

ОПТИМАЛЬНОЕ
УПРАВЛЕНИЕ

УДК 519.658

ПРЯМО-ДВОЙСТВЕННЫЙ МЕТОД НЬЮТОНА С НАИСКОРЕЙШИМ
СПУСКОМ ДЛЯ ЛИНЕЙНОЙ ЗАДАЧИ ПОЛУОПРЕДЕЛЕННОГО
ПРОГРАММИРОВАНИЯ. ИТЕРАЦИОННЫЙ ПРОЦЕСС¹⁾

© 2022 г. **В. Г. Жадан**^{1,2}

¹ 117333 Москва, ул. Вавилова, 40, ФИЦ ИУ РАН, Россия

² 141700 Долгопрудный, М. о., Институтский пер., 9, МФТИ, Россия

e-mail: zhadan@ccas.ru

Поступила в редакцию 02.04.2021 г.
Переработанный вариант 02.04.2021 г.
Принята к публикации 16.12.2021 г.

Рассматривается прямо-двойственный метод Ньютона для решения линейной задачи полуопределенного программирования. При этом как прямые переменные, так и слабые двойственные переменные могут иметь неполный ранг и принадлежать границам допустимых множеств. Приводятся формулы для определения направлений перемещения, а также исследуются свойства этих направлений. Описывается способ выбора шагов перемещения, приводящий к уменьшению ранга симметризованного произведения прямой и слабой двойственной переменных. Библ. 16.

Ключевые слова: линейная задача полуопределенного программирования, прямо-двойственный метод Ньютона, итерационный процесс, наискорейший спуск.

DOI: 10.31857/S0044466922040135

ВВЕДЕНИЕ

Настоящая статья является продолжением работы [1], в которой рассматривался прямо-двойственный метод Ньютона для решения линейной задачи полуопределенного программирования и были выведены уравнения для определения направлений перемещений. Основная особенность при получении данных уравнений заключалась в том, что точки, в которых они выводились, могли принадлежать границам допустимых множеств. В таких точках ньютоновская система линейных уравнений для определения направлений перемещений принимает специальный вид. Было показано, что в пространстве симметричных матриц можно выделить четыре подпространства, которым должны принадлежать данные перемещения.

Цель настоящей работы заключается в получении точных аналитических формул для ньютоновских направлений как в пространстве прямых переменных, так и в пространстве слабых двойственных переменных, а также в построении итерационного процесса. Исследуются свойства ньютоновских направлений, рассматривается вопрос о выборе шага перемещения. Показывается, что в ходе итерационного процесса ранг симметризованного произведения прямой переменной и слабой двойственной переменной может только уменьшаться. Данный ранг характеризует выполнение условия комплементарности, которое входит в условия оптимальности для линейных задач полуопределенного программирования.

Приведем некоторые обозначения, введенные ранее в [1] и используемые в данной работе. Ниже \mathbb{S}^n – пространство симметричных матриц порядка n , и \mathbb{S}_+^n – конус положительно-полуопределенных матриц из \mathbb{S}^n . Конус \mathbb{S}_+^n самосопряженный и задает в \mathbb{S}^n частичный порядок \succeq , используя который будем писать $X \succeq 0$, если $X \in \mathbb{S}_+^n$. Размерность пространства \mathbb{S}^n равняется “треугольному” числу $n^\Delta = \frac{1}{2}n(n+1)$.

¹⁾ Работа выполнена при финансовой поддержке РФФ (проект 21-71-30005).

Скалярное (внутреннее) произведение между матрицами $X \in \mathbb{S}^n$ и $Y \in \mathbb{S}^n$ определяется по Фробениусу: $X \bullet Y = \text{tr}(X^T Y)$. Обычное евклидово скалярное произведение в \mathbb{R}_n обозначается с помощью угловых скобок. Символ J^n обозначает множество индексов от 1 до n . Через $X \circ Y$ обозначается симметризованное произведение между матрицами из \mathbb{S}^n , определяемое как $X \circ Y = \frac{1}{2}(XY + YX)$.

Линейная задача полуопределенного программирования ставится стандартным образом

$$\min C \bullet X, \quad A_i \bullet X = b^i, \quad i \in J^m, \quad X \succeq 0. \quad (1)$$

Матрицы C , X и A_i , $i \in J^m$, принадлежат пространству \mathbb{S}^n . Двойственной к (1) является задача

$$\max \langle b, u \rangle, \quad Y = C - \sum_{i=1}^m u^i A_i, \quad Y \succeq 0, \quad (2)$$

в которой $b = (b^1; \dots; b^m) \in \mathbb{R}^m$. Предполагается, что решения задач (1) и (2) существуют и что матрицы A_i , $i \in J^m$, линейно независимы. Предполагается также, что $C \neq 0_{nn}$ и $b = [b^1; \dots; b^m] \neq 0_n$. Здесь и ниже 0_k есть k -мерный нулевой вектор, 0_{kl} – нулевая матрица размером $k \times l$. Точка с запятой указывает на то, что компоненты вектора располагаются в виде столбца, т.е. одна компонента под другой.

Пусть \mathcal{F}_p и \mathcal{F}_d – допустимые множества в задачах (1) и (2) соответственно. Если $X \in \mathcal{F}_p$ и $[u, Y] \in \mathcal{F}_d$ – решения (1) и (2), то они удовлетворяют системе равенств

$$X \circ Y = 0, \quad A_i \bullet X = b^i, \quad i \in J^m, \quad Y = C - \sum_{i=1}^m u^i A_i. \quad (3)$$

Последнее равенство в (3) можно представить в другом виде, аналогичном равенствам $A_i \bullet X = b^i$. Пусть $l = n^\Delta - m$. Обозначим через K_1, \dots, K_l , матрицы из \mathbb{S}^n , задающий базис в ортогональном дополнении пространства, порождаемого матрицами A_i , $i \in J^m$. Тогда, умножая равенство $Y = C - \sum_{i=1}^m u^i A_i$ скалярно на K_j , $j \in J^l$, получаем $K_j \bullet Y = c^j$, где $c^j = K_j \bullet C$. Условия оптимальности (3) заменяются на следующие:

$$X \circ Y = 0, \quad A_i \bullet X = b^i, \quad i \in J^m, \quad K_j \bullet Y = c^j, \quad j \in J^l. \quad (4)$$

Предлагаемый в [1] метод является вариантом метода Ньютона для решения системы (4).

1. НЬУТОНОВСКАЯ СИСТЕМА В ВЕКТОРНОМ ВИДЕ

Пусть имеется допустимая пара $[X, Y] \in \mathbb{S}^n \times \mathbb{S}^n$. Это означает, что $X \succeq 0$, $Y \succeq 0$ и выполнены последние два равенства из (4). Считаем также, что пара $[X, Y]$ является согласованной, т.е. $X \circ Y \succeq 0$. Если для $[X, Y]$ первое равенство в (4) не выполняется, другими словами, пара не является *комплементарной*, то она неоптимальная, т.е. X и Y вместе с соответствующей двойственной переменной u не являются решениями задач (1) и (2). В этом случае желательно перейти в новую допустимую согласованную пару $[\bar{X}, \bar{Y}]$, используя для этой цели направления перемещения ΔX и ΔY . В качестве ΔX и ΔY могут быть взяты ньютоновские направления, получающиеся из применения метода Ньютона для решения системы (4). В [1] построена специальная система линейных уравнений для нахождения таких ΔX и ΔY . Дадим ее краткое описание.

Для симметричной матрицы $X \circ Y$ имеет место разложение: $X \circ Y = UD(\lambda)U^T$, где U – ортогональная матрица порядка n , $D(\lambda)$ – диагональная матрица с вектором собственных значений $X \circ Y$ на диагонали. Пусть n_p – ранг матрицы $X \circ Y$ и пусть $\lambda = [\lambda_p; \lambda_0]$, где λ_p – ненулевые собственные значения, λ_0 – нулевые собственные значения. Тогда $U = [U_p, U_0]$, где подматрица U_p состоит из n_p столбцов U .

Введем в рассмотрение еще три матрицы U_B, U_N и U_Z , зависящие от образов матриц $X \in \mathbb{S}_+^n$ и $Y \in \mathbb{S}_+^n$, а также от их нуль-пространств. Матрица U_B задает ортонормированный базис в пересечении образа матрицы X с нуль-пространством матрицы Y . Соответственно, матрица U_N задает ортонормированный базис в пересечении образа матрицы Y с нуль-пространством матрицы X . Если пересечение нуль-пространств матриц X и Y имеет положительную размерность, то такие пары $[X, Y]$ называются *особыми*, в противном случае они называются *неособыми*. В неособых парах матрица U_Z отсутствует. Неособая регулярная пара $[X, Y] \in \mathbb{S}_+^n \times \mathbb{S}_+^n$ – это такая пара, когда $U_O = [U_B, U_N]$. В особой регулярной паре $U_O = [U_B, U_N, U_Z]$.

Предположим сначала для простоты, что неоптимальная пара $[X, Y] \in \mathbb{S}_+^n \times \mathbb{S}_+^n$ регулярная и неособая. В этом случае $\lambda_p > 0_{n_p}$. При этих предположениях, как показано в [1], матрицы $X_U = U^T X U$ и $Y_U = U^T Y U$ имеют блочно-диагональный вид

$$X_U = \begin{bmatrix} X_{U_p} & 0 & 0 \\ 0 & X_{U_B} & 0 \\ 0 & 0 & X_{U_N} \end{bmatrix}, \quad Y_U = \begin{bmatrix} Y_{U_p} & 0 & 0 \\ 0 & Y_{U_B} & 0 \\ 0 & 0 & Y_{U_N} \end{bmatrix} \quad (5)$$

с диагональными блоками

$$\begin{aligned} X_{U_p} &= U_p^T X U_p, & X_{U_B} &= U_B^T X U_B, & X_{U_N} &= U_N^T X U_N, \\ Y_{U_p} &= U_p^T Y U_p, & Y_{U_B} &= U_B^T Y U_B, & Y_{U_N} &= U_N^T Y U_N. \end{aligned}$$

Тогда в качестве направлений ΔX и ΔY также целесообразно брать блочно-диагональные матрицы с соответственно блоками $\Delta X_{U_p}, \Delta X_{U_B}, \Delta X_{U_N}$ и $\Delta Y_{U_p}, \Delta Y_{U_B}, \Delta Y_{U_N}$, т.е. полагать

$$\Delta X_U = \begin{bmatrix} \Delta X_{U_p} & 0 & 0 \\ 0 & \Delta X_{U_B} & 0 \\ 0 & 0 & \Delta X_{U_N} \end{bmatrix}, \quad \Delta Y_U = \begin{bmatrix} \Delta Y_{U_p} & 0 & 0 \\ 0 & \Delta Y_{U_B} & 0 \\ 0 & 0 & \Delta Y_{U_N} \end{bmatrix}. \quad (6)$$

Симметричные матрицы X_{U_p}, Y_{U_p} и $\Delta X_{U_p}, \Delta Y_{U_p}$ имеют порядок n_p (ранг матрицы $X \circ Y$). У матриц X_{U_B}, Y_{U_B} и $\Delta X_{U_B}, \Delta Y_{U_B}$ ранг n_B , у матриц X_{U_N}, Y_{U_N} и $\Delta X_{U_N}, \Delta Y_{U_N}$ ранг n_N . Пространство блочно-диагональных симметричных матриц с блоками соответственно порядков n_p, n_B и n_N обозначим через \mathbb{S}_U^n . Его размерность равняется $n_U = n_p^\Delta + n_B^\Delta + n_N^\Delta$.

