В. Г. Жадан, А. А. Орлов, Двойственные методы внутренней точки для линейной задачи полуопределенного программирования, *Ж. вычисл. матем. и матем. физ.*, 2011, том 51, номер 12, 2158–2180

ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2011, том 51, № 12, с. 2158–2180

УДК 519.658.4

ДВОЙСТВЕННЫЕ МЕТОДЫ ВНУТРЕННЕЙ ТОЧКИ ДЛЯ ЛИНЕЙНОЙ ЗАДАЧИ ПОЛУОПРЕДЕЛЕННОГО ПРОГРАММИРОВАНИЯ[1)]

© 2011 г. В. Г. Жадан, А. А. Орлов

(119333 Москва, ул. Вавилова, 40, ВЦ РАН)

e-mail: zhadan@ccas.ru

Поступила в редакцию 31.05.2011 г.

Рассматривается линейная задача полуопределенного программирования. Для ее решения предлагаются двойственные методы внутренней точки, являющиеся обобщением барьерно-проективных двойственных методов для задач линейного программирования. Показывается, что при условии невырожденности решений прямой и двойственной задач методы обладают локальной сходимостью с линейной скоростью. Библ. 20.

Ключевые слова: задача полуопределенного программирования, двойственный метод, метод внутренней точки, локальная сходимость.

1. ВВЕДЕНИЕ

Линейная задача полуопределенного программирования заключается в нахождении положительно-полуопределенной матрицы, доставляющей минимум линейной целевой функции на допустимом множестве с задаваемыми линейными ограничениями-равенствами. К такой постановке сводится решение многих других оптимизационных задач, в частности задач квадратичного программирования при квадратичных ограничениях, задач комбинаторной и структурной оптимизации, оптимизационных задач на собственные значения (см. [1]–[5]). Поэтому методам решения линейных задач полуопределенного программирования в последнее время уделялось огромное внимание. Были разработаны алгоритмы, главным образом типа методов внутренней точки, которые позволяют достаточно эффективно решать такие задачи (см. [5]–[11]).

В настоящей работе рассматриваются два численных метода решения линейной задачи полуопределенного программирования, которые также можно отнести к методам внутренней точки, точнее к двойственным методам аффинно-масштабирующего типа. Предлагаемые методы являются, с одной стороны, обобщением двойственных барьерно-проективных методов, разработанных ранее для решения задач линейного программирования (см. [12]), на более сложные задачи полуопределенного программирования. С другой стороны, их можно также рассматривать как двойственные аналоги прямого метода внутренней точки, предложенного в [13].

Хотя рассматриваемые методы обладают многими свойствами методов внутренней точки, при их построении совершенно не используется идея барьерных функций. Фактически методы получаются как специальные способы решения системы необходимых и достаточных условий для пары взаимно двойственных линейных задач полуопределенного программирования. Главная цель работы состоит в том, чтобы показать сохранение локальной сходимости методов при их перенесении с задач линейного программирования на задачи полуопределенного программирования. Как известно, иногда обобщение методов линейного программирования на линейные задачи полуопределенного программирования нарушают сходимость (см. [14]).

В разд. 2 даются постановки прямой и двойственной линейных задач полуопределенного программирования. В ней также приводятся сведения о некоторых специальных матрицах, используемых в дальнейшем в работе. В разд. 3 описывается двойственный итерационный метод, в котором от начальных точек требуется, чтобы они были допустимыми. Более общий метод, в котором это требование снимается, рассматривается в разд. 4. В разд. 5 показывается, что в случае невырожденности двойственной задачи алгоритмические операторы предлагаемых итерацион-

1) Работа выполнена при финансовой поддержке РФФИ (код проекта 11-01-00786), а также при содействии Программы ведущих научных школ НШ-4006.2010.1 и Программы президиума РАН П-17.

ных процессов корректно определены. Разд. 5 посвящен выводу формул для представления алгоритмических операторов в относительно граничных точках допустимого множества в двойственной задаче, которому, как правило, принадлежит решение. Локальная сходимость метода доказывается в разд. 7 с привлечением результатов из предыдущего разд. 6.

2. ПОСТАНОВКА ЗАДАЧИ И ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

Пусть $\mathbb{S}^n$ – пространство симметричных матриц порядка n, и пусть $\mathbb{S}^n_+$ – конус в $\mathbb{S}^n$, состоящий из положительно-полуопределенных матриц. Внутренность конуса $\mathbb{S}^n_+$, обозначаемая через $\mathbb{S}^n_{++}$, состоит из положительно-определенных матриц и также является конусом. В дальнейшем используются дополнительно неравенства $M \succeq 0$ или $M \succ 0$, чтобы указать на принадлежность квадратной матрицы M, соответственно, конусам $\mathbb{S}^n_+$ или $\mathbb{S}^n_{++}$. Пространство $\mathbb{S}^n$ конечномерно, его размерность равняется так называемому "треугольному числу" $k_\Delta(n) = n(n+1)/2$.

Скалярное (внутреннее) произведение между двумя матрицами L и M одного размера определяется как след матрицы $L^{\mathrm{T}}M$ и обозначается через

$$L \bullet M = \mathrm{Tr}(L^{\mathrm{T}}M) = \sum_{i,j} l_{ij}m_{ij},$$

где l_{ij} и m_{ij} суть (ij)-е элементы, соответственно, матриц L и M. Если L и M – две положительно-полуопределенные матрицы из $\mathbb{S}^n$, то обязательно $L \bullet M \geq 0$. Более того, $L \bullet M = 0$ в том и только том случае, когда $LM = ML = 0_{nn}$.

Рассмотрим линейную задачу полуопределенного программирования

$$\begin{gathered} \min C \bullet X, \\ A_i \bullet X = b^i, \quad i = 1, 2, \ldots, m, \\ X \succeq 0, \end{gathered} \tag{1}$$

где матрицы C, X, а также A_i, $1 \leq i \leq m$, принадлежат пространству $\mathbb{S}^n$. Двойственной к (1) является задача

$$\begin{gathered} \max \langle b, u\rangle, \\ \sum_{i=1}^m u^i A_i + V = 0, \quad V \succeq 0, \end{gathered} \tag{2}$$

в которой $b = (b^1, \ldots, b^m)^{\mathrm{T}} \in \mathbb{R}^m$ и $u \in \mathbb{R}^m$ и $V \in \mathbb{S}^n$. Угловые скобки указывают на обычное евклидово скалярное произведение в конечномерном векторном пространстве. Предполагается, что задачи (1) и (2) имеют решения и что матрицы A_i, $1 \leq i \leq m$, линейно независимы.

Обозначим допустимые множества в прямой и двойственной задачах, соответственно, через $\mathcal{F}_P$ и $\mathcal{F}_D$, т.е.

$$\mathcal{F}_P = \{X \in \mathbb{S}^n_+ : A_i \bullet X = b^i, i = 1, 2, \ldots, m\},$$

$$\mathcal{F}_D = \left\{[u, V] \in \mathbb{R}^m \times \mathbb{S}^n_+ : V = C - \sum_{i=1}^m u^i A_i\right\}.$$

Через $\mathcal{F}_{D,u}$ и $\mathcal{F}_{D,V}$ обозначим также проекции множества $\mathcal{F}_D$ на пространство $\mathbb{R}^m$ и конус $\mathbb{S}^n_+$ соответственно:

$$\mathcal{F}_{D,u} = \{u \in \mathbb{R}^m : [u, V] \in \mathcal{F}_D \text{ для некоторого } V \in \mathbb{S}^n_+\}, \tag{3}$$

$$\mathcal{F}_{D,V} = \{V \in \mathbb{S}^n_+ : [u, V] \in \mathcal{F}_D \text{ для некоторого } u \in \mathbb{R}^m\}. \tag{4}$$

Для любых $X \in \mathcal{F}_P$ и $u \in \mathcal{F}_{D,u}$ выполняется неравенство *слабой двойственности* $C \bullet X \geq \langle b, u\rangle$.

В случае, когда $C \bullet X = \langle b, u\rangle$, для некоторых $X \in \mathcal{F}_P$ и $u \in \mathcal{F}_{D,u}$ имеет место *строгая двойственность*. Условия регулярности ограничений Слейтера для обеих задач (1) и (2), т.е. требование, чтобы множества

$$\mathcal{F}^0_P = \{X \in \mathbb{S}^n_{++} : A_i \bullet X = b^i, i = 1, 2, \ldots, m\},$$

$$\mathcal{F}^0_D = \left\{[u, V] \in \mathbb{R}^m \times \mathbb{S}^n_{++} : V = C - \sum_{i=1}^m u^i A_i\right\}$$

были непустыми, позволяют гарантировать существование строгой двойственности (см., например, [6]). В этом случае решения задач (1) и (2) обязательно существуют.

Пусть $\mathrm{Diag}(a^1, \ldots, a^n)$ или просто $D(a)$ – диагональная матрица с компонентами вектора $a = [a^1, \ldots, a^n]$ на диагонали. Если X_* и $[u_*, V_*]$ – оптимальные решения, соответственно, задач (1) и (2), то $X_* \bullet V_* = 0$. Но для симметричных положительно-полуопределенных матриц X_* и V_* данное равенство выполняется тогда и только тогда, когда $X_*V_* = V_*X_* = 0_{nn}$. Отсюда заключаем, что матрицы X_* и V_* коммутируют. Поэтому найдется такая ортогональная матрица Q, что

$$X_* = QD(\eta_*)Q^{\mathrm{T}}, \quad V_* = QD(\theta_*)Q^{\mathrm{T}}, \tag{5}$$

где $\eta_* = [\eta^1_*, \ldots, \eta^n_*]$ и $\theta_* = [\theta^1_*, \ldots, \theta^n_*]$ – собственные значения матриц X_* и V_* соответственно. Для самих собственных значений η^i_* и θ^i_* выполняется условие дополнительности

$$\eta^i_*\theta^i_* = 0, \quad 1 \leq i \leq n. \tag{6}$$

Условие строгой дополнительности означает, что для каждого $1 \leq i \leq n$ одно из значений η^i_* или θ^i_* строго положительно. В этом случае решения X_* и V_* назовем *строго комплементарными*.

