \min\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1, \mathbf{b}\mathbf{A}\mathbf{x}=0} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{A}\mathbf{x}\}$, причем из (11) $\min\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1, \mathbf{b}\mathbf{A}\mathbf{x}=0} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\mathbf{A}\mathbf{x}\} \geqslant \mathbf{d}_2$, поскольку допустимая область задачи (11) шире своей подобласти, задаваемой ограничениями $\mathbf{x} \geqslant \mathbf{0}$, $\|\mathbf{x}\| = 1$, $\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{x} = 0$.

С учетом отмеченного выше имеем $d_1 > \tau \geqslant d_2$. Отсюда и из теоремы 3 следует утверждение теоремы 11.

Пример 5. Рассмотрим несовместную систему вида (2), (3) со следующими значениями $\mathbf{A}$ и $\mathbf{b}$:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ -2 & 2 & -1 \\ -1 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}.$$

Так как матрица $\mathbf{A}$ является матрицей полного столбцевого ранга, $\operatorname{rank} \mathbf{A} = \operatorname{rank} [\mathbf{A}\ \mathbf{b}] = 3$ и число уравнений больше числа неизвестных, то система вида (1), (3) с параметрами $\mathbf{A}$, $\mathbf{b}$ совместна и имеет единственное решение

$$\mathbf{z} = \begin{bmatrix} 0 \\ 1/3 \\ -1/3 \end{bmatrix},$$

которое не является неотрицательным.

Используя теорему 7, можно показать, что

$$\min_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} = 2.$$

При этом $\operatorname*{Argmin}\limits_{\mathbf{x}:\mathbf{x}\geqslant 0, \|\mathbf{x}\|=1} \{\mathbf{x}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} = \{\mathbf{y}_1, \mathbf{y}_2\}$, где

$$\mathbf{y}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix},$$

$$\mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}_1 = 0, \quad \mathbf{b}^{\mathbf{T}}\mathbf{A}\mathbf{y}_2 = -3\sqrt{2}.$$

Отсюда и из теоремы 1 следует, что $\inf\limits_{\mathscr{X}_{+}(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing} \|\mathbf{H}\|$ не достигается.

В то же время, используя теорему 7, можно показать, что

$$\min_{\mathbf{x}:\mathbf{x}\geqslant 0,\|\mathbf{x}\|=1}\{\mathbf{x}^{\mathrm T}\mathbf{A}^{\mathrm T}\mathbf{A}\mathbf{x}\} = \min_{\mathbf{x}:\mathbf{x}\geqslant 0,\|\mathbf{x}\|=1}\{\mathbf{x}^{\mathrm T}\mathbf{A}^{\mathrm T}(\mathbf{I}-\mathbf{b}\mathbf{b}^{+})\mathbf{A}\mathbf{x}\} = 2.$$

При этом $\operatorname*{Argmin}\limits_{\mathbf{x}:\mathbf{x}\geqslant 0,\|\mathbf{x}\|=1}\{\mathbf{x}^{\mathrm T}\mathbf{A}^{\mathrm T}\mathbf{A}\mathbf{x}\} = \{\mathbf{x}\}$, где

$$\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix},$$

$$\mathbf{b}^{\mathrm T}\mathbf{A}\mathbf{x} = 0.$$

Поэтому $\inf\limits_{\mathscr{X}_+(\mathbf{A},\mathbf{b},\mathbf{H})\neq\varnothing}\|\mathbf{H}\|$ достигается в силу теоремы 3.

В завершение приведем результат, который является очевидным следствием теоремы 5, обобщающим ее на случай коррекции матрицы коэффициентов несовместной линейной модели вида (2), (3).

Теорема 12. *Пусть матрица* $\mathbf{A}$ *коэффициентов несовместной системы вида* (2), (3) *скорректирована с помощью одноранговoго преобразования вида* (20), *в котором вектор* $\mathbf{x}$ *удовлетворяет условию* $\mathbf{x}\geqslant\mathbf{0}$. *Тогда множество допустимых решений скорректированной системы имеет вид* $\mathbf{y}=\mathbf{x}+\Delta\mathbf{x}$ *и состоит более чем из одного элемента тогда и только тогда, когда существует хотя бы один ненулевой вектор* $\Delta\mathbf{x}$ *такой, что*

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{x}^{\mathrm T} \end{bmatrix}\cdot\Delta\mathbf{x} = \mathbf{0},$$

$$\Delta\mathbf{x} = \left(\Delta x_i \,\middle|\, \begin{array}{ll} \Delta x_i \geqslant 0, & \textit{если } x_i = 0, \\ \Delta x_i \geqslant -x_i & \textit{в противном случае} \end{array}\right).$$

Замечание. Несовместная система (2), (3), матрица коэффициентов которой скорректирована с помощью преобразования (20), может иметь более одного решения только тогда, когда $\operatorname{rank}\mathbf{A} < n$.