В случае блочно-диагональных приращений (6) из ньютоновской системы уравнений для равенств (4) следует, что $\Delta X_{U_N} = 0_{n_N}$ и $\Delta Y_{U_B} = 0_{n_B}$, а сама система сводится к следующей:

$$\begin{aligned} Y_{U_p} \circ \Delta X_{U_p} + X_{U_p} \circ \Delta Y_{U_p} &= -D(\lambda_p), \\ A_{i,U_p} \bullet \Delta X_{U_p} + A_{i,U_B} \bullet \Delta X_{U_B} &= 0, \quad i \in J^m, \\ K_{j,U_p} \bullet \Delta Y_{U_p} + K_{j,U_N} \bullet \Delta Y_{U_N} &= 0, \quad j \in J^{l_U}. \end{aligned} \quad (7)$$

Матрицы A_{i,U_p}, A_{i,U_B} и K_{j,U_p}, K_{j,U_N} следующие:

$$A_{i,U_p} = U_p^T A_i U_p, \quad A_{i,U_B} = U_B^T A_i U_B, \quad K_{j,U_p} = U_p^T K_j U_p, \quad K_{j,U_N} = U_N^T K_j U_N,$$

где $i \in J^m, j \in J^{l_U}, l_U = n_U - m$. Блочно-диагональные матрицы с блоками K_{j,U_p}, K_{j,U_B} и $K_{j,U_N}, j \in J^{l_U}$, задают некоторый базис в ортогональном дополнении подпространства из \mathbb{S}_U^n , определяемым блочно-диагональными матрицами с блоками $A_{i,U_p}, A_{i,U_B}, A_{i,U_N}, i \in J^m$.

Чтобы решить систему линейных матричных уравнений (7), представим ее в векторном виде. Для этого введем дополнительные обозначения.

Пусть вес M обозначает *вектор длины n^2* , являющийся прямой суммой столбцов квадратной матрицы M порядка n . В случае симметричной матрицы M будем пользоваться также вектором

svес M длины n^Δ . В данный вектор в отличие от вес M помещаются столбцы не полностью, а только их части, начинающиеся с главного элемента, причем все элементы, не стоящие на диагонали матрицы M , при помещении в свес M умножаются на $\sqrt{2}$.

Для перехода от вес M к свес M и для обратного перехода нам потребуются специальные *элиминационные* и *дублицирующие* матрицы (см. [2], [3]), которые обозначим как $\tilde{\mathcal{L}}_n$ и $\tilde{\mathcal{D}}_n$ соответственно. Обе матрицы $\tilde{\mathcal{L}}_n$ и $\tilde{\mathcal{D}}_n$ являются матрицами полного ранга и имеют соответственно размеры $n^\Delta \times n^2$ и $n^2 \times n^\Delta$. Если M – симметричная матрица порядка n , то

$$\text{svес } M = \tilde{\mathcal{L}}_n \text{ вес } M, \quad \text{вес } M = \tilde{\mathcal{D}}_n \text{ свес } M.$$

Для матриц $\tilde{\mathcal{L}}_n$ и $\tilde{\mathcal{D}}_n$ справедливо равенство $\tilde{\mathcal{L}}_n \tilde{\mathcal{D}}_n = I_{n^\Delta}$. Кроме того, $\tilde{\mathcal{D}}_n^T \tilde{\mathcal{D}}_n = I_{n^\Delta}$. Здесь и ниже I_k – единичная матрица порядка k .

Матрице $M \in \mathbb{S}^n$ можно поставить в соответствие ее кронекеровскую сумму M^\otimes , определяемую как

$$M^\otimes = \frac{1}{2}(M \otimes I_n + I_n \otimes M),$$

где знак \otimes между матрицами обозначает их произведение по Кронекеру. Матрица M^\otimes симметричная и имеет порядок n^2 . Ниже через \tilde{M}^\otimes обозначается квадратная матрица $\tilde{\mathcal{L}}_n M^\otimes \tilde{\mathcal{D}}_n$, имеющая уже порядок n^Δ . Если M – неособая матрица, то матрицы M^\otimes и \tilde{M}^\otimes также неособые. Более того, неособой оказывается и матрица $\tilde{\mathcal{L}}_n(M \otimes M)\tilde{\mathcal{D}}_n$, причем для обратной матрицы справедлива формула (см. [2])

$$[\tilde{\mathcal{L}}_n(M \otimes M)\tilde{\mathcal{D}}_n]^{-1} = \tilde{\mathcal{L}}_n(M^{-1} \otimes M^{-1})\tilde{\mathcal{D}}_n. \tag{8}$$

Для любых матриц M_1, M_2 и M_3 , если определено их произведение $M_1 M_2 M_3$, имеет место равенство

$$\text{вес } M_1 M_2 M_3 = (M_3^T \otimes M_1) \text{ вес } M_2. \tag{9}$$

С помощью формулы (9) получаем, что равенство $X_{U_p} \circ Y_{U_p} = D(\lambda_p)$ может быть записано следующим образом:

$$\text{svес}(X_{U_p} \circ Y_{U_p}) = \tilde{X}_{U_p}^\otimes \text{svес } Y_{U_p} \tilde{Y}_{U_p}^\otimes \text{svес } X_{U_p} = \text{svес } D(\lambda_p). \tag{10}$$

Здесь $\tilde{X}_{U_p}^\otimes = \tilde{\mathcal{L}}_{n_p} X_{U_p}^\otimes \tilde{\mathcal{D}}_{n_p}$, $\tilde{Y}_{U_p}^\otimes = L_{n_p} Y_{U_p}^\otimes \tilde{\mathcal{D}}_{n_p}$.

Если матрицы $\tilde{X}_{U_p}^\otimes$ и $\tilde{Y}_{U_p}^\otimes$ неособые, то неформально через $\tilde{X}_{U_p}^{-\otimes}$ и $\tilde{Y}_{U_p}^{-\otimes}$ будем обозначать их обратные матрицы, т.е. $(\tilde{X}_{U_p}^\otimes)^{-1}$ и $(\tilde{Y}_{U_p}^\otimes)^{-1}$ соответственно. Обозначим также для упрощения записи $\mathcal{G}_p = \tilde{Y}_{U_p}^{-\otimes} \tilde{X}_{U_p}^\otimes$. Тогда на основании (10) получаем

$$\text{svес } X_{U_p} = \mathcal{G}_p \text{ свес } Y_{U_p} = \tilde{Y}_{U_p}^{-\otimes} \text{ свес } D(\lambda_p), \tag{11}$$

$$\text{svес } Y_{U_p} = \mathcal{G}_p^{-1} \text{ свес } X_{U_p} = \tilde{X}_{U_p}^\otimes \text{ свес } D(\lambda_p). \tag{12}$$

Пусть $\mathcal{A}_{\text{svес}}^{U_p}$ и $\mathcal{A}_{\text{svес}}^{U_B}$ – матрицы, строками которых являются соответственно векторы свес A_{i,U_p} и свес A_{i,U_B} , $i \in J^m$. Аналогично, $\mathcal{K}_{\text{svес}}^{U_p}$ и $\mathcal{K}_{\text{svес}}^{U_N}$ – матрицы, строками которых являются соответственно векторы свес K_{j,U_p} и свес K_{j,U_N} , $j \in J^{l_U}$. Тогда, используя равенства

$$\text{svес}(Y_{U_p} \circ \Delta X_{U_p}) = \tilde{Y}_{U_p}^\otimes \text{ свес } \Delta X_{U_p}, \quad \text{svес}(X_{U_p} \circ \Delta Y_{U_p}) = \tilde{X}_{U_p}^\otimes \text{ свес } \Delta Y_{U_p}, \tag{13}$$

получающиеся аналогично (11) и (12), приходим к тому, что ньютоновская система (7) может быть представлена в следующем виде:

$$\tilde{Y}_{U_P}^{\otimes} \text{svec} \Delta X_{U_P} + \tilde{X}_{U_P}^{\otimes} \text{svec} \Delta Y_{U_P} = -\text{svec} D(\lambda_p), \tag{14}$$

$$\mathcal{A}_{\text{svec}}^{U_P} \text{svec} \Delta X_{U_P} + \mathcal{A}_{\text{svec}}^{U_B} \text{svec} \Delta X_{U_B} = 0_m, \tag{15}$$

$$\mathcal{H}_{\text{svec}}^{U_P} \text{svec} \Delta Y_{U_P} + \mathcal{H}_{\text{svec}}^{U_N} \text{svec} \Delta Y_{U_N} = 0_{l_U}. \tag{16}$$

Обозначим через Ψ матрицу системы (14)–(16), через f – ее правую часть, т.е. положим

$$\Psi = \begin{bmatrix} \tilde{Y}_{U_P}^{\otimes} & 0 & \tilde{X}_{U_P}^{\otimes} & 0 \\ \mathcal{A}_{\text{svec}}^{U_P} & \mathcal{A}_{\text{svec}}^{U_B} & 0 & 0 \\ 0 & 0 & \mathcal{H}_{\text{svec}}^{U_P} & \mathcal{H}_{\text{svec}}^{U_N} \end{bmatrix}, \quad f = \begin{bmatrix} -\text{svec} D(\lambda_p) \\ 0_m \\ 0_{l_U} \end{bmatrix}.$$

Тогда система (14)–(16) переписывается в виде $\Psi z = f$, где

$$z = [\text{svec} \Delta X_{U_P}; \text{svec} \Delta X_{U_B}; \text{svec} \Delta Y_{U_P}; \text{svec} \Delta Y_{U_N}].$$

В [1] дано определение невырожденной пары $[X, Y]$, где $X \in \mathbb{S}_U^n, Y \in \mathbb{S}_U^n$. В невырожденной паре $[X, Y]$ строки матриц $[\mathcal{A}_{\text{svec}}^{U_P}, \mathcal{A}_{\text{svec}}^{U_B}]$ и $[\mathcal{H}_{\text{svec}}^{U_P}, \mathcal{H}_{\text{svec}}^{U_N}]$ линейно независимы. Более того, пара $[X, Y]$ является невырожденной в том и только в том случае, когда выполняются неравенства $m \geq n_B^\Delta \geq m - n_P^\Delta$. Если $n_B > 0$, то матрицы $A_{i,U_B}, i \in J^m$, порождают все пространство \mathbb{S}^{n_B} . Аналогично, если $n_N > 0$, то матрицы $K_{j,U_N}, j \in J^{l_U}$, порождают все пространство \mathbb{S}^{n_N} .

Утверждение 1. Пусть допустимая пара $[X, Y] \in \mathbb{S}_U^n \times \mathbb{S}_U^n$ регулярная и согласованная. Пусть, кроме того, она является невырожденной парой. Тогда матрица Ψ неособая.

Доказательство. Так как при сделанных предположениях матрицы X_{U_P} и Y_{U_P} , имеющие порядок n_P , неособые, то матрицы $\tilde{X}_{U_P}^{\otimes}$ и $\tilde{Y}_{U_P}^{\otimes}$ также неособые порядка n_P^Δ . Учитывая это, умножим первую строку матрицы Ψ на матрицу $\mathcal{A}_{\text{svec}}^{U_P} \tilde{Y}_{U_P}^{\otimes}$ и вычтем ее из второй строки. В результате приходим к матрице

$$\Psi_1 = \begin{bmatrix} \tilde{Y}_{U_P}^{\otimes} & 0 & \tilde{X}_{U_P}^{\otimes} & 0 \\ 0 & \mathcal{A}_{\text{svec}}^{U_B} & -\mathcal{A}_{\text{svec}}^{U_P} G_P & 0 \\ 0 & 0 & \mathcal{H}_{\text{svec}}^{U_P} & \mathcal{H}_{\text{svec}}^{U_N} \end{bmatrix}.$$

Определитель получившейся матрицы Ψ_1 совпадает с определителем матрицы Ψ . Отсюда следует, что $\det \Psi = \det \tilde{Y}_{U_P}^{\otimes} \det \Psi_2$, где

$$\Psi_2 = \begin{bmatrix} \mathcal{A}_{\text{svec}}^{U_B} & -\mathcal{A}_{\text{svec}}^{U_P} G_P & 0 \\ 0 & \mathcal{H}_{\text{svec}}^{U_P} & \mathcal{H}_{\text{svec}}^{U_N} \end{bmatrix}.$$

Следовательно, $\det \Psi \neq 0$, если $\det \Psi_2 \neq 0$.