Приведем некоторые обозначения, которые будут использоваться в дальнейшем. Единичная матрица порядка n обозначается через I_n. Символ $\otimes$ между матрицами означает их произведение по Кронекеру.

Если M – квадратная матрица порядка n, то символом $\mathrm{vec}\,M$ обозначается прямая сумма ее столбцов, т.е. вектор-столбец длины n^2, в котором последовательно один под другим располагаются столбцы матрицы M. Для симметричных матриц имеет смысл вместо вектор-столбца $\mathrm{vec}\,M$ рассматривать вектор-столбец $\mathrm{vech}\,M$. В него также помещаются последовательно сверху вниз столбцы матрицы M, но не полностью, а только их нижние части, начинающиеся с диагонального элемента. Аналогичным образом определяется вектор-столбец $\mathrm{vecs}\,M$. От $\mathrm{vech}\,M$ он отличается только тем, что все элементы, не стоящие на диагонали матрицы M, при помещении в $\mathrm{vecs}\,M$ умножаются на два. Как вектор $\mathrm{vech}\,M$, так и вектор $\mathrm{vecs}\,M$ имеют длину, равную $k_\Delta(n)$.

Для перехода от вектора $\mathrm{vec}\,M$ к вектору $\mathrm{vech}\,M$ и для обратного перехода используются специальные *элиминационные* и *дуплицирующие* матрицы (см. [15], [16]). Элиминационная матрица $\mathcal{L}_n$ для каждой симметричной матрицы M порядка n осуществляет преобразование $\mathcal{L}_n\mathrm{vec}\,M = \mathrm{vech}\,M$. Напротив, дуплицирующая матрица $\mathcal{D}_n$ для каждой симметричной матрицы M порядка n осуществляет обратное преобразование $\mathcal{D}_n\mathrm{vech}\,M = \mathrm{vec}\,M$. Матрица $\mathcal{L}_n$ имеет размер $k_\Delta(n) \times n^2$, матрица $\mathcal{D}_n$ – размер $n^2 \times k_\Delta(n)$. Обе матрицы $\mathcal{L}_n$ и $\mathcal{D}_n$ являются матрицами полного ранга, равного $k_\Delta(n)$. Матрица $\mathcal{L}_n$ полуортогональна, т.е. $\mathcal{L}_n\mathcal{L}_n^{\mathrm{T}} = I_{k_\Delta(n)}$. Кроме того, $\mathcal{L}_n\mathcal{D}_n = I_{k_\Delta(n)}$.

Для любой квадратной матрицы M порядка n справедливы формулы (см. [15])

$$\mathcal{D}_n\mathcal{L}_n(M \otimes M)\mathcal{D}_n = (M \otimes M)\mathcal{D}_n, \tag{7}$$

$$\mathcal{D}_n\mathcal{L}_n(M \otimes I_n + I_n \otimes M)\mathcal{D}_n = (M \otimes I_n + I_n \otimes M)\mathcal{D}_n. \tag{8}$$

Более того, если матрица M неособая, то

$$[\mathcal{L}_n(M \otimes M)\mathcal{D}_n]^{-1} = \mathcal{L}_n(M^{-1} \otimes M^{-1})\mathcal{D}_n. \tag{9}$$

При работе с векторами вида $\operatorname{vec} M$ удобно нумеровать их элементы не числами натурального ряда, а парами индексов из следующего набора:

$$J = \{(1, 1), \ldots, (n, 1), (1, 2), \ldots, (n, 2), \ldots, (1, n), \ldots, (n, n)\}. \tag{10}$$

Всего в таком наборе n^2 пар индексов. Номера вида (i, i), в которых первый индекс совпадает со вторым, будем называть *диагональными*. Набор J содержит n таких диагональных номеров, а именно $(1, 1), (2, 2), \ldots, (n, n)$. Те же номера (i, j) из J, в которых индексы i и j отличны друг от друга, будем называть *внедиагональными*. О внедиагональных номерах (i, j) и (j, i) будем говорить как о *симметричных внедиагональных* номерах. При этом если $i > j$, то такой номер (i, j) называется *младшим внедиагональным* номером. Напротив, симметричный номер (j, i) называется *старшим внедиагональным* номером.

Наряду с набором (10) используется также набор

$$J_\Delta = \{(1, 1), \ldots, (n, 1), (2, 2), \ldots, (n, 2), (3, 3), \ldots, (n, n)\},$$

соержащий $k_\Delta(n)$ пар индексов. В него входят только диагональные номера и младшие внедиагональные номера. Данный набор парных индексов используется главным образом при работе с векторами $\operatorname{vech} M$ или $\operatorname{vecs} M$.

3. ДОПУСТИМЫЙ ИТЕРАЦИОННЫЙ ПРОЦЕСС

Мы предположили, что решения обеих задач (1) и (2) существуют. Тогда, в силу необходимых и достаточных условий, система равенств и неравенств

$$\begin{gathered} X \bullet V = 0, \\ A_i \bullet X = b^i, \quad 1 \le i \le m, \\ V = C - \sum_{i=1}^{m} u^i A_i, \\ X \succeq 0, \quad V \succeq 0, \end{gathered} \tag{11}$$

обязательно имеет решение.

Обозначим через $X * V$ симметризованное произведение матриц X и V из $\mathbb{S}^n$, т.е. матрицу $X * V = (XV + VX)/2$.

Утверждение 1. *Для симметричных матриц $X \succeq 0$ и $V \succeq 0$ равенство $X * V = 0_{nn}$ возможно в том и только том случае, когда $XV = VX = 0_{nn}$.*

Доказательство. Покажем только *необходимость*. Пусть $X * V = 0_{nn}$. Тогда $XV = -(XV)^{\mathrm{T}}$, следовательно, матрица XV кососимметричная. У вещественной кососимметричной матрицы на диагонали расположены нулевые элементы, поэтому $X \bullet V = \operatorname{tr}(XV) = 0$. Так как $X \succeq 0$ и $V \succeq 0$, отсюда заключаем, что $XV = VX = 0_{nn}$. Утверждение доказано.

С учетом утверждения 1 система (11) может быть переписана в виде

$$\begin{gathered} X * V = 0_{nn}, \\ A_i \bullet X = b^i, \quad 1 \le i \le m, \\ V = C - \sum_{i=1}^{m} u^i A_i, \\ X \succeq 0, \quad V \succeq 0. \end{gathered}$$

Заменяя в ней матричные равенства на их векторные аналоги, получаем

$$\begin{aligned} \mathrm{vec}(X * V) &= 0_{nn}, \\ \mathcal{A}_{\mathrm{vec}}\mathrm{vec}\,X &= b, \\ \mathrm{vec}\,V &= \mathrm{vec}\,C - \mathcal{A}_{\mathrm{vec}}^{\mathrm{T}}u, \\ X \succeq 0, &\quad V \succeq 0. \end{aligned} \tag{12}$$

Здесь и ниже через $\mathcal{A}_{\mathrm{vec}}$ обозначается $m \times n^2$-матрица, строками которой являются векторы $\mathrm{vec}A_i$, $1 \le i \le m$.

Но в силу известной формулы

$$\mathrm{vec}(ABC) = (C^{\mathrm{T}} \otimes A)\mathrm{vec}\,B, \tag{13}$$

справедливой для любых матриц A, B и C, для которых определено произведение ABC, получаем, что $\mathrm{vec}(X * V) = V^{\otimes}\mathrm{vec}X$, где

$$V^{\otimes} = [V \otimes I_n + I_n \otimes V]/2$$

есть кронекеровская сумма матрицы V. Таким образом, система условий (12) может быть представлена в виде

$$\begin{aligned} V^{\otimes}\mathrm{vec}X &= 0_{n^2}, \\ \mathcal{A}_{\mathrm{vec}}\mathrm{vec}\,X &= b, \\ \mathrm{vec}\,V &= \mathrm{vec}\,C - \mathcal{A}_{\mathrm{vec}}^{\mathrm{T}}u, \\ X \succeq 0, &\quad V \succeq 0. \end{aligned} \tag{14}$$

Умножим теперь второе уравнение в (14) на матрицу $\mathcal{A}_{\mathrm{vec}}^{\mathrm{T}}$:

$$\mathcal{A}_{\mathrm{vec}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vec}}\mathrm{vec}X = \mathcal{A}_{\mathrm{vec}}^{\mathrm{T}}b. \tag{15}$$

Складывая данное равенство с первым равенством из (14), получаем линейное уравнение относительно $\mathrm{vec}\,X$:

$$\tilde{\Phi}(V)\mathrm{vec}X = \mathcal{A}_{\mathrm{vec}}^{\mathrm{T}}b, \tag{16}$$

где $\tilde{\Phi}(V) = \mathcal{A}_{\mathrm{vec}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vec}} + V^{\otimes}$ – симметричная матрица порядка n^2.