Пример 6. Пусть

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix},\ \mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix},\ \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix},\ \mathbf{z} = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$$

Так как $\operatorname{rank}\mathbf{A} = 1$ и $\operatorname{rank}[\mathbf{A}\ \mathbf{b}] = 2$, то система (2), (3) несовместна. Несложно также убедиться, что $\mathbf{x}\perp\mathbf{y}\perp\mathbf{z}$, $\mathbf{A}\cdot[\mathbf{x}\ \mathbf{y}\ \mathbf{z}] = \mathbf{0}$.

Выберем $\mathbf{x}$ в качестве вектора, по которому в соответствии с формулой (20) строится матрица коррекции:

$$\mathbf{H} = (\mathbf{b}-\mathbf{A}\mathbf{x})\mathbf{x}^{+} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\widetilde{\mathbf{A}} = \mathbf{A} + \mathbf{H} = \begin{bmatrix} 3/2 & -1/2 & 1 & -1 \\ -2 & 0 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}.$$

Тогда согласно теореме 12 множество

$$\mathrm{T} = \{\mathbf{t} \mid \widetilde{\mathbf{A}}\mathbf{t} = \mathbf{b}, \mathbf{t} \geqslant \mathbf{0}\}$$

неотрицательных решений системы $\widetilde{\mathbf{A}}\mathbf{x} = \mathbf{b}$ состоит более чем из одного элемента. Более того, для рассматриваемого примера несложно показать, что любое решение скорректированной системы представимо в виде

$$\mathbf{t} = \mathbf{x} + \alpha \cdot \mathbf{y} + \beta \cdot \mathbf{z},$$

где

$$\alpha \in [0, +\infty), \quad |\beta| \leqslant \min\{1, \alpha\}.$$

Автор выражает признательность организаторам конференции «Дискретный анализ и исследование операций» (DAOR′2000) за возможность представления предварительных результатов данной работы и участникам секции математического программирования этой конференции за высказанные замечания и пожелания.

ЛИТЕРАТУРА

1. **Алберт А.** Регрессия, псевдоинверсия и рекуррентное оценивание. М.: Наука, 1977.
2. **Ватолин А. А.** Аппроксимация несобственных задач линейного программирования по критерию евклидовой нормы // Журн. вычисл. математики и мат. физики. 1984. Т. 24, № 12. С. 1907–1908.
3. **Воеводин В. В., Кузнецов Ю. А.** Матрицы и вычисления. М.: Наука, 1984.
4. **Гантмахер Ф. Р.** Теория матриц. М.: Наука, 1988.
5. **Голуб Дж., Ван Лоун Ч.** Матричные вычисления. М.: Мир, 1999.
6. **Горелик В. А., Ерохин В. И.** О верхней оценке минимума квадратичной формы на пересечении единичной сферы с областью неотрицательных координат // Моделирование, оптимизация и декомпозиция сложных динамических процессов. М.: ВЦ РАН, 2001. С. 30–50.
7. **Горелик В. А., Кондратьева В. А.** Параметрическое программирование и несобственные задачи линейной оптимизации // Моделирование, оптимизация и декомпозиция сложных динамических процессов. М.: ВЦ РАН, 1999. С. 57–82.
8. **Еремин И. И., Мазуров В. Д., Астафьев Н. Н.** Несобственные задачи линейного и выпуклого программирования. М.: Наука, 1983.

9. **Ерохин В. И.** Обобщение тождества Шермана — Моррисона на случай одноранговой модификации псевдообратной матрицы полного ранга // Журн. вычисл. математики и мат. физики. 1999. Т. 39, № 8. С. 1280–1282.
10. **Икрамов Х. Д.** Задачник по линейной алгебре. М.: Наука, 1975.
11. **Кондратьева В. А.** Несобственные задачи линейной оптимизации и параметрическое программирование: Дис. ... канд. физ.-мат. наук. М., 2000.
12. **Лоусон Ч., Хенсон Р.** Численное решение задач метода наименьших квадратов. М.: Наука, 1986.
13. **Хорн Р., Джонсон Ч.** Матричный анализ. М.: Мир, 1989.
14. **Vatolin A. A.** Parametric approximation of inconsistent systems of linear equations and inequalities // Parametric optimization and ill-posed problems in mathematical optimization. Seminar Berichte, 81. Berlin: Humboldt Univ., 1986. P. 145–154.

Адрес автора:

Борисоглебский
государственный
педагогический институт,
ул. Народная, 43,
397160 Борисоглебск,
Воронежская обл., Россия.
E-mail: bgpi@mail.ru,
bgpi@vmail.ru

Статья поступила
24 января 2002 г.