В случае, когда $n_B = 0$ и, следовательно, $n_B^\Delta = 0$, матрица Ψ_2 имеет вид

$$\Psi_3 = \begin{bmatrix} -\mathcal{A}_{\text{svec}}^{U_P} G_P & 0 \\ \mathcal{H}_{\text{svec}}^{U_P} & \mathcal{H}_{\text{svec}}^{U_N} \end{bmatrix}.$$

Более того, поскольку $[X, Y]$ – невырожденная пара и $n_B = 0$, то $n_P^\Delta \geq m$. Поэтому матрица $\mathcal{A}_{\text{svec}}^{U_P}$ имеет полный ранг, равный m . Матрица $\mathcal{A}_{\text{svec}}^{U_P} G_P$ также имеет полный ранг, равный m .

Обратимся к линейной однородной системе

$$\Psi_3 \mu = 0_{n_U}. \tag{17}$$

Пусть $\mu = [\mu_P; \mu_N]$. Тогда, как следует из уравнения $\mathcal{H}_{\text{svec}}^{U_P} \mu_P + \mathcal{H}_{\text{svec}}^{U_N} \mu_N = 0$, вектор μ должен принадлежать ортогональному дополнению к подпространству, порожденному строками матрицы $[K_{\text{svec}}^{U_P}, K_{\text{svec}}^{U_N}]$. Но данное ортогональное дополнение согласно определению матриц K_{j,U_P} и K_{j,U_N} , $j \in J^{l_U}$, порождается строками матрицы $[A_{\text{svec}}^{U_P}, A_{\text{svec}}^{U_N}]$. Поэтому решение μ должно иметь вид

$$\mu = \begin{bmatrix} \mu_P \\ \mu_N \end{bmatrix} \begin{bmatrix} (\mathcal{A}_{\text{svec}}^{U_P})^T \\ (\mathcal{A}_{\text{svec}}^{U_N})^T \end{bmatrix} \nu$$

для некоторого вектора $\nu \in \mathbb{R}_m$. Отсюда, в частности, получаем $\mu_P = (\mathcal{A}_{\text{svec}}^{U_P})^T \nu$. После подстановки данного μ_P в уравнение в (17), приходим к заключению, что должно выполняться равенство $\mathcal{A}_{\text{svec}}^{U_P} \mathcal{G}_P (\mathcal{A}_{\text{svec}}^{U_P})^T \nu = 0_m$. Так как матрица $\mathcal{A}_{\text{svec}}^{U_P} \mathcal{G}_P (\mathcal{A}_{\text{svec}}^{U_P})^T$ неособая, то $\nu = 0_m$. Таким образом, однородная система (17) имеет только нулевое решение. Следовательно, матрица этой системы неособая, ее ранг отличен от нулевого.

Перейдем теперь к рассмотрению общего случая, когда $n_B > 0$. Так как пара $[X, Y]$ невырожденная, то выполняется неравенство $n_B^\Delta \leq m$.

Если $n_B^\Delta = m$, то матрица $\mathcal{A}_{\text{svec}}^{U_B}$ квадратная порядка m . Опять же в силу невырожденности пары $[X, Y]$ матрицы A_{i,U_B} , $i \in J^m$, порождают все пространство \mathbb{S}^{n_B} . Поэтому матрица $\mathcal{A}_{\text{svec}}^{U_B}$, стоящая в левом верхнем углу Ψ_2 , обязательно должна быть неособой. Кроме того, матрица $[\mathcal{H}_{\text{svec}}^{U_P}, \mathcal{H}_{\text{svec}}^{U_N}]$, стоящая в нижней строке Ψ_2 , также оказывается квадратной порядка $l_U = n_P^\Delta + n_N^\Delta$, причем в силу линейной независимости матриц $K_j^{U_N}$, $j \in J^{l_U}$, следующей из невырожденности пары $[X, Y]$, она неособая. Поэтому вся матрица Ψ_2 , будучи квадратной и имеющей порядок $n_P^\Delta + n_B^\Delta + n_N^\Delta$, оказывается неособой.

Рассмотрим наконец случай, когда $0 < n_B^\Delta < m$, и покажем, что однородная система уравнений $\Psi_2 \mu = 0_{n_U}$ с квадратной матрицей Ψ_2 порядка n_U также может иметь только тривиальное решение.

Пусть $d_B = m - n_B^\Delta > 0$. Разобьем вектор μ на три подвектора, а именно, $\mu = [\mu_B; \mu_P; \mu_N] \in \mathbb{R}^{n_U}$. Система $\Psi_2 \mu = 0$ распадается на две подсистемы:

$$\mathcal{A}_{\text{svec}}^{U_B} \mu_B - \mathcal{A}_{\text{svec}}^{U_P} \mathcal{G}_P \mu_P = 0_m, \quad (18)$$

$$\mathcal{H}_{\text{svec}}^{U_P} \mu_P + \mathcal{H}_{\text{svec}}^{U_N} \mu_N = 0_{l_U}. \quad (19)$$

Так как пара $[X, Y]$ невырожденная, то матрица $[\mathcal{H}_{\text{svec}}^{U_P}, \mathcal{H}_{\text{svec}}^{U_N}]$ подсистемы (19) размером $l_U \times (l_U + d_B)$ имеет полный ранг, равный l_U . Следовательно, общее решение этой подсистемы есть d_B -мерное линейное подпространство.

Пусть \mathcal{L}_B – подпространство в \mathbb{R}^m , порожденное столбцами матрицы $\mathcal{A}_{\text{svec}}^{U_B}$ и пусть E_B – матрица некоторого базиса в подпространстве \mathcal{L}_B^\perp , являющемся ортогональным дополнением \mathcal{L}_B . Поскольку опять в силу невырожденности пары $[X, Y]$ матрицы A_{i,U_B} , $i \in J^m$, порождают пространство \mathbb{S}^{n_B} , то матрица $\mathcal{A}_{\text{svec}}^{U_B}$ имеет полный ранг, равный n_B^Δ . Поэтому у матрицы E_B размер $m \times d_B$.

Покажем, что общее решение подсистемы (19) имеет вид

$$\begin{bmatrix} \mu_P \\ \mu_N \end{bmatrix} = \begin{bmatrix} (\mathcal{A}_{\text{svec}}^{U_P})^T \\ (\mathcal{A}_{\text{svec}}^{U_N})^T \end{bmatrix} E_B \nu,$$

где $v \in \mathbb{R}^{d_B}$. Действительно, подставляя данное решение в (19), в силу ортогональности матриц \tilde{A}_i^U и \tilde{K}_j^U , $i \in J^m$, $j \in J^{l_U}$, получаем

$$\begin{aligned} & \left[\mathcal{H}_{\text{svec}}^{U_P} \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T + \mathcal{H}_{\text{svec}}^{U_N} \left(\mathcal{A}_{\text{svec}}^{U_N} \right)^T \right] E_B v = \\ & = \left[\mathcal{H}_{\text{svec}}^{U_B} \left(\mathcal{A}_{\text{svec}}^{U_B} \right)^T + \mathcal{H}_{\text{svec}}^{U_P} \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T + \mathcal{H}_{\text{svec}}^{U_N} \left(\mathcal{A}_{\text{svec}}^{U_N} \right)^T \right] E_B v = 0. \end{aligned}$$

При этом матрица $\left[\mathcal{A}_{\text{svec}}^{U_P}, \mathcal{A}_{\text{svec}}^{U_N} \right]^T E_B$ имеет полный ранг, равный d_B . Чтобы в этом убедиться, достаточно показать, что полный ранг, равный d_B , имеет подматрица $\left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T E_B$.

Число столбцов матрицы $\left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T E_B$ не больше числа строк. Действительно, иначе выполнялось бы неравенство $d_B = m - n_B^\Delta > n_P^\Delta$ или $m > n_B^\Delta + n_P^\Delta$, которое противоречит неравенству $m \leq n_B^\Delta + n_P^\Delta$, следующему из предположения о невырожденности пары $[X, Y]$. Таким образом, ранг матрицы $\left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T E_B$ не превышает d_B . На самом деле, он равен d_B . Пусть это не так, тогда $\left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T E_B h = 0_{n_P^\Delta}$ для некоторого ненулевого вектора $h \in \mathbb{R}^{d_B}$. Но это означает, что ненулевой вектор $E_B h$, принадлежащий нуль-пространству матрицы $\left(\mathcal{A}_{\text{svec}}^{U_B} \right)^T$, принадлежит одновременно нуль-пространству матрицы $\left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T$. Последнее невозможно, из-за того, что в силу предположения о невырожденности пары $[X, Y]$ строки матрицы $\left[\mathcal{A}_{\text{svec}}^{U_P}, \mathcal{A}_{\text{svec}}^{U_B} \right]$ линейно независимы.

Подставим найденное выражение для μ_P в подсистему (18), в результате получаем следующее уравнение:

$$\mathcal{A}_{\text{svec}}^{U_B} \mu_B - \mathcal{A}_{\text{svec}}^{U_P} \mathcal{G}_P \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T E_B v = 0_m, \tag{20}$$

которое после умножения слева на матрицу E_B^T принимает вид

$$E_B^T \mathcal{A}_{\text{svec}}^{U_P} \mathcal{G}_P \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T E_B v = 0_m.$$

Матрица этой системы, как следует из вышесказанного, неособая, поэтому $v = 0_{d_B}$.

Мы установили, что $\mu_P = 0_{n_P^\Delta}$ и $\mu_N = 0_{n_N^\Delta}$. Кроме того, из (20) в силу линейной независимости столбцов матрицы $\mathcal{A}_{\text{svec}}^{U_B}$ получаем, что $\mu_B = 0_{n_B^\Delta}$. Таким образом, однородная система $\Psi_2 \mu = 0_{n_U}$ имеет только нулевое решение. Следовательно, матрица Ψ_2 неособая. Утверждение доказано.

2. РЕШЕНИЕ НЬЮТОНОВСКОЙ СИСТЕМЫ

Рассмотрим разные случаи решения системы (14)–(16) в зависимости от ранга n_B матрицы U_B . Предполагаем, что текущая допустимая пара $[X, Y] \in \mathbb{S}_U^n \times \mathbb{S}_U^n$ является невырожденной, а также регулярной и согласованной.

Случай 1: $n_B^\Delta = m$. При выполнении этого равенства матрица $\mathcal{A}_{\text{svec}}^{U_B}$ квадратная порядка m , причем из-за предположения о невырожденности пары $[X, Y]$ данная матрица неособая. Имеем

$$\Delta Y_U = - \sum_{i=1}^m \Delta u^i \begin{bmatrix} A_{i,U_P} & 0 & 0 \\ 0 & A_{i,U_B} & 0 \\ 0 & 0 & A_{i,U_N} \end{bmatrix} \tag{21}$$

для некоторого $\Delta u = [\Delta u^1; \dots, \Delta u^m] \in \mathbb{R}^m$. Кроме того, как уже отмечалось, $\Delta Y_{U_B} = 0_{n_B n_B}$. Отсюда, поскольку квадратные симметричные матрицы A_{i,U_B} порядка n_B линейно независимые и их ровно m штук, приходим к выводу, что $\Delta u = 0_m$. В этом случае согласно (21) выполняется следую-

щее: $\Delta Y_U = 0$ и, в частности, $\Delta Y_{U_p} = 0_{n_p n_p}$. Таким образом, первое уравнение (14) сводится к $\tilde{Y}_{U_p}^{\otimes} \text{svec} \Delta X_{U_p} = -\text{svec} D(\lambda_p)$, разрешая которое, получаем $\text{svec} \Delta X_{U_p} = -\tilde{Y}_{U_p}^{-\otimes} \text{svec} D(\lambda_p)$.