Пусть $X = X(V)$ – решение уравнения (16). Так как матрица V и все матрицы A_i, $1 \le i \le m$, симметричные, то матрицу $X(V)$ также можно взять симметричной. Тогда от вектора $\mathrm{vec}\,X$ целесообразно перейти к вектору $\mathrm{vech}X$. Имеем, согласно сказанному выше, $\mathrm{vec}\,X = \mathcal{D}_n\mathcal{L}_n\mathrm{vec}\,X = \mathcal{D}_n\mathrm{vech}X$. После умножения левой и правой части (16) на матрицу $\mathcal{L}_n$ приходим к уравнению

$$\Phi(V)\mathrm{vech}X = \mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}b. \tag{17}$$

Здесь через $\Phi(V)$ обозначена матрица $\mathcal{L}_n\tilde{\Phi}(V)\mathcal{D}_n$, приводимая к виду

$$\Phi(V) = \mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vecs}} + \mathcal{L}_nV^{\otimes}\mathcal{D}_n. \tag{18}$$

В (17) и (18) символы $\mathcal{A}_{\mathrm{vech}}$ и $\mathcal{A}_{\mathrm{vecs}}$ используются для обозначения $m \times k_{\Delta}(n)$-матриц, строками которых являются, соответственно, векторы $\mathrm{vech}A_i$ и $\mathrm{vecs}A_i$, $1 \le i \le m$.

Матрица $\Phi(V)$ квадратная порядка $k_{\Delta}(n)$. Если она неособая, то, разрешая уравнение (17), получаем $\mathrm{vech}X = \Phi^{-1}(V)\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}b$, или

$$\mathrm{vec}\,X = \mathcal{D}_n\Phi^{-1}(V)\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}b. \tag{19}$$

Таким образом, чтобы удовлетворить условию (16), в качестве $X = X(V)$ может быть взята такая симметричная матрица, прямая сумма столбцов которой есть вектор (19).

Пусть теперь $V = V(u) = C - \Sigma_{i=1}^{m} u^i A_i$. Тогда зависимость $X(V)$ фактически становится зависимостью $X(u) = X(V(u))$. Если подставить ее во второе равенство из (14) и учесть, что $\mathcal{A}_{\mathrm{vec}}\mathcal{D}_n = \mathcal{A}_{\mathrm{vecs}}$, то приходим к системе m уравнений относительно m переменных:

$$[\mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(V(u))\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} - I_m]b = 0_m. \tag{20}$$

Применим для решения системы (20) метод простой итерации:

$$u_{k+1} = u_k + \alpha_k[I_m - \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(V_k))\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}]b, \quad V_k = V(u_k), \tag{21}$$

где $\alpha_k > 0$. Пусть точка $u_0 \in \mathbb{R}^m$ такова, что $V_0 = V(u_0) \succ 0$. Тогда за счет выбора шага α_k можно добиться выполнения неравенства $V_k \succ 0$ на всех итерациях. Метод в этом случае оказывается методом внутренней точки, на всех итерациях двойственная переменная u_k принадлежит множеству $\mathcal{F}_{D,u}^{0} = \{u \in \mathcal{F}_{D,u} : V(u) \succ 0\}$.

Пусть $\mathcal{F}_{D,V}^{0} = \{V \in \mathcal{F}_{D,V} : V \succ 0\}$. В пространстве двойственной переменной V итерационный процесс (21) принимает вид

$$\mathrm{vech}\,V_{k+1} = \mathrm{vech}\,V_k - \alpha_k\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}[I_m - \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(V_k))\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}]b. \tag{22}$$

От начальной точки V_0 теперь требуется, чтобы $V_0 \in \mathcal{F}_{D,V}^{0}$. При надлежащем выборе шага α_k опять итерационный процесс (22) является методом внутренней точки, т.е. $V_k \in \mathcal{F}_{D,V}^{0}$ на всех итерациях.

Положительная определенность всех матриц V_k, когда $V_k \in \mathcal{F}_{D,V}^{0}$, позволяет несколько упростить правую часть в итерационном процессе (22). Действительно, в этом случае матрица $V_k^{\otimes}$ также является положительно-определенной и у матрицы $\mathcal{L}_n V_k^{\otimes}\mathcal{D}_n$ существует обратная матрица, которая согласно [15] имеет вид $(\mathcal{L}_n V_k^{\otimes}\mathcal{D}_n)^{-1} = \mathcal{L}_n(V_k^{\otimes})^{-1}\mathcal{D}_n$.

Отсюда, используя формулу Шермана–Моррисона–Вудберри, получаем

$$\begin{gathered}\Phi^{-1}(V_k) = (\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vecs}} + \mathcal{L}_n V_k^{\otimes}\mathcal{D}_n)^{-1} = \\ = \mathcal{L}_n(V_k^{\otimes})^{-1}\mathcal{D}_n - \mathcal{L}_n(V_k^{\otimes})^{-1}\mathcal{D}_n\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}(I_m + \mathcal{A}_{\mathrm{vecs}}\mathcal{L}_n(V_k^{\otimes})^{-1}\mathcal{D}_n\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}})^{-1}\mathcal{A}_{\mathrm{vecs}}\mathcal{L}_n(V_k^{\otimes})^{-1}\mathcal{D}_n.\end{gathered} \tag{23}$$

Положим $\mathcal{X} = \mathcal{A}_{\mathrm{vecs}}\mathcal{L}_n(V_k^{\otimes})^{-1}\mathcal{D}_n\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}$. Тогда на основании (23) имеем

$$\mathcal{A}_{\mathrm{vech}}\Phi^{-1}(V_k)\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} = \mathcal{X} - \mathcal{X}(I_m + \mathcal{X})^{-1}\mathcal{X} = I_m - (I_m + \mathcal{X})^{-1}.$$

Поэтому формула пересчета (22) в точках $V_k \in \mathcal{F}_{D,V}^{0}$ может быть записана в виде

$$\mathrm{vech}\,V_{k+1} = \mathrm{vech}\,V_k - \alpha_k\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}[I_m + \mathcal{A}_{\mathrm{vecs}}\mathcal{L}_n(V_k^{\otimes})^{-1}\mathcal{D}_n\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}]^{-1}b, \tag{24}$$

где, как уже отмечалось, $V_0 \in \mathcal{F}_{D,V}^{0}$.

Соответствующий процесс (21) относительно переменной u также может быть представлен в более простом виде:

$$u_{k+1} = u_k + \alpha_k[I_m + \mathcal{A}_{\mathrm{vecs}}\mathcal{L}_n(V_k^{\otimes})^{-1}\mathcal{D}_n\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}]^{-1}b; \tag{25}$$

здесь $u_0 \in \mathcal{F}_{D,u}^{0}$.

4. ОБЩИЙ ИТЕРАЦИОННЫЙ ПРОЦЕСС

Рассмотрим другую, более общую зависимость переменной X от пары двойственных переменных u и V. Это позволит построить соответствующий итерационный процесс в объединенном пространстве $\mathbb{R}^m \times \mathbb{S}^n$, в котором уже не обязательно сохраняется равенство $V_k = V(u_k)$. С этой целью сложим вместе уравнение (15), первое равенство из (14) и третье равенство из (14), предварительно умноженное на произвольную константу $\tau > 0$. В результате получим

$$\tilde{\Phi}(V)\mathrm{vec}X = \mathcal{A}_{\mathrm{vec}}^{\mathrm{T}} b + \tau(\mathrm{vec}V + \mathcal{A}_{\mathrm{vec}}^{\mathrm{T}} u - \mathrm{vec}C) \tag{26}$$

или, после перехода от вектора $\mathrm{vec}X$ к вектору $\mathrm{vech}X$,

$$\Phi(V)\mathrm{vech}X = \mathcal{F}(u, \mathrm{vech}V), \tag{27}$$

где для упрощения записи введено обозначение

$$\mathcal{F}(u, \mathrm{vech}V) = \mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} b + \tau(\mathrm{vech}V + \mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} u - \mathrm{vech}C).$$

Разрешая уравнение (27) относительно X, получаем

$$\mathrm{vec}X = \mathcal{D}_n\Phi^{-1}(V)\mathcal{F}(u, \mathrm{vech}V). \tag{28}$$

Переменная X становится теперь функцией $X = X(u, V)$. Данная зависимость переходит в прежнюю зависимость (19), если в (26) положить $\tau = 0$ или если взять $V = V(u)$.

Подставим найденную согласно (28) зависимость $X(u, V)$ в первые два равенства из (14):

$$\begin{aligned} \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(V)\mathcal{F}(u, \mathrm{vech}V) - b &= 0_m, \\ V^{\otimes}\mathcal{D}_n\Phi^{-1}(V)\mathcal{F}(u, \mathrm{vech}V) &= 0_{n^2}. \end{aligned} \tag{29}$$

Данная система состоит из $m + n^2$ уравнений. На самом деле она может быть редуцирована к системе из меньшего числа уравнений. Это следует из того, что левая часть второго уравнения в (29) является прямой суммой столбцов симметричной матрицы. Поэтому систему (29) можно заменить на систему из $m + k_\Delta(n)$ уравнений вида

$$\begin{aligned} \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(V)\mathcal{F}(u, \mathrm{vech}V) - b &= 0_m, \\ \mathcal{L}_n V^{\otimes}\mathcal{D}_n\Phi^{-1}(V)\mathcal{F}(u, \mathrm{vech}V) &= 0_{k_\Delta(n)}, \end{aligned} \tag{30}$$

в которой число неизвестных также равно $m + k_\Delta(n)$.