Так как $D(\lambda_p) = X_{U_p} \circ Y_{U_p}$, то в силу (10) $\text{svec} D(\lambda_p) = \tilde{Y}_{U_p}^{\otimes} \text{svec} X_{U_p}$. Поэтому выражение для $\text{svec} \Delta X^{U_p}$ упрощается

$$\text{svec} \Delta X_{U_p} = -\text{svec} X_{U_p}. \quad (22)$$

Зная $\text{svec} \Delta X_{U_p}$, из второго уравнения (15) находим

$$\text{svec} \Delta X_{U_B} = \left(\mathcal{A}_{\text{svec}}^{U_B} \right)^{-1} \mathcal{A}_{\text{svec}}^{U_p} \text{svec} X_{U_p} = \left(\mathcal{A}_{\text{svec}}^{U_B} \right)^{-1} b - \text{svec} X_{U_B}. \quad (23)$$

Случай 2: $0 < n_B^{\Delta} < m$. При этом предположении матрица $\mathcal{A}_{\text{svec}}^{U_B}$ становится прямоугольной, причем число столбцов в ней меньше числа строк. Обозначим через \mathcal{M} пространство столбцов матрицы $\mathcal{A}_{\text{svec}}^{U_B}$, через \mathcal{M}^{\perp} – его ортогональное дополнение. Пусть E_B – матрица некоторого базиса в подпространстве \mathcal{M}^{\perp} , она имеет размер $m \times d_B$, где $d_B = m - n_B^{\Delta}$. Так как, по-прежнему, $\Delta Y_{U_B} = 0_{n_B n_B}$, то

$$\text{svec} \Delta Y_{U_B} = -\left(\mathcal{A}_{\text{svec}}^{U_B} \right)^{\top} \Delta u = 0_{n_B^{\Delta}}.$$

Отсюда следует, что $\Delta u = E_B q$ для некоторого $q \in \mathbb{R}^{d_B}$. Поэтому

$$\text{svec} \Delta Y_{U_p} = -\left(\mathcal{A}_{\text{svec}}^{U_p} \right)^{\top} E_B q, \quad \text{svec} \Delta Y_{U_N} = -\left(\mathcal{A}_{\text{svec}}^{U_N} \right)^{\top} E_B q. \quad (24)$$

Умножим обе части уравнения (14) слева на матрицу $\tilde{Y}_{U_p}^{-\otimes}$ и подставим в получившееся уравнение $\text{svec} \Delta Y_{U_p}$ из (24). В результате приходим к следующему выражению для $\text{svec} \Delta X_{U_p}$:

$$\text{svec} \Delta X_{U_p} = \tilde{Y}_{U_p}^{-\otimes} \left[\tilde{X}_{U_p}^{\otimes} \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^{\top} E_B q - \text{svec} D(\lambda_p) \right].$$

Если учесть (11), то

$$\text{svec} \Delta X_{U_p} = \mathcal{G}_p \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^{\top} E_B q - \text{svec} X_{U_p}.$$

Подставляя найденное $\text{svec} \Delta X_{U_p}$ в уравнение (15) и умножая далее (15) слева на матрицу E_B^{\top} , с учетом того, что $E_B^{\top} \mathcal{A}_{\text{svec}}^{U_B} = 0_{d_B}$, получаем

$$\left(E_B^{\top} \mathcal{A}_{\text{svec}}^{U_p} \mathcal{G}_p \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^{\top} E_B \right) q = E_B^{\top} \mathcal{A}_{\text{svec}}^{U_p} \text{svec} X_{U_p}.$$

Отсюда находим $q = \Phi_p^{-1} E_B^{\top} \mathcal{A}_{\text{svec}}^{U_p} \text{svec} X_{U_p}$, где $\Phi_p = E_B^{\top} \mathcal{H}_p E_B$, $\mathcal{H}_p = \mathcal{A}_{\text{svec}}^{U_p} \mathcal{G}_p \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^{\top}$. Таким образом, вводя еще одно обозначение $\tilde{\Phi}_p^{-1} = E_B \Phi_p^{-1} E_B^{\top}$, приходим к следующей формуле для приращения $\text{svec} \Delta X_{U_p}$:

$$\text{svec} \Delta X_{U_p} = \left[\mathcal{G}_p \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^{\top} \tilde{\Phi}_p^{-1} \mathcal{A}_{\text{svec}}^{U_p} - I_{n_p^{\Delta}} \right] \text{svec} X_{U_p}. \quad (25)$$

Но $\mathcal{A}_{\text{svec}}^{U_p} \text{svec} X_{U_p} = b - \mathcal{A}_{\text{svec}}^{U_B} \text{svec} X_{U_B}$, поэтому

$$\text{svec} \Delta X_{U_p} = \mathcal{G}_p \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^{\top} \tilde{\Phi}_p^{-1} b - \text{svec} X_{U_p}. \quad (26)$$

Кроме того, из (15) получаем

$$\text{svec} \Delta X_{U_B} = -\left[\left(\mathcal{A}_{\text{svec}}^{U_B} \right)^{\top} \mathcal{A}_{\text{svec}}^{U_B} \right]^{-1} \left(\mathcal{A}_{\text{svec}}^{U_B} \right)^{\top} \mathcal{A}_{\text{svec}}^{U_p} \text{svec} \Delta X_{U_p}.$$

Следовательно, согласно (26),

$$\mathcal{A}_{\text{svec}}^{U_P} \text{svec } \Delta X_{U_P} = \mathcal{H}_P \tilde{\Phi}_P^{-1} b + \mathcal{A}_{\text{svec}}^{U_B} \text{svec } X_{U_B} - b,$$

что влечет

$$\text{svec } \Delta X_{U_B} = \left[\left(\mathcal{A}_{\text{svec}}^{U_B} \right)^T \mathcal{A}_{\text{svec}}^{U_B} \right]^{-1} \left(\mathcal{A}_{\text{svec}}^{U_B} \right)^T \left(I_m - \mathcal{H}_P \tilde{\Phi}_P^{-1} \right) b - \text{svec } X_{U_B}. \quad (27)$$

Далее, используя соотношение $\text{svec } \Delta Y_U = - \left(\mathcal{A}_{\text{svec}}^U \right)^T \Delta u$, вычисляем соответствующие направления в Y -пространстве:

$$\text{svec } \Delta Y_{U_P} = - \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T \tilde{\Phi}_P^{-1} b, \quad \text{svec } \Delta Y_{U_N} = - \left(\mathcal{A}_{\text{svec}}^{U_N} \right)^T \tilde{\Phi}_P^{-1} b. \quad (28)$$

Как уже отмечалось, $\text{svec } \Delta X_{U_N} = 0_{n_N^\Delta}$ и $\text{svec } \Delta Y_{U_B} = 0_{n_B^\Delta}$.

Случай 3: $n_B = 0$. При этом предположении $\mathcal{A}_{\text{svec}}^{U_P} \text{svec } X_{U_P} = b$, а из уравнения (15) следует, что $\mathcal{A}_{\text{svec}}^{U_P} \text{svec } \Delta X_{U_P} = 0_m$. Имеем $\text{svec } \Delta Y_{U_P} = - \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T \Delta u$ для некоторого m -мерного вектора Δu . Тогда уравнение (14) с учетом (11) сводится к

$$\text{svec } \Delta X_{U_P} - \mathcal{G}_P \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T \Delta u = - \text{svec } X_{U_P}. \quad (29)$$

Умножим уравнение (29) слева на матрицу $\mathcal{A}_{\text{svec}}^{U_P}$, в результате находим $\Delta u = \mathcal{H}_P^{-1} b$. После подстановки найденного Δu в (29) получаем

$$\text{svec } \Delta X_{U_P} = \mathcal{G}_P \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T \mathcal{H}_P^{-1} b - \text{svec } X_{U_P}. \quad (30)$$

Отсюда, в частности, следует, что $\text{svec } \Delta X_{U_P} = 0_{n_P^\Delta}$, когда $n_P^\Delta = m$. Для $\text{svec } \Delta X_{U_N}$ по-прежнему имеем $\text{svec } \Delta X_{U_N} = 0_{n_N^\Delta}$.

С помощью соотношения $\text{svec } \Delta Y_U = - \left(\mathcal{A}_{\text{svec}}^U \right)^T \Delta u$ определяем дополнительно

$$\text{svec } \Delta Y_{U_P} = - \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T \mathcal{H}_P^{-1} b, \quad \text{svec } \Delta Y_{U_N} = - \left(\mathcal{A}_{\text{svec}}^{U_N} \right)^T \mathcal{H}_P^{-1} b.$$

Обратим внимание, что матрица \mathcal{H}_P^{-1} переходит в введенную ранее матрицу $\tilde{\Phi}_P^{-1}$, если в последней взять $E_B = I_m$.

Нами получены формулы для приращений ΔX и ΔY , в которых присутствуют матрицы A_{i,U_P} и A_{i,U_B} , $i \in J^m$. Можно также вывести формулы для приращений ΔX и ΔY , в которых присутствуют только матрицы K_{j,U_P} и K_{j,U_N} , $j \in J^{l_U}$. При нахождении таких приращений вместо (15) используется уравнение (16).

3. СВОЙСТВА НЬЮТОНОВСКИХ НАПРАВЛЕНИЙ

Рассмотрим некоторые простейшие свойства, которыми обладают ньютоновские направления ΔX_U и ΔY_U . Эти свойства являются характерными для метода Ньютона, когда он используется для решения оптимизационных задач.

Утверждение 2. *Направления ΔX_U и ΔY_U ортогональны друг другу, т.е.*

$$\Delta X_U \bullet \Delta Y_U = 0. \quad (31)$$

Доказательство. Равенство (31) в эквивалентной форме записывается в виде

$$\langle \text{svec } \Delta X_U, \text{svec } \Delta Y_U \rangle = 0. \quad (32)$$

Учтем далее, что $\text{svec } \Delta X_{U_N} = 0_{n_N^\Delta}$ и $\text{svec } \Delta Y_{U_B} = 0_{n_B^\Delta}$. Тогда (32) выполняется, если

$$\langle \text{svec } \Delta X_{U_P}, \text{svec } \Delta Y_{U_P} \rangle = 0. \quad (33)$$

Покажем, что равенство (33) действительно имеет место, причем считаем, что направления $\text{svec } \Delta X_{U_p}$ и $\text{svec } \Delta Y_{U_p}$ берутся равными соответственно (26) и (28). Подставляя их, получаем

$$\begin{aligned} & \langle \text{svec } \Delta X_{U_p}, \text{svec } \Delta Y_{U_p} \rangle = \\ & = - \left\langle \left[\mathcal{G}_P \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^T \tilde{\Phi}_P^{-1} \mathcal{A}_{\text{svec}}^{U_p} - I_{n_p^\Delta} \right] \text{svec } X_{U_p}, \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^T \tilde{\Phi}_P^{-1} b \right\rangle = \\ & = \left\langle \left[\tilde{\Phi}_P^{-1} \mathcal{A}_{\text{svec}}^{U_p} \mathcal{G}_P \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^T \tilde{\Phi}_P^{-1} \mathcal{A}_{\text{svec}}^{U_p} - \tilde{\Phi}_P^{-1} \mathcal{A}_{\text{svec}}^{U_p} \right] \text{svec } X_{U_p}, b \right\rangle = 0, \end{aligned}$$

так как $\tilde{\Phi}_P^{-1} \mathcal{A}_{\text{svec}}^{U_p} \mathcal{G}_P \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^T \tilde{\Phi}_P^{-1} = \tilde{\Phi}_P^{-1}$. Мы пришли к (33) и, следовательно, к (31). Утверждение доказано.