Снова применим для решения системы (30) метод простой итерации, тогда придем к следующему итерационному процессу:

$$\begin{aligned} u_{k+1} &= u_k + \alpha_k[b - \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(V_k)\mathcal{F}(u_k, \mathrm{vech}V_k)], \\ \mathrm{vech}V_{k+1} &= \mathrm{vech}V_k - \alpha_k\mathcal{L}_n V_k^{\otimes}\mathcal{D}_n\Phi^{-1}(V_k)\mathcal{F}(u_k, \mathrm{vech}V_k). \end{aligned} \tag{31}$$

Потребуем от начальной точки $[u_0, V_0]$, чтобы $V_0 \succ 0$. Если $V_k \succ 0$, то за счет выбора шага α_k можно добиться, чтобы имело место $V_{k+1} \succ 0$. В этом случае аналогично тому, как это делалось в первом варианте метода, можно упростить формулы пересчета. А именно, подставляя выражение (23) для $\Phi^{-1}(V_k)$ и проводя последовательно выкладки, получаем

$$u_{k+1} = u_k + \alpha_k\{(I + \mathcal{Z}_k)^{-1}b + \tau[u_k - (I + \mathcal{Z}_k)^{-1}(u_k - \mathcal{A}_{\mathrm{vecs}}\mathcal{R}_k^{-1}\mathrm{vech}(V_k - C))]\}, \tag{32}$$

$$\begin{aligned} \mathrm{vech}V_{k+1} = {} & \mathrm{vech}V_k - \alpha_k\{\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}(I + \mathcal{Z}_k)^{-1}b + \\ & + \tau[\mathrm{vech}(V_k - C) + \mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}(I_m + \mathcal{Z}_k)^{-1}(u_k - \mathcal{A}_{\mathrm{vecs}}\mathcal{R}_k^{-1}\mathrm{vech}(V_k - C))]\}. \end{aligned} \tag{33}$$

Здесь для сокращения записи использованы обозначения

$$\mathcal{R}_k = \mathcal{L}_n V_k^{\otimes}\mathcal{D}_n, \quad \mathcal{Z}_k = \mathcal{A}_{\mathrm{vecs}}\mathcal{R}_k^{-1}\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}.$$

Заметим, что итерационный процесс (31) полностью соответствует процессу в векторно-матричной форме:

$$u_{k+1} = u_k + \alpha_k[b - \mathcal{A} \bullet X_k], \quad V_{k+1} = V_k - \alpha_k V_k * X_k, \tag{34}$$

в котором $X_k = X(u_k, V_k)$, $\mathcal{A}$ есть $(mn \times n)$-матрица, составленная из матриц A_i, $1 \le i \le m$:

$$\mathcal{A} = \begin{bmatrix} A_1 \\ \dots \\ A_m \end{bmatrix}.$$

Запись $\mathcal{A} \bullet X$ означает попарное умножение матриц A_i и X, т.е. результатом операции $\mathcal{A} \bullet X$ является m-мерный вектор с компонентами $A_i \bullet X$, $1 \le i \le m$.

Итерационную формулу для пересчета векторов V_k в процессе (34) можно представить в несколько ином виде. Это, в частности, позволит оценить величину шага α_k, который сохранял бы положительную определенность матрицы V_{k+1}.

Пусть $Y_k = V_k * X_k$. Обе матрицы V_k и Y_k являются симметричными. Предположим дополнительно, что матрица V_k положительно-определенная. Тогда, используя конгруэнтное приведение двух симметричных матриц (см. [17]), одна из которых положительно-определенная, для некоторой невырожденной матрицы P получаем

$$P^{\mathrm{T}} V_k P = I_n, \quad P^{\mathrm{T}} Y_k P = D(\omega_k), \tag{35}$$

где $\omega_k = [\omega_k^1, \ldots, \omega_k^n]$ – вектор, составленный из собственных значений матрицы $V_k^{-1} Y_k$. Все эти собственные значения действительные и, согласно (35), $V_k = (P^{-1})^{\mathrm{T}} P^{-1}$, $Y_k = (P^{-1})^{\mathrm{T}} D(\omega_k) P^{-1}$. Поэтому правую формулу в (34) можно переписать в виде

$$V_{k+1} = (P^{-1})^{\mathrm{T}} [I_n - \alpha_k D(\omega_k)] P^{-1}.$$

Отсюда видно, что матрица V_{k+1} будет положительно-определенной в том и только том случае, когда $e - \alpha_k \omega_k > 0_n$, где e есть n-мерный вектор, состоящий из единиц. Пусть $\hat{\omega}_k$ – максимальное положительное собственное значение из набора $\omega_k^1, \ldots, \omega_k^n$, и пусть $\hat{\alpha}_k = 1/\hat{\omega}_k$. Если положительных собственных значений нет, то полагаем $\hat{\alpha}_k = +\infty$. Тогда указанное неравенство $e - \alpha_k \omega_k > 0_n$ выполняется, если $\alpha_k < \hat{\alpha}_k$.

5. НЕВЫРОЖДЕННОСТЬ МАТРИЦЫ $\Phi(V)$

Предположим, что ранг матрицы $V \in \mathbb{S}_+^n$ равняется r. Тогда она может быть представлена в виде

$$V = Q \mathrm{Diag}(\theta^1, \ldots, \theta^r, 0, \ldots, 0) Q^{\mathrm{T}}, \tag{36}$$

где Q – ортогональная матрица, $\theta^i > 0$, $1 \le i \le r$. Если $r < n$, то V принадлежит границе конуса $\mathbb{S}_+^n$. Касательное пространство к $\mathbb{S}_+^n$ в этой точке V имеет следующий вид (см. [18]):

$$\mathcal{T}_V = \left\{ Q \begin{bmatrix} G & H \\ H^{\mathrm{T}} & 0 \end{bmatrix} Q^{\mathrm{T}} : G \in \mathbb{S}^r, H \in \mathbb{R}^{r \times (n-r)} \right\}.$$

Здесь $\mathbb{R}^{k \times l}$ – пространство $(k \times l)$-матриц.

Предположим далее, что матрица X ранга $n - r$ также принадлежит $\mathbb{S}_+^n$ и для нее имеет место представление

$$X = Q \mathrm{Diag}(0, \ldots, 0, \eta^{r+1}, \ldots, \eta^n) Q^{\mathrm{T}},$$

в котором по-прежнему Q – ортогональная матрица (быть может, отличная от матрицы Q из (36)), а $\eta^i > 0$, $r < i \le n$. Если $1 \le r < n$, то точка X также является граничной точкой $\mathbb{S}^n_+$, но касательное пространство к $\mathbb{S}^n_+$ в этой точке теперь заменяется на следующее:

$$\mathcal{T}_X = \left\{ Q\begin{bmatrix} 0 & H \\ H^{\mathrm{T}} & G \end{bmatrix} Q^{\mathrm{T}} : G \in \mathbb{S}^{n-r}, H \in \mathbb{R}^{r\times(n-r)} \right\}.$$

Пусть $\mathcal{R}_A$ – подпространство в $\mathbb{S}^n$, порожденное матрицами A_i, $1 \le i \le m$, и пусть $\mathcal{R}_A^\perp$ – его ортогональное дополнение. Следуя [19], даем определение невырожденной точки V в двойственной задаче (2) и невырожденной точки X в прямой задаче (1).

Определение 1. Точка $V \in \mathcal{F}_{D,V}$ называется *невырожденной* в двойственной задаче (2), если $\mathcal{T}_V + \mathcal{R}_A = \mathbb{S}^n$. Соответственно, точка $X \in \mathcal{F}_P$ называется *невырожденной* в прямой задаче (1), если $\mathcal{T}_X + \mathcal{R}_A^\perp = \mathbb{S}^n$.

В силу того, что размерности $\mathbb{S}^n$ и $\mathcal{R}_A$ равны, соответственно, $k_\Delta(n)$ и m, а для размерности $\mathcal{T}_V$ имеем $\dim\mathcal{T}_V = k_\Delta(n) - k_\Delta(n-r)$, то точка $V \in \mathcal{F}_{D,V}$ ранга r будет невырожденной в том и только том случае, когда

$$k_\Delta(n) - k_\Delta(n-r) + m \ge k_\Delta(n),$$

откуда следует неравенство

$$k_\Delta(n-r) \le m. \tag{37}$$

Из аналогичных соображений получаем, что точка X ранга $n-r$ будет невырожденной в том и только том случае, когда

$$m \le k_\Delta(n) - k_\Delta(r). \tag{38}$$

Пусть Q_1 и Q_2 – подматрицы матрицы Q в (36), состоящие, соответственно, из первых r и последующих $n-r$ столбцов. Имеет место следующее

Утверждение 2 (см. [19]). *Точка $V \in \mathcal{F}_{D,V}$ будет невырожденной тогда и только тогда, когда матрицы $B_i = Q_2^{\mathrm{T}} A_i Q_2$, $1 \le i \le m$, порождают пространство $\mathbb{S}^{n-r}$. Соответственно, точка $X \in \mathcal{F}_P$ будет невырожденной тогда и только тогда, когда матрицы*

$$B_i = \begin{bmatrix} 0 & Q_1^{\mathrm{T}} A_i Q_2 \\ Q_2^{\mathrm{T}} A_i Q_1 & Q_2^{\mathrm{T}} A_i Q_2 \end{bmatrix}, \quad 1 \le i \le m,$$

линейно независимы.

Для того чтобы правые части систем (22) или (31) были корректно определены, требуется невырожденность матрицы $\Phi(V)$, имеющей вид (18). Покажем, что данная матрица является неособой, если точка $V \in \mathcal{F}_{D,V}$ является невырожденной.

Из свойств кронекеровских произведений матриц вытекает следующее утверждение относительно кронекеровской суммы $V^\otimes$ матрицы V.

Утверждение 3. *Пусть для матрицы $V \in \mathbb{S}^n$ имеет место разложение $V = QD(\theta)Q^{\mathrm{T}}$, где Q – ортогональная матрица, θ есть n-мерный вектор, составленный из собственных значений матрицы V. Тогда*

$$V^\otimes = (Q\otimes Q)D(\theta^\otimes)(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}}). \tag{39}$$

Здесь $\theta^\otimes$ – диагональ матрицы $D^\otimes(\theta) = [D(\theta)\otimes I_n + I_n \otimes D(\theta)]/2$.