Введем в рассмотрение функцию

$$\varphi(\alpha) = (X + \alpha \Delta X) \bullet (Y + \alpha \Delta Y),$$

где $\alpha \geq 0$.

Утверждение 3. Пусть $[X, Y]$ – регулярная допустимая согласованная пара. Тогда

$$\varphi(\alpha) = (1 - \alpha) \varphi(0). \tag{34}$$

Доказательство. Поскольку $X \succeq 0, Y \succeq 0$, то $\varphi(0) = X \bullet Y \geq 0$, причем равенство возможно только в том случае, когда пара $[X, Y]$ оптимальная. Имеем

$$\varphi(\alpha) = X_U \bullet Y_U + \alpha(Y_U \bullet \Delta X_U + X_U \bullet \Delta Y_U) + \alpha^2 \Delta X_U \bullet \Delta Y_U,$$

или, если учесть (31),

$$\varphi(\alpha) = X_U \bullet Y_U + \alpha(Y_U \bullet \Delta X_U + X_U \bullet \Delta Y_U).$$

Примем далее во внимание, что $X_{U_N} = 0, \Delta X_{U_N} = 0$ и $Y_{U_B} = 0, \Delta Y_{U_B} = 0$. Тогда $\varphi(\alpha) = X_{U_p} \bullet Y_{U_p} + \alpha h$, где $h = Y_{U_p} \bullet \Delta X_{U_p} + X_{U_p} \bullet \Delta Y_{U_p}$.

Так как

$$Y_{U_p} \bullet \Delta X_{U_p} = \text{tr } Y_{U_p} \Delta X_{U_p} = \text{tr } \Delta X_{U_p} Y_{U_p}$$

и

$$X_{U_p} \bullet \Delta Y_{U_p} = \text{tr } X_{U_p} \Delta Y_{U_p} = \text{tr } \Delta Y_{U_p} X_{U_p},$$

то $h = \text{tr}(Y_{U_p} \circ \Delta X_{U_p}) + \text{tr}(X_{U_p} \circ \Delta Y_{U_p})$. Поэтому в силу первого уравнения из (7) справедливо равенство $h = -\text{tr } D(\lambda_p) = -\sum_{i=1}^{n_p} \lambda_p^i$. Поскольку $\text{tr } D(\lambda_p) = \text{tr}(X_{U_p} \circ Y_{U_p}) = \varphi(0)$, то выполняется (34). Утверждение доказано.

Приращение $\text{svec } \Delta Y_U$ связано с приращением двойственной переменной Δu формулой: $\text{svec } \Delta Y_U = -\left(A_{\text{svec}}^U\right)^T \Delta u$, причем, как следует из вышесказанного, имеем

$$\Delta u = \begin{cases} 0_m, & n_B^\Delta = m, \\ E_B \Phi_P^{-1} E_B^T b, & 0 < n_B^\Delta < m, \\ \mathcal{H}_P^{-1} b, & n_B = 0. \end{cases} \tag{35}$$

Покажем, что направление Δu приводит к увеличению значения целевой функции в двойственной задаче (2). Для этого сделаем дополнительное предположение относительно задачи (1). Пусть

$$\mathcal{F}_{\mathcal{A}} = \{Z \in \mathbb{S}^n : A_i \bullet Z = b^i, 1 \leq i \leq m\}.$$

Условие A_1 . Не существует матрицы $Z \in \mathcal{F}_{\mathcal{A}}$ ранга r такой, что $r^\Delta < m$.

Утверждение 4. Пусть $[X, Y] \in \mathbb{S}_U^n \times \mathbb{S}_U^n$ – регулярная допустимая согласованная пара. Пусть, кроме того, выполнено условие A_1 . Тогда $\langle b, \Delta u \rangle \geq 0$, причем равенство возможно только при $\Delta u = 0_m$.

Доказательство. Рассмотрим только общий случай формулы (35) для приращения Δu , когда $0 < n_B^\Delta < m$. Как уже отмечалось, третий случай $n_B^\Delta = m$ сводится ко второму, если в качестве матрицы E_B взять единичную матрицу I_m . Имеем

$$\langle b, \Delta u \rangle = \langle E_B^T b, \Phi_P^{-1} E_B^T b \rangle \geq 0, \tag{36}$$

поскольку матрица Φ_P положительно определена. Неравенство (36) переходит в равенство, когда $E_B^T b = 0_{d_B}$. Но выполнение данного равенства означает, что вектор b принадлежит нуль-пространству матрицы E_B^T , совпадающему с пространством столбцов матрицы $\mathcal{A}_{\text{svec}}^{U_B}$. Но тогда для некоторого ненулевого $z \in \mathbb{R}^{n_B^\Delta}$ имеет место равенство

$$\mathcal{A}_{\text{svec}}^{U_B} z = b. \tag{37}$$

Пусть $Z \in \mathbb{S}^n$ – матрица, у которой соответствующая матрица $Z_U = U^T Z U$ принадлежит пространству \mathbb{S}_U^n , является блочно-диагональной, причем с единственным ненулевым диагональным блоком Z_{U_B} . Считаем также, что $\text{svec } Z_{U_B} = z$. Ранг r матрицы Z не больше, чем n_B , и на основании (37) получаем

$$A_i \bullet Z = A_i^U \bullet Z_U = \langle \text{svec } A_i^{U_B}, \text{svec } Z_{U_B} \rangle = b^i, \quad 1 \leq i \leq m,$$

т.е. $Z \in \mathcal{F}_{\mathcal{A}}$. Но $r^\Delta \leq n_B^\Delta < m$, что противоречит условию A_1 . Таким образом, возможен только случай, когда $n_B = m$. Но согласно (35) тогда выполняется равенство $\Delta u = 0_m$. Утверждение доказано.

Утверждение 5. Пусть $[X, Y]$ – регулярная допустимая согласованная пара. Тогда верно

$$C \bullet \Delta X \leq 0. \tag{38}$$

причем равенство возможно только при $\Delta X_U = 0$.

Доказательство. Рассмотрим общий случай, когда направления $\text{svec } \Delta X_{U_p}$ и $\text{svec } \Delta Y_{U_p}$ вычисляются согласно (26) и (28). Неравенство (38) перепишем в виде

$$\langle \text{svec } C_U, \text{svec } \Delta X_U \rangle \leq 0. \tag{39}$$

Так как $\Delta X_{U_N} = 0_{n_N n_N}$, то

$$\langle \text{svec } C_U, \text{svec } \Delta X_U \rangle = \langle \text{svec } C_{U_p}, \text{svec } \Delta X_{U_p} \rangle + \langle \text{svec } C_{U_B}, \text{svec } \Delta X_{U_B} \rangle. \tag{40}$$

Имеем с учетом того, что $Y_{U_B} = 0_{n_B n_B}$,

$$\text{svec } C_{U_p} = \text{svec } Y_{U_p} + \left(\mathcal{A}_{\text{svec}}^{U_p} \right)^T u, \quad \text{svec } C_{U_B} = \left(\mathcal{A}_{\text{svec}}^{U_B} \right)^T u. \tag{41}$$

Следовательно,

$$\langle \text{svec } C_{U_p}, \text{svec } \Delta X_{U_p} \rangle = \langle \text{svec } Y_{U_p}, \text{svec } \Delta X_{U_p} \rangle + \langle u, \mathcal{A}_{\text{svec}}^{U_p} \text{svec } \Delta X_{U_p} \rangle, \tag{42}$$

причем $\text{svec } Y_{U_p}$ удовлетворяют равенству (12).

Для второго слагаемого в (40) выполняется

$$\langle \text{svec } C_{U_B}, \text{svec } \Delta X_{U_B} \rangle = \langle u, \mathcal{A}_{\text{svec}}^{U_B} \text{svec } \Delta X_{U_B} \rangle. \tag{43}$$

Подставляя (42), (43) в (40) и учитывая (15), получаем

$$\begin{aligned} \langle \text{svec } C_U, \text{svec } \Delta X_U \rangle &= \langle \text{svec } Y_{U_p}, \text{svec } \Delta X_{U_p} \rangle + \langle u, \mathcal{A}_{\text{svec}}^{U_p} \text{svec } \Delta X_{U_p} + \mathcal{A}_{\text{svec}}^{U_B} \text{svec } \Delta X_{U_B} \rangle = \\ &= \langle \text{svec } Y_{U_p}, \text{svec } \Delta X_{U_p} \rangle. \end{aligned}$$

На основании (26) имеем

$$\begin{aligned} \langle \text{svect} Y_{U_p}, \text{svect} \Delta X_{U_p} \rangle &= \langle \text{svect} Y_{U_p}, \mathcal{G}_p \left(\mathcal{A}_{\text{svect}}^{U_p} \right)^T \tilde{\Phi}_p^{-1} b \rangle - \langle \text{svect} Y_{U_p}, \text{svect} X_{U_p} \rangle = \\ &= \langle \text{svect} Y_{U_p}, \mathcal{G}_p \left(\mathcal{A}_{\text{svect}}^{U_p} \right)^T \tilde{\Phi}_p^{-1} \mathcal{A}_{\text{svect}}^{U_p} \text{svect} X_{U_p} \rangle - \langle \text{svect} Y_{U_p}, \text{svect} X_{U_p} \rangle. \end{aligned}$$

Здесь мы учли, что $b = \mathcal{A}_{\text{svect}}^{U_p} \text{svect} X_{U_p} + \mathcal{A}_{\text{svect}}^{U_B} \text{svect} X_{U_B}$.

Из (12) и (26) следует, что

$$\begin{aligned} \langle \text{svect} Y_{U_p}, \text{svect} \Delta X_{U_p} \rangle &= \left\langle \mathcal{G}_p^{-1} \text{svect} X_{U_p}, \mathcal{G}_p \left(\mathcal{A}_{\text{svect}}^{U_p} \right)^T \tilde{\Phi}_p^{-1} \mathcal{A}_{\text{svect}}^{U_p} \mathcal{G}_p \mathcal{G}_p^{-1} \text{svect} X_{U_p} \right\rangle - \\ &- \left\langle \mathcal{G}_p^{-1} \text{svect} X_{U_p}, \mathcal{G}_p \mathcal{G}_p^{-1} \text{svect} X_{U_p} \right\rangle, \end{aligned}$$

или, обозначая $z = \mathcal{G}_p^{-1} \text{svect} X_{U_p}$, $Z = Z_1 - \mathcal{G}_p$, $Z_1 = \mathcal{G}_p \left(\mathcal{A}_{\text{svect}}^{U_p} \right)^T \tilde{\Phi}_p^{-1} \mathcal{A}_{\text{svect}}^{U_p} \mathcal{G}_p$,

$$\langle \text{svect} Y_{U_p}, \text{svect} \Delta X_{U_p} \rangle = \langle z, Zz \rangle.$$

Найдем собственные значения ξ матрицы Z . Они удовлетворяют уравнению

$$\left(\mathcal{G}_p \left(\mathcal{A}_{\text{svect}}^{U_p} \right)^T \tilde{\Phi}_p^{-1} \mathcal{A}_{\text{svect}}^{U_p} \mathcal{G}_p - \mathcal{G}_p \right) y = \xi y, \quad (44)$$

где y – ненулевой собственный вектор Z . Умножая равенство (44) слева на матрицу $E_B^T \mathcal{A}_{\text{svect}}^{U_p}$, приходим к равенству $\xi E_B^T \mathcal{A}_{\text{svect}}^{U_p} y = 0$. Следовательно, если $E_B^T \mathcal{A}_{\text{svect}}^{U_p} y \neq 0$, т.е. вектор y не принадлежит нуль-пространству матрицы $E_B^T \mathcal{A}_{\text{svect}}^{U_p}$, то $\xi = 0$.