Используя (39), матрицу $\Phi(V)$ представим в виде

$$\Phi(V) = \mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vecs}} + \mathcal{L}_n(Q\otimes Q)D(\theta^\otimes)(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})\mathcal{D}_n, \tag{40}$$

или, с учетом равенства (7),

$$\Phi(V) = \mathcal{A}^{\mathrm{T}}_{\mathrm{vech}}\mathcal{A}_{\mathrm{vecs}} + \mathcal{L}_n(Q\otimes Q)D(\theta^{\otimes})\mathcal{D}_n\mathcal{L}_n(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})\mathcal{D}_n. \tag{41}$$

Уточним вид матрицы $D(\theta^{\otimes})\mathcal{D}_n\mathcal{L}_n$, входящей во второе слагаемое в выражении (41) для матрицы $\Phi(V)$.

Лемма 1. *Справедливо равенство*

$$D(\theta^{\otimes})\mathcal{D}_n\mathcal{L}_n = \mathcal{D}_n\mathcal{L}_nD(\theta^{\otimes})\mathcal{D}_n\mathcal{L}_n. \tag{42}$$

Доказательство. Матрица $\mathcal{L}_n$ имеет размер $k_\Delta(n)\times n^2$, причем все ее столбцы с номерами, равными старшим внедиагональным номерам из J, являются нулевыми. В столбцах с номерами, равными младшим внедиагональным номерам из J, имеется единственный ненулевой единичный элемент, который расположен в строке с тем же номером, что и номер столбца. В матрице $\mathcal{D}_n$ размера $n^2\times k_\Delta(n)$ во всех столбцах с номерами, равными внедиагональным номерам из J_Δ, имеются только два отличных от нуля единичных элемента, которые расположены в строках с симметричными младшими и старшими внедиагональными номерами из J, причем номер первой строки равняется номеру столбца. Поэтому в результате умножения матрицы $\mathcal{D}_n$ на $\mathcal{L}_n$ получается квадратная матрица порядка n^2, в которой во всех столбцах расположены векторы, соответствующие прямой сумме столбцов симметричной матрицы, быть может нулевой. Умножение матрицы $\mathcal{D}_n\mathcal{L}_n$ слева на диагональную матрицу $D(\theta^{\otimes})$, в которой вектор $\theta^{\otimes}$ является прямой суммой столбцов симметричной матрицы, не меняет этого свойства столбцов результирующей матрицы. Так как $\mathcal{D}_n\mathcal{L}_n\operatorname{vec}M = \operatorname{vec}M$ для симметричных матриц M, то отсюда приходим к равенству (42). Лемма доказана.

Учтем утверждение леммы 1 и перейдем к нахождению более удобного для дальнейшего исследования виду матрицы $\Phi(V)$. Поскольку матрица $Q\otimes Q$ ортогональна, то

$$\mathcal{A}^{\mathrm{T}}_{\mathrm{vec}} = (Q\times Q)(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})\mathcal{A}^{\mathrm{T}}_{\mathrm{vec}} = (Q\times Q)(\mathcal{A}^{Q}_{\mathrm{vec}})^{\mathrm{T}}. \tag{43}$$

Здесь и ниже через $\mathcal{A}^{Q}_{\mathrm{vec}}$ обозначена $m\times n^2$-матрица, строками которой являются векторы $\operatorname{vec}(Q^{\mathrm{T}}A_iQ)$, $1\le i\le m$. Так как все матрицы $Q^{\mathrm{T}}A_iQ$ симметричные, то в матрице $(\mathcal{A}^{Q}_{\mathrm{vec}})^{\mathrm{T}}$ имеются одинаковые строки, а именно строки с младшими и старшими симметричными внедиагональными номерами из J. Следовательно, всем столбцам матрицы $(\mathcal{A}^{Q}_{\mathrm{vec}})^{\mathrm{T}}$ соответствуют симметричные матрицы. Поэтому наряду с равенством (43) справедливо

$$\mathcal{A}^{\mathrm{T}}_{\mathrm{vec}} = (Q\times Q)(\mathcal{A}^{Q}_{\mathrm{vec}})^{\mathrm{T}} = (Q\times Q)\mathcal{D}_n\mathcal{L}_n(\mathcal{A}^{Q}_{\mathrm{vec}})^{\mathrm{T}}. \tag{44}$$

Обозначим через $\mathcal{A}^{Q}_{\mathrm{vech}}$ матрицу размера $m\times k_\Delta(n)$, строками которой являются, соответственно, векторы $\operatorname{vech}(Q^{\mathrm{T}}A_iQ)$, $1\le i\le m$. Тогда $\mathcal{L}_n(\mathcal{A}^{Q}_{\mathrm{vec}})^{\mathrm{T}} = (\mathcal{A}^{Q}_{\mathrm{vech}})^{\mathrm{T}}$ и равенство (44) теперь может быть записано как $\mathcal{A}^{\mathrm{T}}_{\mathrm{vec}} = (Q\times Q)\mathcal{D}_n(\mathcal{A}^{Q}_{\mathrm{vech}})^{\mathrm{T}}$, откуда

$$\mathcal{A}^{\mathrm{T}}_{\mathrm{vech}} = \mathcal{L}_n\mathcal{A}^{\mathrm{T}}_{\mathrm{vec}} = \mathcal{L}_n(Q\times Q)\mathcal{D}_n(\mathcal{A}^{Q}_{\mathrm{vech}})^{\mathrm{T}}. \tag{45}$$

Далее, транспонируя равенство (43), получаем $\mathcal{A}_{\mathrm{vec}} = \mathcal{A}^{Q}_{\mathrm{vec}}(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})$. Поэтому $\mathcal{A}_{\mathrm{vecs}} = \mathcal{A}_{\mathrm{vec}}\mathcal{D}_n =$ $= \mathcal{A}^{Q}_{\mathrm{vec}}(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})\mathcal{D}_n$. Но согласно формуле (7) имеем $(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})\mathcal{D}_n = \mathcal{D}_n\mathcal{L}_n(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})\mathcal{D}_n$. Используя данное соотношение, приходим к равенству $\mathcal{A}_{\mathrm{vecs}} = \mathcal{A}^{Q}_{\mathrm{vec}}\mathcal{D}_n\mathcal{L}_n(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})\mathcal{D}_n$, которое переписываем в виде

$$\mathcal{A}_{\mathrm{vecs}} = \mathcal{A}^{Q}_{\mathrm{vecs}}\mathcal{L}_n(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})\mathcal{D}_n, \tag{46}$$

где через $\mathcal{A}^{Q}_{\mathrm{vecs}}$ обозначена матрица, строками которой являются векторы $\operatorname{vecs}(Q^{\mathrm{T}}A_iQ)$, $1\le i\le m$.

Введем квадратную матрицу $\mathscr{H}$, положив $\mathscr{H} = \mathscr{L}_n(Q \otimes Q)\mathscr{D}_n$. Ее порядок равен $k_\Delta(n)$. Согласно формуле (8) и ортогональности матрицы Q имеем

$$\mathscr{H}^{-1} = (\mathscr{L}_n(Q \otimes Q)\mathscr{D}_n)^{-1} = \mathscr{L}_n(Q^{\mathrm{T}} \otimes Q^{\mathrm{T}})\mathscr{D}_n.$$

Таким образом, на основании (45) и (46) для матрицы (40) справедливо представление

$$\Phi(V) = \mathscr{H}\mathscr{W}(\theta^{\otimes})\mathscr{H}^{-1}, \tag{47}$$

где $\mathscr{W}(\theta^{\otimes}) = (\mathscr{A}^{Q}_{\mathrm{vech}})^{\mathrm{T}}\mathscr{A}^{Q}_{\mathrm{vecs}} + \mathscr{L}_n D(\theta^{\otimes})\mathscr{D}_n$. Матрица $\mathscr{W}(\theta^{\otimes})$ также имеет порядок $k_\Delta(n)$.

Пусть $\tilde{\theta}^{\otimes}$ есть $k_\Delta(n)$-мерный вектор, определяемый согласно формуле $\tilde{\theta}^{\otimes} = \mathscr{L}_n\theta^{\otimes}$. Как несложно проверить, $D(\tilde{\theta}^{\otimes}) = \mathscr{L}_n D(\theta^{\otimes})\mathscr{D}_n$ и, следовательно, матрица $\mathscr{W}(\theta^{\otimes})$ фактически зависит от вектора $\tilde{\theta}^{\otimes}$ и может быть записана в виде

$$\mathscr{W}(\tilde{\theta}^{\otimes}) = (\mathscr{A}^{Q}_{\mathrm{vech}})^{\mathrm{T}}\mathscr{A}^{Q}_{\mathrm{vecs}} + D(\tilde{\theta}^{\otimes}). \tag{48}$$

Утверждение 4. *В невырожденной точке $V \in \mathscr{F}_{D,\,V}$ матрица $\Phi(V)$ является неособой.*

Доказательство. В силу (47) достаточно показать, что неособой является матрица $\mathscr{W}(\tilde{\theta}^{\otimes})$, имеющая вид (48).

Обозначим через D_2 диагональную матрицу порядка $k_\Delta(n)$. В ней компоненты вектора, стоящего на диагонали, равны единице, если их парный номер из J_Δ является диагональным. В противном случае данная компонента равна 2. Матрица D_2 положительно-определенная. Пусть $D_2^{1/2}$ – квадратный корень из D_2. Имеем, очевидно, $D(\tilde{\theta}^{\otimes}) = D_2^{-1/2}D(\tilde{\theta}^{\otimes})D_2^{1/2}$. Кроме того, $\mathscr{A}^{Q}_{\mathrm{vecs}} = \mathscr{A}^{Q}_{\mathrm{vech}}D_2$.

Положим для сокращения записи $\mathscr{A}^{Q}_{\mathrm{vec2}} = \mathscr{A}^{Q}_{\mathrm{vech}}D_2^{1/2}$. Тогда матрицу $\mathscr{W}(\tilde{\theta}^{\otimes})$ можно представить в виде

$$\mathscr{W}(\tilde{\theta}^{\otimes}) = D_2^{-1/2}\mathscr{W}(\tilde{\theta}^{\otimes})D_2^{1/2}, \tag{49}$$

где

$$\mathscr{W}_1(\tilde{\theta}^{\otimes}) = (\mathscr{A}^{Q}_{\mathrm{vec2}})^{\mathrm{T}}\mathscr{A}^{Q}_{\mathrm{vec2}} + D(\tilde{\theta}^{\otimes}).$$

Матрица $\mathscr{W}_1(\tilde{\theta}^{\otimes})$ является симметричной положительно-полуопределенной. При этом матрица $\mathscr{W}(\tilde{\theta}^{\otimes})$ будет неособой в том и только том случае, когда $\mathscr{W}_1(\tilde{\theta}^{\otimes})$ – положительно-определенная матрица.

Для того чтобы проверить положительную определенность матрицы $\mathscr{W}_1(\tilde{\theta}^{\otimes})$, покажем, что линейная однородная система уравнений

$$\mathscr{W}_1(\tilde{\theta}^{\otimes})x = 0_{k_\Delta(n)} \tag{50}$$

имеет только тривиальное решение $x = 0_{k_\Delta(n)}$.

Действительно, после умножения левой и правой частей соотношений (50) на x^{T} получаем

$$\langle x, (\mathscr{A}^{Q}_{\mathrm{vec2}})^{\mathrm{T}}\mathscr{A}^{Q}_{\mathrm{vec2}}x\rangle + \langle x, D(\tilde{\theta}^{\otimes})x\rangle = 0. \tag{51}$$

Так как обе матрицы $(\mathscr{A}^{Q}_{\mathrm{vec2}})^{\mathrm{T}}\mathscr{A}^{Q}_{\mathrm{vec2}}$ и $D(\tilde{\theta}^{\otimes})$ положительно-полуопределенные, то равенство (51) имеет место тогда и только тогда, когда

$$\langle x, (\mathscr{A}^{Q}_{\mathrm{vec2}})^{\mathrm{T}}\mathscr{A}^{Q}_{\mathrm{vec2}}x\rangle = 0, \quad \langle x, D(\tilde{\theta}^{\otimes})x\rangle = 0. \tag{52}$$

Если все собственные значения у матрицы V строго положительны, то $\tilde{\theta}^{\otimes} > 0_{k_\Delta(n)}$ и из второго равенства (52) сразу следует, что $x = 0_{k_\Delta(n)}$.

Далее предполагаем, что матрица V такова, что $\operatorname{rank} V = r < n$. Тогда среди собственных значений матрицы V имеются нулевые. Пусть, для определенности, для V имеет место разложение (36), в котором $\theta^i > 0$, $1 \le i \le r$, и $\theta^j = 0$, $r < j \le n$. Положим для сокращения записи $n_1 = k_\Delta(n) - k_\Delta(n-r)$, $n_2 = k_\Delta(n-r)$. Из невырожденности точки $V \in \mathcal{F}_{D,V}$ согласно (37) вытекает неравенство $n_2 \le m$.

При сделанном предположении относительно собственных значений матрицы V у вектора $\tilde{\theta}^{\otimes}$ первые n_1 компонент строго положительны, а остальные компоненты в количестве n_2 штук нулевые. В этом случае для вектора $\tilde{\theta}^{\otimes}$ справедливо разбиение

$$\tilde{\theta}^{\otimes} = \begin{bmatrix} (\tilde{\theta}^{\otimes})^N \\ (\tilde{\theta}^{\otimes})^B \end{bmatrix}, \tag{53}$$

где $(\tilde{\theta}^{\otimes})^N > 0_{n_1}$, $(\tilde{\theta}^{\otimes})^B = 0_{n_2}$.

Матрицу $\mathcal{A}^Q_{\mathrm{vec2}}$ разобьем также на две подматрицы в соответствии с разбиением (53):

$$\mathcal{A}^Q_{\mathrm{vec2}} = [(\mathcal{A}^Q_{\mathrm{vec2}})^N \,|\, (\mathcal{A}^Q_{\mathrm{vec2}})^B].$$

Здесь первая матрица $(\mathcal{A}^Q_{\mathrm{vec2}})^N$ имеет размер $m \times n_1$, вторая матрица $(\mathcal{A}^Q_{\mathrm{vec2}})^B$ – размер $m \times n_2$.

Если опять с целью сокращения записи положить $N = (\mathcal{A}^Q_{\mathrm{vec2}})^N$, $B = (\mathcal{A}^Q_{\mathrm{vec2}})^B$, $\tilde{\Theta}_N = D((\tilde{\theta}^{\otimes})^N)$, то матрица $\mathcal{W}_1(\tilde{\theta}^{\otimes})$ в этих обозначениях принимает вид

$$\mathcal{W}_1(\tilde{\theta}^{\otimes}) = \begin{bmatrix} N^{\mathrm{T}}N + \tilde{\Theta}_N & N^{\mathrm{T}}B \\ B^{\mathrm{T}}N & B^{\mathrm{T}}B \end{bmatrix}. \tag{54}$$

Разобьем также вектор x, являющийся решением уравнения (50), на две части $x = [x^N, x^B]$ в соответствии с разбиением вектора $\tilde{\theta}^{\otimes}$. На основании второго равенства (52) заключаем, что $x^N = 0_{n_1}$. Поэтому первое равенство (52) сводится к $\langle x^B, B^{\mathrm{T}}Bx^B\rangle = 0$. Отсюда следует, что $Bx^B = 0_{n_2}$. Но в невырожденных точках $V \in \mathcal{F}_{D,V}$ согласно утверждению 2 матрица B имеет полный ранг по столбцам, равный n_2. Таким образом, $x^B = 0$. Мы пришли к выводу, что уравнение (50) имеет только тривиальное решение $x = 0_{k_\Delta(n)}$, что означает невырожденность матрицы $\mathcal{W}_1(\tilde{\theta}^{\otimes})$. Утверждение доказано.

Ниже предполагается, что двойственная задача (2) является невырожденной, т.е. все точки $V \in \mathcal{F}_{D,V}$ невырожденные. Тогда, как следует из утверждения 4, решение системы (17) существует и единственно для любых $V \in \mathcal{F}_{D,V}$. Соответственно, решение системы (27) также существует и единственно для любых $[u, V] \in \mathcal{F}_D$. В силу непрерывности, эти решения будут существовать и в некоторых окрестностях множество $\mathcal{F}_{D,V}$ и $\mathcal{F}_D$. В этом случае правые части систем (22) и (31) полностью определены в некоторой области, содержащей допустимое множество $\mathcal{F}_D$.

Пусть точка $u_* \in \mathcal{F}_{D,u}$ такова, что u_* вместе с $V_* = V(u_*)$ есть решение двойственной задачи (2). Пусть, кроме того, X_* – соответствующее решение прямой задачи (1). Считаем также, что для X_* и V_* имеют место представления (5), причем, для определенности, у матрицы V_* первые собственные значения в количестве r штук строго положительные, а остальные значения нулевые. У матрицы X_*, напротив, ненулевыми могут быть только последние $n - r$ собственных значения. Поэтому у вектора $\operatorname{vech} D(\eta_*)$ ненулевыми могут оказаться лишь элементы, входящие в число последних $k_\Delta(n-r)$ элементов, расположенные на диагональных парных номерах из J_Δ. Но тогда с учетом утверждения леммы 1 получаем, что

$$D(\theta_*^{\otimes})\mathcal{D}_n\mathcal{L}_n(Q^{\mathrm{T}}\otimes Q^{\mathrm{T}})\mathcal{D}_n\operatorname{vech}X_* = D(\theta_*^{\otimes})\mathcal{D}_n\operatorname{vech}D(\eta_*) = \mathcal{D}_nD(\tilde{\theta}_*^{\otimes})\operatorname{vech}D(\eta_*) = 0.$$

Отсюда на основании (41) приходим к равенству

$$\Phi(V_*)\operatorname{vech}X_* = \mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vecs}}\operatorname{vech}X_* = \mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}b.$$

Таким образом, $\operatorname{vech}X_*$ удовлетворяет уравнению (27) при $[u, V] = [u_*, V_*]$. Поэтому в силу единственности решения этого уравнения заключаем, что $X(u_*, V_*) = X_*$.

Так как $V_* * X_* = 0_{nn}$ и $\mathcal{A}\bullet X_* = b$, то пара $[u_*, V_*]$ является стационарной точкой итерационного процесса (34). Соответственно, вектор u_* является стационарной точкой процесса (21).

6. АЛГОРИТМИЧЕСКИЕ ОПЕРАТОРЫ И ИХ ПРЕДСТАВЛЕНИЯ

Пусть $\mathcal{G}_1(u)$ обозначает алгоритмический оператор итерационного процесса (21). Он задается формулой

$$\mathcal{G}_1(u) = [I_m - \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(V(u))\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}]b,$$

а сам итерационный процесс (21) с его помощью записывается в виде

$$u_{k+1} = u_k + \alpha_k\mathcal{G}_1(u_k).$$

В тех точках $u \in \mathcal{F}_{D,u}$, в которых $V(u) \succ 0$, для оператора $\mathcal{G}_1(u)$, как это следует из (25), справедливо представление

$$\mathcal{G}_1(u) = [I_m - \mathcal{A}_{\mathrm{vecs}}\mathcal{L}_n(V^{\otimes}(u))^{-1}\mathcal{D}_n\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}]b.$$

Уточним теперь вид данного оператора в точках множества $\mathcal{F}_{D,u}$, принадлежащих его относительной границе. В них выполняется только $V(u) \succeq 0$. Для определенности опять считаем, что матрица V имеет ранг r и для нее имеет место разложение (36).