Предположим далее, что $E_B^T \mathcal{A}_{\text{svect}}^{U_p} y = 0$. В этом случае вектор $\mathcal{A}_{\text{svect}}^{U_p} y$ принадлежит подпространству \mathcal{M} , являющемуся пространством столбцов матрицы $\mathcal{A}_{\text{svect}}^{U_B}$, т.е. $\mathcal{A}_{\text{svect}}^{U_p} y = \mathcal{A}_{\text{svect}}^{U_B} \rho$ для некоторого ρ .

Умножим скалярно равенство (44) на вектор $\mathcal{G}_p^{-T} y$. Тогда

$$\left\langle \mathcal{G}_p^{-T} y, \left(\mathcal{G}_p \left(\mathcal{A}_{\text{svect}}^{U_p} \right)^T \tilde{\Phi}_p^{-1} \mathcal{A}_{\text{svect}}^{U_p} \mathcal{G}_p - \mathcal{G}_p \right) y \right\rangle = \xi \langle \mathcal{G}_p^{-T} y, y \rangle.$$

Так как $\mathcal{A}_{\text{svect}}^{U_p} y = \mathcal{A}_{\text{svect}}^{U_B} \rho$, то данное равенство сводится к $\langle y, y \rangle + \xi \langle \mathcal{G}_p^{-T} y, y \rangle = 0$. Откуда $\xi = -\|y\|^2 / \langle \mathcal{G}_p^{-T} y, y \rangle < 0$.

Таким образом, у матрицы Z все собственные значения действительные и неположительные. Следовательно, неравенство $C \bullet \Delta X \leq 0$ действительно имеет место.

Равенство $C \bullet \Delta X = 0$ выполняется тогда и только тогда, когда $Zz = 0$. В этом случае

$$\mathcal{G}_p \left(\mathcal{A}_{\text{svect}}^{U_p} \right)^T \tilde{\Phi}_p^{-1} \mathcal{A}_{\text{svect}}^{U_p} \text{svect} X_{U_p} = \text{svect} X_{U_p}.$$

Отсюда согласно (25) $\text{svect} \Delta X_{U_p} = 0$. Поэтому согласно (15) имеем

$$\mathcal{A}_{\text{svect}}^{U_B} \text{svect} \Delta X_{U_B} = 0_m.$$

Поскольку в невырожденной паре $[X, Y]$ столбцы матрицы $\mathcal{A}_{\text{svect}}^{U_B}$ линейно независимы, то $\text{svect} \Delta X_{U_B} = 0$. Таким образом, $\Delta X_U = 0$. Утверждение доказано.

4. ИТЕРАЦИОННЫЙ ПРОЦЕСС И ВЫБОР ШАГА

Предположим, что задана начальная допустимая регулярная согласованная пара $[X^0, Y^0]$. Пусть, кроме того, на k -м шаге итерационного процесса найдена пара $[X^k, Y^k]$, также являющаяся допустимой регулярной согласованной парой. Обозначим через U ортогональную матрицу порядка n такую, что $X^k \circ Y^k = UD(\lambda)U^T$. Считаем, для определенности, что $U = [U_p, U_B, U_N]$, где

ранги матриц U_P, U_B и U_N равны соответственно n_P, n_B и n_N . Для вектора собственных значений λ матрицы U выполняется: $\lambda = [\lambda_P; 0]$, $\lambda_P \in \mathbb{R}^{n_P}$. Для X^k и Y^k справедливы представления

$$X_U^k = \begin{bmatrix} X_{U_P}^k & 0 & 0 \\ 0 & X_{U_B}^k & 0 \\ 0 & 0 & X_{U_N}^k \end{bmatrix}, \quad Y_U^k = \begin{bmatrix} Y_{U_P}^k & 0 & 0 \\ 0 & Y_{U_B}^k & 0 \\ 0 & 0 & Y_{U_N}^k \end{bmatrix}, \quad (45)$$

где

$$X_{U_P}^k = U_P^T X^k U_P, \quad X_{U_B}^k = U_B^T X^k U_B, \quad X_{U_N}^k = U_N^T X^k U_N$$

и

$$Y_{U_P}^k = U_P^T Y^k U_P, \quad Y_{U_B}^k = U_B^T Y^k U_B, \quad Y_{U_N}^k = U_N^T Y^k U_N.$$

Возьмем симметричные блочно-диагональные матрицы

$$\Delta X_U^k = \begin{bmatrix} \Delta X_{U_P}^k & 0 & 0 \\ 0 & \Delta X_{U_B}^k & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \Delta Y_U^k = \begin{bmatrix} \Delta Y_{U_P}^k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Delta Y_{U_N}^k \end{bmatrix},$$

которым соответствуют векторы $\text{svec } \Delta X_{U_P}^k$, $\text{svec } \Delta X_{U_B}^k$ и $\text{svec } \Delta Y_{U_P}^k$, $\text{svec } \Delta Y_{U_N}^k$.

Имея формулы для ньютоновских направлений ΔX_U^k и ΔY_U^k , построим итерационный процесс

$$X_U^{k+1} = X_U^k + \alpha_k \Delta X_U^k, \quad Y_U^{k+1} = Y_U^k + \alpha_k \Delta Y_U^k, \quad (46)$$

где шаг $\alpha_k \geq 0$ подбирается таким образом, чтобы он был максимально возможным при выполнении следующих условий: $X_U^{k+1} \succeq 0$, $Y_U^{k+1} \succeq 0$ и $X_U^{k+1} \circ Y_U^{k+1} \succeq 0$. Тогда, если все пары $[X^k, Y^k]$, $k \geq 0$, оказываются допустимыми, регулярными, согласованными и неособыми, то итерационный процесс (46) полностью определен.

Для того, чтобы $X_U^{k+1} \succeq 0$, должны выполняться неравенства

$$X_{U_P}^{k+1} = X_{U_P}^k + \alpha_k \Delta X_{U_P}^k \succeq 0, \quad X_{U_B}^{k+1} = X_{U_B}^k + \alpha_k \Delta X_{U_B}^k \succeq 0. \quad (47)$$

Для матриц $X_{U_P}^k$ и $X_{U_B}^k$ справедливы представления (см. [1]):

$$X_{U_P}^k = F_{X_P} D(\eta_{BN}) F_{X_P}^T, \quad X_{U_B}^k = D(\eta_{BB}), \quad (48)$$

где F_{X_P} — ортогональная матрица порядка n_P , η_{BN} и η_{BB} — векторы собственных значений матрицы X^k со всеми строго положительными компонентами.

Пусть

$$\Delta \hat{X}_{U_P}^k = \Delta X_{U_P}^k + X_{U_P}^k, \quad \Delta \hat{X}_{U_B}^k = \Delta X_{U_B}^k + X_{U_B}^k.$$

Тогда на основании (22), (26) и (30) имеем

$$\text{svec } \Delta \hat{X}_{U_P}^k = \begin{cases} 0_m, & n_B^\Delta = m, \\ \mathcal{G}_P \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T E_B \left(E_B^T \mathcal{H}_P E_B \right)^{-1} E_B^T b, & 0 < n_B^\Delta < m, \\ \mathcal{G}_P \left(\mathcal{A}_{\text{svec}}^{U_P} \right)^T \mathcal{H}_P^{-1} b, & n_B = 0, \end{cases}$$

а на основании (23) и (27), получаем

$$\text{svec} \Delta \hat{X}_{U_B}^k = \begin{cases} (\mathcal{A}_{\text{svec}}^{U_B})^{-1} b, & n_B^\Delta = m, \\ \left[\left(\mathcal{A}_{\text{svec}}^{U_B} \right)^T \mathcal{A}_{\text{svec}}^{U_B} \right]^{-1} \left(\mathcal{A}_{\text{svec}}^{U_B} \right)^T (I_m - \mathcal{K}_p \tilde{\Phi}_p^{-1}) b, & 0 < n_B^\Delta < m. \end{cases}$$

В введенных обозначениях первое рекуррентное соотношение в итерационном процессе (46) расщепляется на следующие два:

$$X_{U_p}^{k+1} = (1 - \alpha_k) X_{U_p}^k + \alpha_k \hat{\Delta} X_{U_p}^k, \quad X_{U_B}^{k+1} = (1 - \alpha_k) X_{U_B}^k + \alpha_k \hat{\Delta} X_{U_B}^k.$$

Тогда, как следует из (48), имеем

$$X_{U_p}^{k+1} = F_{X_p} [(1 - \alpha_k) D(\eta_{BN}) + \alpha_k W_{X_p}] F_{X_p}^T,$$

где через W_{X_p} обозначена матрица $W_{X_p} = F_{X_p}^T \hat{\Delta} X_{U_p}^k F_{X_p}$.

Обе матрицы $D(\eta_{BN})$ и W_{X_p} являются симметричными, причем первая матрица положительно-определенная. Поэтому с помощью конгруэнтного преобразования обе они могут быть приведены к диагональному виду [16]. Пусть \mathcal{T}_{X_p} – неособая матрица, осуществляющая такое преобразование:

$$\mathcal{T}_{X_p} D(\eta_{BN}) \mathcal{T}_{X_p}^{-1} = D(\bar{e}), \quad \mathcal{T}_{X_p} W_{X_p} \mathcal{T}_{X_p}^{-1} = D(\sigma_p),$$

где \bar{e} есть n_p -мерный вектор, состоящий из единиц. Тогда

$$X_{U_p}^{k+1} = F_{X_p} \mathcal{T}_{X_p}^{-1} [(1 - \alpha_k) D(\bar{e}) + \alpha_k D(\sigma_p)] \mathcal{T}_{X_p} F_{X_p}^T. \quad (49)$$

Схожим образом получаем

$$X_{U_B}^{k+1} = \mathcal{T}_{X_B}^{-1} [(1 - \alpha_k) D(\bar{e}) + \alpha_k D(\sigma_B)] \mathcal{T}_{X_B}, \quad (50)$$

где \mathcal{T}_{X_B} – неособая матрица, осуществляющая конгруэнтное преобразование:

$$\mathcal{T}_{X_B} D(\eta_{BB}) \mathcal{T}_{X_B}^{-1} = D(\bar{e}), \quad \mathcal{T}_{X_B} \Delta \hat{X}_B^k \mathcal{T}_{X_B}^{-1} = D(\sigma_B).$$

Теперь \bar{e} есть n_B -мерный вектор, состоящий из единиц. Из (49) и (50) видно, что если минимальная из компонент векторов σ_p и σ_B меньше единицы, то шаг α_k конечный. Более того, если среди компонент векторов σ_p и σ_B имеются отрицательные, то обязательно шаг α_k удовлетворяет неравенству: $0 < \alpha_k < 1$.

Обратимся ко второму рекуррентному соотношению из (46). Наряду с (47) должны выполняться соотношения

$$Y_{U_p}^{k+1} = Y_{U_p}^k + \alpha_k \Delta Y_{U_p}^k \succeq 0, \quad Y_{U_N}^{k+1} = Y_{U_N}^k + \alpha_k \Delta Y_{U_N}^k \succeq 0,$$

причем для матриц $Y_{U_p}^k$ и $Y_{U_N}^k$ справедливы представления, аналогичные (48):

$$Y_{U_p}^k = F_{Y_p} D(\theta_{NB}) F_{Y_p}^T, \quad Y_{U_N}^k = D(\theta_{NN}),$$

где F_{Y_p} – ортогональная матрица порядка n_p , θ_{NB} и θ_{NN} – векторы собственных значений матрицы Y^k со всеми строго положительными компонентами.