Так как в невырожденной точке $V \in \mathcal{F}_{D,V}$ матрица $\mathcal{W}_1(\tilde{\theta}^{\otimes})$ неособая, то у нее существует обратная матрица. Найдем ее. Воспользуемся для этой цели блочным представлением (54) матрицы $\mathcal{W}_1(\tilde{\theta}^{\otimes})$. Опять же из-за невырожденности точки V, согласно утверждению 2, правая нижняя подматрица $B^{\mathrm{T}}B$ матрицы $\mathcal{W}_1(\tilde{\theta}^{\otimes})$ неособая. Обозначим через $\mathcal{K}$ неособую матрицу

$$\mathcal{K} = N^{\mathrm{T}}N + \tilde{\Theta}_N - N^{\mathrm{T}}B(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}N.$$

Тогда с помощью формулы Фробениуса получаем

$$\mathcal{W}_1^{-1}(\tilde{\theta}_*) = \begin{bmatrix}\mathcal{W}_{11}^{-1} & \mathcal{W}_{12}^{-1}\\ \mathcal{W}_{21}^{-1} & \mathcal{W}_{23}^{-1}\end{bmatrix}, \tag{55}$$

где $\mathcal{W}_{11}^{-1} = \mathcal{K}^{-1}$ и

$$\mathcal{W}_{21}^{-1} = -(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}N\mathcal{K}^{-1}, \quad \mathcal{W}_{12}^{-1} = -\mathcal{K}^{-1}N^{\mathrm{T}}B(B^{\mathrm{T}}B)^{-1}, \tag{56}$$

$$\mathcal{W}_{22}^{-1} = (B^{\mathrm{T}}B)^{-1} + (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}N\mathcal{K}^{-1}N^{\mathrm{T}}B(B^{\mathrm{T}}B)^{-1}. \tag{57}$$

Пусть $\mathcal{L}_B$ обозначает пространство столбцов матрицы B, и пусть $\mathcal{L}_B^{\perp}$ – его ортогональное дополнение. Тогда матрицу $\mathcal{K}$ можно записать в виде

$$\mathcal{K} = \tilde{\Theta}_N + N^{\mathrm{T}}[I_m - \mathcal{P}]N = \tilde{\Theta}_N + N^{\mathrm{T}}\mathcal{P}_{\perp}N,$$

где $\mathcal{P} = B(B^{\text{т}}B)^{-1}B^{\text{т}}$ и $\mathcal{P}_{\perp} = I_m - \mathcal{P}$ — матрицы ортогонального проектирования на, соответственно, подпространства $\mathcal{L}_B$ и $\mathcal{L}_B^{\perp}$. Матрица $\mathcal{P}_{\perp}$ является симметрической и идемпотентной, т.е. $\mathcal{P}_{\perp} = \mathcal{P}_{\perp}^2$. Поэтому $\mathcal{K} = \tilde{\Theta}_N + N^{\text{т}}\mathcal{P}_{\perp}^2 N$. Чтобы вычислить обратную матрицу $\mathcal{K}^{-1}$, воспользуемся формулой Шермана–Моррисона–Вудберри. Так как $\tilde{\Theta}_N$ — положительно-определенная матрица, то согласно этой формуле имеем

$$\mathcal{K}^{-1} = \tilde{\Theta}_N^{-1} - \tilde{\Theta}_N^{-1}N^{\text{т}}\mathcal{P}_{\perp}(I + \mathcal{P}_{\perp}N\tilde{\Theta}_N^{-1}N^{\text{т}}\mathcal{P}_{\perp})^{-1}\mathcal{P}_{\perp}N\tilde{\Theta}_N^{-1} \tag{58}$$

и, далее,

$$D_2^{1/2}\mathcal{H}^{-1}\mathcal{A}_{\text{vech}}^{\text{т}} = D_2^{1/2}\mathcal{L}_n(Q^{\text{т}} \times Q^{\text{т}})\mathcal{D}_n\mathcal{A}_{\text{vech}}^{\text{т}} = (\mathcal{A}_{\text{vec2}}^{Q})^{\text{т}}.$$

Кроме того,

$$\mathcal{A}_{\text{vecs}}\mathcal{H}D_2^{-1/2} = \mathcal{A}_{\text{vecs}}\mathcal{L}_n(Q \times Q)\mathcal{D}_n D^{-1/2} = \mathcal{A}_{\text{vec2}}^{Q}.$$

Отсюда следует, что

$$\mathcal{A}_{\text{vecs}}\Phi^{-1}(V)\mathcal{A}_{\text{vech}}^{\text{т}} = \mathcal{A}_{\text{vec2}}^{Q}\mathcal{W}_1^{-1}(\tilde{\theta})(\mathcal{A}_{\text{vec2}}^{Q})^{\text{т}}.$$

Вычислим сначала $M_1 = \mathcal{W}_{11}^{-1}N^{\text{т}} + \mathcal{W}_{12}^{-1}B^{\text{т}}$ и $M_2 = \mathcal{W}_{21}^{-1}N^{\text{т}} + \mathcal{W}_{22}^{-1}B^{\text{т}}$. Проведя последовательно выкладки, получим

$$M_1 = \mathcal{K}^{-1}N^{\text{т}}[I - B(B^{\text{т}}B)^{-1}B^{\text{т}}] = \mathcal{K}^{-1}N^{\text{т}}\mathcal{P}_{\perp},$$

$$M_2 = -(B^{\text{т}}B)^{-1}B^{\text{т}}N\mathcal{K}^{-1}N^{\text{т}} + (B^{\text{т}}B)^{-1}B^{\text{т}} + (B^{\text{т}}B)^{-1}B^{\text{т}}N\mathcal{K}^{-1}N^{\text{т}}B(B^{\text{т}}B)^{-1}B^{\text{т}} =$$
$$= (B^{\text{т}}B)^{-1}B^{\text{т}}[I - N\mathcal{K}^{-1}N^{\text{т}} + N\mathcal{K}^{-1}N^{\text{т}}B(B^{\text{т}}B)^{-1}B^{\text{т}}] = (B^{\text{т}}B)^{-1}B^{\text{т}}[I - N\mathcal{K}^{-1}N^{\text{т}}\mathcal{P}_{\perp}].$$

Таким образом,

$$\mathcal{W}_1^{-1}(\tilde{\theta})(\mathcal{A}_{\text{vec2}}^{Q})^{\text{т}} = \begin{bmatrix} \mathcal{K}^{-1}N^{\text{т}}\mathcal{P}_{\perp} \\ (B^{\text{т}}B)^{-1}B^{\text{т}}[I - N\mathcal{K}^{-1}N^{\text{т}}\mathcal{P}_{\perp}] \end{bmatrix}.$$

Далее находим

$$\mathcal{A}_{\text{vec2}}^{Q}\mathcal{W}_1^{-1}(\tilde{\theta})(\mathcal{A}_{\text{vec2}}^{Q})^{\text{т}} = N\mathcal{K}^{-1}N^{\text{т}}\mathcal{P}_{\perp} + B(B^{\text{т}}B)^{-1}B^{\text{т}}[I - N\mathcal{K}^{-1}N^{\text{т}}\mathcal{P}_{\perp}] =$$
$$= N\mathcal{K}^{-1}N^{\text{т}}\mathcal{P}_{\perp} + \mathcal{P}[I - N\mathcal{K}^{-1}N^{\text{т}}\mathcal{P}_{\perp}] = \mathcal{P} + \mathcal{P}_{\perp}N\mathcal{K}^{-1}N^{\text{т}}\mathcal{P}_{\perp}.$$

Введем обозначение $\mathcal{Y} = \mathcal{P}_{\perp}N\tilde{\Theta}_N^{-1}N^{\text{т}}\mathcal{P}_{\perp}$. Тогда после подстановки выражения (58) для матрицы $\mathcal{K}^{-1}$ получаем

$$\mathcal{P}_{\perp}NK^{-1}N^{\text{т}}\mathcal{P}_{\perp} = \mathcal{P}_{\perp}N[\tilde{\Theta}_N^{-1} - \tilde{\Theta}_N^{-1}N^{\text{т}}\mathcal{P}_{\perp}(I + \mathcal{Y})^{-1}\mathcal{P}_{\perp}N\tilde{\Theta}_N^{-1}]N^{\text{т}}\mathcal{P}_{\perp} =$$
$$= \mathcal{Y} - \mathcal{Y}(I + \mathcal{Y})^{-1}\mathcal{Y} = I - (I + \mathcal{Y})^{-1}.$$

Отсюда следует, что

$$I - \mathcal{A}_{\text{vecs}}\Phi^{-1}(V)\mathcal{A}_{\text{vech}}^{\text{т}} = \mathcal{P}_{\perp}(I + \mathcal{Y})^{-1}\mathcal{P}_{\perp}.$$

Окончательно имеем

$$\mathcal{G}_1(u) = \mathcal{P}_{\perp}[I_m + \mathcal{P}_{\perp}(\mathcal{A}_{\text{vec2}}^{Q})^{N}\tilde{\Theta}_N^{-1}((\mathcal{A}_{\text{vec2}}^{Q})^{N})^{\text{т}}\mathcal{P}_{\perp}]^{-1}\mathcal{P}_{\perp}b,$$

где

$$\mathcal{P}_{\perp} = I_m - (\mathcal{A}_{\text{vec2}}^{Q})^{B}[((\mathcal{A}_{\text{vec2}}^{Q})^{B})^{\text{т}}(\mathcal{A}_{\text{vec2}}^{Q})^{B}]^{-1}((\mathcal{A}_{\text{vec2}}^{Q})^{B})^{\text{т}}.$$

Отметим, что зависимость алгоритмического отображения $\mathcal{G}_1(u)$ от двойственной переменной u здесь опосредованная через связь $V = V(u)$ и через разложение (36).