Проведем с помощью неособых матриц \mathcal{T}_{Y_p} и \mathcal{T}_{Y_N} конгруэнтные преобразования:

$$\mathcal{T}_{Y_p} D(\theta_{NB}) \mathcal{T}_{Y_p}^{-1} = D(\bar{e}), \quad \mathcal{T}_{Y_p} W_{Y_p} \mathcal{T}_{Y_p}^{-1} = D(\omega_p),$$

$$\mathcal{T}_{Y_N} D(\theta_{NN}) \mathcal{T}_{Y_N}^{-1} = D(\bar{e}), \quad \mathcal{T}_{Y_N} \Delta Y_N^k \mathcal{T}_{Y_N}^{-1} = D(\omega_N),$$

где $W_{Y_p} = F_{Y_p}^T \Delta Y_{U_p}^k F_{Y_p}$. Тогда получаем

$$Y_{U_p}^{k+1} = F_{Y_p} \mathcal{T}_{Y_p}^{-1} [D(\bar{e}) + \alpha_k D(\omega_p)] \mathcal{T}_{Y_p} F_{Y_p}^T, \quad Y_{U_N}^{k+1} = \mathcal{T}_{Y_N}^{-1} [D(\bar{e}) + \alpha_k D(\omega_N)] \mathcal{T}_{Y_N}.$$

Таким образом, шаг α_k должен быть ограничен, если среди компонент векторов ω_p или ω_N имеются отрицательные. Более того, если какая-либо компонента ω_p или ω_N окажется меньше -1 , то $0 < \alpha_k < 1$. Так как $\Delta Y = 0_{nn}$ при $n_B^\Delta = m$, то в данном случае ограничение на шаг α_k становится несущественным.

Пусть $Z^k = X^k \circ Y^k$. Из разложения $Z^k = U_p D(\lambda_p) U_p^T$ и равенства

$$U_p^T Z^k U_p + U_p^T \Delta Z^k U_p = 0$$

следует, что $\Delta Z_{U_p}^k = U_p^T \Delta Z^k U_p = -D(\lambda_p)$. Поэтому $Z^k + \alpha_k \Delta Z^k \succeq 0$ для $0 \leq \alpha_k \leq 1$.

На основании вышесказанного приходим к следующему результату.

Утверждение 6. Пусть $0 \leq \alpha_k \leq 1$. Тогда

$$\text{rank } X_{U_p}^{k+1} \leq \text{rank } X_{U_p}^k, \quad \text{rank } Y_{U_p}^{k+1} \leq \text{rank } Y_{U_p}^k.$$

Более того, $\text{rank } Z^{k+1} \leq \text{rank } Z^k$, т.е. ранг матрицы $Z = X \circ Y$ не увеличивается в ходе итераций.

Определение 1. k -ю итерацию назовем активной, если $\text{rank } Z^{k+1} < \text{rank } Z^k$.

Так как допустимая пара $[X, Y]$ является оптимальной, если $\text{rank } X \circ Y = 0$, то в случае, когда метод находит решение задач (1) и (2), то он находит это решение за не более чем n активных итераций.

Согласно утверждению 7, если пара $[X^k, Y^k]$ была допустимой, регулярной, согласованной и неособой, то следующая пара $[X^{k+1}, Y^{k+1}]$ является допустимой, регулярной и согласованной. Однако она может оказаться особой. В этом случае, следует решить вспомогательную линейную полуопределенную задачу дополненности, о которой пойдет речь в следующем разделе.

В заключение данного раздела рассмотрим вопрос о матрице $\tilde{\Phi}_p^{-1}$. Чтобы ее вычислить, надо знать матрицу $\mathcal{G}_p = \tilde{Y}_{U_p}^{-\otimes} \tilde{X}_{U_p}^{\otimes}$. Найдем выражение для $Y_{U_p}^{-\otimes}$.

Матрица $\tilde{Y}_{U_p}^{\otimes}$ представима в виде

$$\tilde{Y}_{U_p}^{\otimes} = \tilde{\mathcal{L}}_{n_p} (F_{Y_p} D(\theta_{NB}) F_{Y_p}^T)^{\otimes} \tilde{\mathcal{D}}_{n_p},$$

причем, так как F_{Y_p} – ортогональная матрица, то справедлива следующая формула:

$$(F_{Y_p} D(\theta_{NB}) F_{Y_p}^T)^{\otimes} = (F_{Y_p} \otimes F_{Y_p}) D^{\otimes}(\theta_{NB}) (F_{Y_p}^T \otimes F_{Y_p}^T), \tag{51}$$

где матрица $F_{Y_p} \otimes F_{Y_p}$ также ортогональная, $(F_{Y_p} \otimes F_{Y_p})^{-1} = F_{Y_p}^T \otimes F_{Y_p}^T$.

Воспользуемся далее формулами, имеющими место для любой квадратной матрицы M порядка n_p (см. [2]):

$$\tilde{\mathcal{D}}_{n_p} \tilde{\mathcal{L}}_{n_p} (M \otimes M) \tilde{\mathcal{D}}_{n_p} = (M \otimes M) \tilde{\mathcal{D}}_{n_p}, \tag{52}$$

$$\tilde{\mathcal{D}}_{n_p} \tilde{\mathcal{L}}_{n_p} M^{\otimes} \tilde{\mathcal{D}}_{n_p} = M^{\otimes} \tilde{\mathcal{D}}_{n_p}. \tag{53}$$

Тогда согласно (51), (52) и (53) получаем

$$\begin{aligned} \tilde{Y}_{U_p}^{\otimes} &= \tilde{\mathcal{L}}_{n_p} (F_{Y_p} \otimes F_{Y_p}) D^{\otimes}(\theta_{NB}) (F_{Y_p}^T \otimes F_{Y_p}^T) \tilde{\mathcal{D}}_{n_p} = \\ &= \tilde{\mathcal{L}}_{n_p} (F_{Y_p} \otimes F_{Y_p}) D^{\otimes}(\theta_{NB}) \tilde{\mathcal{D}}_{n_p} \tilde{\mathcal{L}}_{n_p} (F_{Y_p}^T \otimes F_{Y_p}^T) \tilde{\mathcal{D}}_{n_p} = \\ &= \tilde{\mathcal{L}}_{n_p} (F_{Y_p} \otimes F_{Y_p}) \tilde{\mathcal{D}}_{n_p} \tilde{\mathcal{L}}_{n_p} D^{\otimes}(\theta_{NB}) \tilde{\mathcal{D}}_{n_p} \tilde{\mathcal{L}}_{n_p} (F_{Y_p}^T \otimes F_{Y_p}^T) \tilde{\mathcal{D}}_{n_p}. \end{aligned}$$

Отсюда, используя формулу (8), приходим к следующему выражению для обратной матрицы:

$$\tilde{Y}_{U_p}^{-\otimes} = \mathcal{W} \left(\tilde{\mathcal{L}}_{n_p} D^{\otimes}(\theta_{NB}) \tilde{\mathcal{D}}_{n_p} \right)^{-1} \mathcal{W}^{-1},$$

где введено обозначение ${}^{\circ}W = \tilde{\mathcal{L}}_{n_p}(F_{Y_p} \otimes F_{Y_p})\tilde{\mathcal{D}}_{n_p}$, причем для обратной матрицы ${}^{\circ}W^{-1}$ выполняется следующее: ${}^{\circ}W^{-1} = \tilde{\mathcal{L}}_{n_p}(F_{Y_p}^T \otimes F_{Y_p}^T)\tilde{\mathcal{D}}_{n_p} = {}^{\circ}W^T$. Матрица $\tilde{\mathcal{L}}_{n_p}D^{\otimes}(\Theta_{NB})\tilde{\mathcal{D}}_{n_p}$ – диагональная порядка n_p^{Δ} со строго положительными диагональными элементами.

5. ВСПОМОГАТЕЛЬНАЯ ЗАДАЧА ДОПОЛНИТЕЛЬНОСТИ

Полученные в разд. 2 ньютоновские направления ΔX_U и ΔY_U вычислялись в парах $[X, Y]$, которые являлись неособыми, т.е. пересечение нуль-пространств матриц X и Y имело нулевую размерность. Предположим, теперь, что пара $[X, Y]$ особая, в таких парах размерность указанного подпространства положительная. Предполагаем также, что $U = [U_p, U_B, U_N, U_Z]$ и ранг матрицы U_Z равен $n_Z > 0$.

Ньютоновская система (7) в особой паре видоизменяется

$$\begin{aligned} Y_{U_p} \circ \Delta X_{U_p} + X_{U_p} \circ \Delta Y_{U_p} &= -D(\lambda_p), \\ A_{i,U_p} \bullet \Delta X_{U_p} + A_{i,U_B} \bullet \Delta X_{U_B} + A_{i,U_Z} \bullet \Delta X_{U_Z} &= 0, \quad i \in J^m, \\ K_{j,U_p} \bullet \Delta Y_{U_p} + K_{j,U_N} \bullet \Delta Y_{U_N} + K_{j,U_Z} \bullet \Delta Y_{U_Z} &= 0, \quad j \in J^{l_v}. \end{aligned} \quad (54)$$

Здесь $A_{i,U_Z} = U_Z^T A_i U_Z$, $K_{j,U_Z} = U_Z^T K_j U_Z$, $\Delta X_{U_Z} = U_Z^T \Delta X U_Z$, $\Delta Y_{U_Z} = U_Z^T \Delta Y U_Z$. Векторное представление системы следующее:

$$\tilde{Y}_{U_p}^{\otimes} \text{svec} \Delta X_{U_p} + \tilde{X}_{U_p}^{\otimes} \text{svec} \Delta Y_{U_p} = -\text{svec} D(\lambda_p), \quad (55)$$

$$\mathcal{A}_{\text{svec}}^{U_p} \text{svec} \Delta X_{U_p} + \mathcal{A}_{\text{svec}}^{U_B} \text{svec} \Delta X_{U_B} + \mathcal{A}_{\text{svec}}^{U_Z} \text{svec} \Delta X_{U_Z} = 0_m, \quad (56)$$

$$\mathcal{K}_{\text{svec}}^{U_p} \text{svec} \Delta Y_{U_p} + \mathcal{K}_{\text{svec}}^{U_N} \text{svec} \Delta Y_{U_N} + \mathcal{K}_{\text{svec}}^{U_Z} \text{svec} \Delta Y_{U_Z} = 0_{l_v}, \quad (57)$$

где $\mathcal{A}_{\text{svec}}^{U_Z}$ и $\mathcal{K}_{\text{svec}}^{U_Z}$ – матрицы, строками которых являются соответственно векторы $\text{svec} A_i^{U_Z}$ и $\text{svec} K_j^{U_Z}$.

Система (55)–(57) недоопределена и, следовательно, имеет целое множество решений. Зафиксируем $\text{svec} \Delta X_{U_Z}$ и $\text{svec} \Delta Y_{U_Z}$. Тогда решение системы относительно остальных переменных единственное и может быть получено аналогично тому, как это делается в разд. 2. При этом выражения для ΔX_{U_p} , ΔX_{U_B} , ΔY_{U_p} и ΔY_{U_N} полностью сохраняются с единственной заменой b на $b_1 = b - \mathcal{A}_{\text{svec}}^{U_Z} \text{svec} \Delta X_{U_Z}$, т.е. становятся зависимыми от $\text{svec} \Delta X_{U_Z}$.

Пусть ΔX_{U_Z} и ΔY_{U_Z} – матрицы, которым соответствуют векторы $\text{svec} \Delta X_{U_Z}$ и $\text{svec} \Delta Y_{U_Z}$. Распорядимся приращением ΔX_{U_Z} таким образом, чтобы

$$\Delta X_{U_Z} \succeq 0, \quad \Delta Y_{U_Z} \succeq 0, \quad \Delta X_{U_Z} \bullet \Delta Y_{U_Z} = 0.$$

Обозначим через \mathcal{A}_{U_Z} матрицу $\mathcal{A}_{U_Z} = [A_{1,U_Z}; \dots; A_{m,U_Z}]$. Здесь, как и в случае с векторами, знак точки с запятой между матрицами означает, что эти матрицы помещаются одна под другой в виде матрицы-столбца. Под $\mathcal{A}_{U_Z} \bullet \Delta X_{U_Z}$ будем понимать m -мерный вектор

$$\mathcal{A}_{U_Z} \bullet \Delta X_{U_Z} = [A_{1,U_Z} \bullet \Delta X_{U_Z}; \dots; A_{m,U_Z} \bullet \Delta X_{U_Z}].$$

Согласно своему определению

$$\Delta Y_{U_Z} = -\sum_{i=1}^m A_{i,U_Z} \Delta u^i.$$

Теперь для Δu при $0 \leq n_B^{\Delta} < m$ выполняется

$$\Delta u = \tilde{\Phi}_p^{-1} \left(b - \mathcal{A}_{\text{svec}}^{U_Z} \text{svec} \Delta X_{U_Z} \right),$$

причем матрица $\tilde{\Phi}_p^{-1}$ заменяется на \mathcal{H}_p^{-1} , когда $n_B = 0$. Поэтому верно

$$\Delta Y_{U_Z} = -\sum_{i=1}^m (\tilde{\Phi}_p^{-1} b)^i A_{i,U_Z} + \sum_{i=1}^m [\tilde{\Phi}_p^{-1} (\mathcal{A}_{U_Z} \bullet \Delta X_{U_Z})]^i A_{i,U_Z}. \quad (58)$$

Обозначим далее

$$\lambda_{\mathcal{A},\tilde{\Phi}_p}(\Delta X_{U_Z}) = \tilde{\Phi}_p^{-1} [\mathcal{A}_{U_Z} \bullet \Delta X_{U_Z}] \in \mathbb{R}^m, \quad \Theta = -\sum_{i=1}^m (\tilde{\Phi}_p^{-1} b)^i A_{i,U_Z}$$

и определим линейное относительно ΔX_{U_Z} отображение $\Lambda_{\mathcal{A},\tilde{\Phi}_p}(\Delta X_{U_Z}) : S^{n_Z} \rightarrow S^{n_Z}$, положив

$$\Lambda_{\mathcal{A},\tilde{\Phi}_p}(\Delta X_{U_Z}) = \sum_{i=1}^m \lambda_{\mathcal{A},\tilde{\Phi}_p}^i(\Delta X_{U_Z}) A_{i,U_Z}.$$

Тогда согласно (58) имеем $\Delta Y_{U_Z} = \Lambda_{\mathcal{A},\tilde{\Phi}_p}(\Delta X_{U_Z}) + \Theta$.

Введем в рассмотрение вспомогательную задачу, которая является линейной полуопределенной задачей дополнителности

$$\begin{aligned} \Delta X_{U_Z} \bullet \Delta Y_{U_Z} &= 0, \\ \Delta Y_{U_Z} &= \Lambda_{\mathcal{A},\tilde{\Phi}_p}(\Delta X_{U_Z}) + \Theta, \\ \Delta X_{U_Z} &\succeq 0, \quad \Delta Y_{U_Z} \succeq 0. \end{aligned} \quad (59)$$

Утверждение 7. Пусть допустимая особая пара $[X, Y]$ такова, что матрицы

$$\begin{bmatrix} A_{i,U_B} & 0 \\ 0 & A_{i,U_Z} \end{bmatrix}, \quad i \in J^m, \quad (60)$$

линейно независимы. Тогда решение задачи (59) существует, причем единственное.

Доказательство. Поставим в соответствие задаче дополнителности (59) задачу квадратичного программирования

$$\begin{aligned} \Delta X_{U_Z} \bullet (\Lambda_{\mathcal{A},\tilde{\Phi}_p}(\Delta X_{U_Z}) + \Theta) &\rightarrow \min, \\ \Lambda_{\mathcal{A},\tilde{\Phi}_p}(\Delta X_{U_Z}) + \Theta &\succeq 0, \quad \Delta X_{U_Z} \succeq 0. \end{aligned} \quad (61)$$

Если взять $q = \mathcal{A}_{U_Z} \bullet \Delta X_{U_Z} \in \mathbb{R}^m$, то целевую функцию в задаче (61) можно записать как $\tilde{f}(q) = \langle q, \tilde{\Phi}_p^{-1} q \rangle - \langle \tilde{\Phi}_p^{-1} b, q \rangle$.

Покажем, что матрица $\tilde{\Phi}_p^{-1}$ такова, что $\langle q, \tilde{\Phi}_p^{-1} q \rangle > 0$ для любых ненулевых q . Действительно, так как матрица Φ_p^{-1} положительно определена, то $\langle q, \tilde{\Phi}_p^{-1} q \rangle = 0$ в том и только в том случае, когда $E_B^T q = 0$, т.е. вектор $q = \mathcal{A}_{U_Z} \bullet \Delta X_{U_Z} = \mathcal{A}_{\text{svec}}^{U_Z} \text{svec } \Delta X_{U_Z}$ принадлежит ядру матрицы E_B^T . Но данное ядро совпадает с образом матрицы $\mathcal{A}_{\text{svec}}^{U_B}$. Поэтому столбцы матрицы $[\mathcal{A}_{\text{svec}}^{U_B}, \mathcal{A}_{\text{svec}}^{U_Z}]$ оказываются линейно зависимыми, что противоречит предположению о линейной независимости матриц (60).

Матрица Φ_p^{-1} является положительно-определенной, поэтому квадратичная функция $\langle w, \Phi_p^{-1} w \rangle$, где $w = E_B^T q$, является сильно выпуклой и, следовательно, достигает своего минимального значения по w , причем единственного, на любом выпуклом замкнутом множестве. Отсюда следует достижение минимального значения функции

$$f(\Delta X_{U_Z}) = \Delta X_{U_Z} \bullet (\Lambda_{\mathcal{A},\tilde{\Phi}_p}(\Delta X_{U_Z}) + \Theta)$$

на допустимом множестве в задаче (61), т.е. задача (61) также имеет решение.

Согласно необходимым условиям минимума

$$f_{\Delta X_{U_Z}}(\Delta X_{U_Z}) \succeq 0, \quad X \bullet f_{\Delta X_{U_Z}}(\Delta X_{U_Z}) = 0,$$

где $f_{\Delta X_{U_Z}}(\Delta X_{U_Z})$ – градиент функции $f(\Delta X_{U_Z})$ по ΔX_{U_Z} . Данные условия можно переписать в виде

$$\Delta Y_{U_Z} = f_{\Delta X_{U_Z}}(\Delta X_{U_Z}) = \Lambda_{s,l,\Phi_p}(\Delta X_{U_Z}) + \Theta \succeq 0, \quad \Delta X_{U_Z} \bullet \Delta Y_{U_Z} = 0.$$

Таким образом, решение задачи квадратичного программирования (61) является одновременно решением задачи дополнителности (59). Утверждение доказано.

Замечание 1. Пусть $\tilde{\Phi}_p^{-1} = [\phi^{kl}]$, $1 \leq k, l \leq m$. Тогда производная компоненты $\lambda_{s,l,\Phi_p}^i(\Delta X_{U_Z})$ по $\Delta X_{U_Z}^{pq}$, $1 \leq p, 1 \leq q, \leq n_z$, равна

$$\frac{d\lambda_{s,l,\Phi_p}^i(\Delta X_{U_Z})}{dX_{U_Z}^{pq}} = \sum_{j=1}^m \phi^{ij} A_j^{pq}.$$

Отсюда следует, что (k, l) -й элемент градиента линейной функции $\Lambda_{s,l,\Phi_p}(\Delta X_{U_Z}) \bullet \Delta X_{U_Z}$ по $\Delta X_{U_Z}^{pq}$ равняется $\sum_{i,j=1}^m \phi^{ij} A_i^{kl} A_j^{pq}$.

Замечание 2. Если пара $[X, Y]$ невырожденная, то $n_p^\Delta + n_B^\Delta \geq m$. При выполнении предположения (60) должно выполняться неравенство $n_B^\Delta + n_Z^\Delta \leq m$. Отсюда следует оценка сверху на число n_z , определяющее размерность задачи дополнителности (59), а именно, $n_z \leq n_p$.

Решение задачи дополнителности (59) единственное. Подставляя найденные ΔX_{U_Z} и ΔY_{U_Z} в систему (54), получаем, что она становится зависимой только от остальных переменных и однозначным образом разрешима относительно них. Как уже отмечалось, формулы для определения этих направлений отличаются от полученных в разд. 2 только заменой вектора правых частей b на вектор b_1 .

ЗАКЛЮЧЕНИЕ

Предложенный алгоритм предназначен для решения линейных задач полуопределенного программирования в стандартной постановке. Вычисления ведутся в пространствах прямых переменных и слабых двойственных переменных, хотя несложно переписать формулы для вычислений в пространстве двойственных переменных. Все текущие приближения в итерационном процессе принадлежат допустимым множествам, поэтому метод можно трактовать как метод внутренней точки, но не строго внутренней точки, так как точки могут принадлежать границам допустимых множеств. Это принципиальное отличие от методов внутренней точки в традиционном понимании позволяет уменьшать вычислительные затраты при приближении к решениям прямой и двойственной задач. Делается это за счет более точного выполнения условия комплементарности между ними. Поэтому особенно эффективно алгоритм должен работать на конечной стадии вычислений, когда текущие точки находятся вблизи решений задач.

СПИСОК ЛИТЕРАТУРЫ

1. Жадан В.Г. Прямо-двойственный метод Ньютона с наискорейшим спуском для линейной задачи полуопределенного программирования. Ньютоновская система уравнений // Ж. вычисл. матем. и матем. физ. 2022. Т. 62. № 2. С. 22–38.
2. Magnus J.R., Neudecker H. The elimination matrix: some lemmas and applications // SIAM J. Alg. Disc. Meth. 1980. V. 1. № 4. P. 422–449.
3. Магнус Я.Р., Нейдеккер Ч. Матричное дифференциальное исчисление с приложениями к статистике и эконометрике. М.: Физматлит, 2002.
4. Nesterov Yu.E., Nemirovski A.S. Interior Point Polynomial Algorithms in Convex Programming // Philadelphia: SIAM Publications, SIAM, 1994.
5. Alizadeh F. Interior point methods in semidefinite programming with applications to combinatorial optimization // SIAM J. Optim. 1995. V. 5. P. 13–51.
6. Vandenberghe L., Boyd S. Semidefinite programming // SIAM Rev. 1996. V. 38. P. 49–95.
7. Todd M.J. Semidefinite optimization // Acta Numer. 2001. V. 10. P. 515–560.
8. Wolkowicz H., Saigal R., Vandenberghe L. (eds). Handbook of Semidefinite Programming. Boston: Kluwer Academic Publishers, 2000.

9. *Laurent M., Rendle F.* Semidefinite Programming and Integer Programming. Centrum voor Wiskunde en Informatica, 2002.
10. *de Klerk E.* Aspects of Semidefinite Programming. Interior Point Algorithms and Selected Applications. Kluwer Academic Publishers, 2004.
11. *Monteiro R.D.C.* First- and second-order methods for semidefinite programming // Mathematical programming, Series B. 2003. V. 97. № 1–2. P. 209–244.
12. *Жадан В.Г.* Прямо-двойственный метод Ньютона для задач линейного программирования // Ж. вычисл. матем. и матем. физ. 1999. Т. 39. № 1. С. 17–32.
13. *Жадан В.Г.* О сходимости прямо-двойственного метода Ньютона для задачи линейного программирования // Ж. вычисл. матем. и матем. физ. 1999. Т. 39. № 5. С. 431–445.
14. *Арнольд В.И.* О матрицах, зависящих от параметров // Успехи матем. наук. 1971. Т. 26. вып. 2 (158). С. 101–114.
15. *Alizadeh F., Haeblerly J.-P.F., Overton M.L.* Complementarity and nondegeneracy in semidefinite programming // Math. Program., Series B. 1997. V. 77. № 2. P. 129–162.
16. *Тыртышников Е.Е.* Матричный анализ и линейная алгебра. М.: Физматлит, 2007.