Если обратиться к итерационному процессу (31), то для него алгоритмический оператор имеет вид

$$\mathcal{G}_2(u, V) = \begin{bmatrix} b - \mathcal{A}_{\mathrm{vecs}}\mathrm{vech}X(u, V) \\ -\mathrm{vech}(V * X(u, V)) \end{bmatrix}.$$

В точках $[u, V] \in \mathcal{F}_D$, так как тогда $V = V(u)$, имеем

$$\mathcal{G}_2(u, V) = \begin{bmatrix} \mathcal{G}_1(u) \\ -\mathrm{vech}(V * X(u, V)) \end{bmatrix}. \tag{59}$$

Объединим пару переменных u и $\mathrm{vech}\,V$ в единый $(m + k_\Delta(n))$-мерный вектор. Пусть $Z = [u, V]$. Тогда с помощью отображения (59) итерационный процесс (31) может быть записан в виде

$$Z_{k+1} = Z_k + \alpha_k \mathcal{G}_2(Z_k). \tag{60}$$

В решении двойственной задачи (2), т.е. в паре $Z_* = [u_*, \mathrm{vech}V_*]$, как было показано, имеем $\mathcal{G}_2(Z_*) = 0_{m + k_\Delta(n)}$.

Отметим также, что из равенства $\mathcal{A}_{\mathrm{vec}}\mathrm{vec}X_* = b$, где X_* – решение прямой задачи (1), следует, что

$$b = \mathcal{A}_{\mathrm{vec}}\mathrm{vec}X_* = \mathcal{A}_{\mathrm{vec}}(Q^{\mathrm{T}} \otimes Q^{\mathrm{T}})\mathrm{vec}D(\eta_*) = \mathcal{A}^Q_{\mathrm{vec}}\mathrm{vec}D(\eta_*) = \mathcal{A}^Q_{\mathrm{vec}2}\mathcal{L}_n\mathrm{vec}D(\eta_*).$$

Отсюда видно, что вектор b есть линейная комбинация последних $k_\Delta(n - r)$ столбцов матрицы $\mathcal{A}^Q_{\mathrm{vec}2}$ (точнее, $n - r$ из них). Поэтому он принадлежит подпространству $\mathcal{L}_B$. Отсюда делаем вывод, что $\mathcal{P}_\perp b = 0_m$ и, стало быть, $\mathcal{G}_1(u_*) = 0_m$.

7. МАТРИЦА ЯКОБИ

Обозначим через $F(u, V)$ матричную функцию $F(u, V) = V * X(u, V)$ и введем отображение

$$\mathcal{Q}(u, V) = \begin{bmatrix} \mathcal{Q}^1(u, V) \\ \mathcal{Q}^2(u, V) \end{bmatrix},$$

где

$$\mathcal{Q}^1(u, V) = \mathcal{A}_{\mathrm{vecs}}\mathrm{vech}X(u, V), \quad \mathcal{Q}^2(u, V) = \mathrm{vech}F(u, V).$$

Ниже нам потребуются матрицы Якоби отображений $\mathcal{Q}^1(u, V)$ и $\mathcal{Q}^2(u, V)$ относительно переменных u и V, точнее, от u и $\mathrm{vech}\,V$. Чтобы их получить, надо знать, разумеется, матрицы Якоби X_u и X_V от матричной функции $X(u, V)$. Согласно определению из [16] имеем

$$X_u(u, V) = \frac{\partial \mathrm{vec}X(u, V)}{\partial u}, \quad X_V(u, V) = \frac{\partial \mathrm{vec}X(u, V)}{\partial \mathrm{vec}V}.$$

Рассмотрим также матрицы Якоби

$$X^\Delta_u(u, V) = \mathcal{L}_n X_u(u, V) = \frac{\partial \mathrm{vech}X(u, V)}{\partial u},$$

$$X^\Delta_V(u, V) = \mathcal{L}_n X_V(u, V)\mathcal{D}_n = \frac{\partial \mathrm{vech}X(u, V)}{\partial \mathrm{vech}V},$$

учитывающие симметричность матриц $X(u, V)$ и V.

Имеем для отображения $\mathfrak{Q}^1(u, V)$

$$\frac{\partial \mathfrak{Q}^1(u, V)}{\partial u} = \mathcal{A}_{\mathrm{vec}} \frac{\partial X(u, V)}{\partial u} = \mathcal{A}_{\mathrm{vecs}} X_u^{\Delta}(u, V), \tag{61}$$

$$\frac{\partial \mathfrak{Q}^1(u, V)}{\partial \operatorname{vech} V} = \mathcal{A}_{\mathrm{vec}} \frac{\partial \operatorname{vec} X(u, V)}{\partial \operatorname{vech} V} = \mathcal{A}_{\mathrm{vecs}} X_V^{\Delta}(u, V). \tag{62}$$

Соответственно, для отображения $\mathfrak{Q}^2(u, V)$

$$\frac{\partial \mathfrak{Q}^2(u, V)}{\partial u} = \frac{\partial \operatorname{vech} F(u, V)}{\partial u} = \mathcal{L}_n F_u(u, V), \tag{63}$$

$$\frac{\partial \mathfrak{Q}^2(u, V)}{\partial \operatorname{vech} V} = \frac{\partial \operatorname{vech} F(u, V)}{\partial \operatorname{vech} V} = \mathcal{L}_n F_V(u, V) \mathfrak{D}_n, \tag{64}$$

где

$$F_u(u, V) = \frac{\partial \operatorname{vec} F(u, V)}{\partial u}, \quad F_V(u, V) = \frac{\partial \operatorname{vec} F(u, V)}{\partial \operatorname{vec} V}$$

суть матрицы Якоби матричной функции $F(u, V)$ относительно переменных u и V соответственно.

Из вида функции $F(u, V)$ следует, что

$$F_u(u, V) = \frac{1}{2} \frac{\partial \operatorname{vec}[X(u, V)V + VX(u, V)]}{\partial u},$$

или, если воспользоваться формулой (13),

$$\frac{\partial \operatorname{vec}(X(u, V)V)}{\partial u} = \frac{\partial [(V \otimes I_n) \operatorname{vec} X(u, V)]}{\partial u}, \tag{65}$$

$$\frac{\partial \operatorname{vec}(VX(u, V))}{\partial u} = \frac{\partial [(I_n \otimes V) \operatorname{vec} X(u, V)]}{\partial u}. \tag{66}$$

Кроме того, так как матрица $X(u, V)$ симметричная, то $X_u(u, V) = \mathfrak{D}_n X_u^{\Delta}(u, V)$. Отсюда и из (65), (66), используя также (63), получаем

$$\frac{\partial \mathfrak{Q}^2(u, V)}{\partial u} = \mathcal{L}_n V^{\otimes} \mathfrak{D}_n X_u^{\Delta}(u, V). \tag{67}$$

Чтобы вычислить производную $F_V(u, V)$, воспользуемся следующим утверждением.

Лемма 2. *Матрица Якоби $F_V(u, V)$ для матричной функции $F(u, V)$ имеет вид*

$$F_V(u, V) = X^{\otimes} + V^{\otimes} X_V(u, V). \tag{68}$$

Доказательство. Возьмем матричную функцию $H(u, V) = VX(u, V)$, ее дифференциал относительно переменной V имеет вид

$$dH(u, V) = dVX(u, V) + VdX(u, V).$$

В векторной форме это соотношение запишется в виде

$$d\operatorname{vec}(H(u, V)) = \operatorname{vec}(dVX(u, V)) + \operatorname{vec}(VdX(u, V)).$$

Тогда на основании формулы (13) получаем

$$d(\operatorname{vec} H(u, V)) = (X^{\mathrm{T}}(u, V) \otimes I_n) d\operatorname{vec} V + (I_n \otimes V) \operatorname{vec}(dX(u, V)). \tag{69}$$

Но по определению первого дифференциала имеем

$$\operatorname{vec}(dX(u, V)) = X_V(u, V) d\operatorname{vec} V.$$

Подставляя данное равенство в (69), приходим к соотношению

$$d(\mathrm{vec}H) = [(X^{\mathrm{T}}(u, V) \otimes I_n) + (I_n \otimes V)X_V(u, V)]d\mathrm{vec}V.$$

Таким образом, с учетом симметричности матрицы $X(u, V)$, имеем

$$H_V(u, V) = X(u, V) \otimes I_n + (I_n \otimes V)X_V(u, V). \tag{70}$$

Аналогичным образом показывается, что матрица Якоби для матричной функции $G(u, V) = X(u, V)V$ относительно переменной V имеет вид

$$G_V(u, V) = (V \otimes I_n)X_V(u, V) + I_n \otimes X(u, V). \tag{71}$$

На основании (70) и (71) приходим к (68). Лемма доказана.

Поскольку $X_V(u, V) = \mathcal{D}_n\mathcal{L}_nX_V(u, V)$, то, согласно лемме 2, имеем

$$\frac{\partial\mathcal{Q}^2(u, V)}{\partial\mathrm{vech}V} = \mathcal{L}_nX^{\otimes}\mathcal{D}_n + \mathcal{L}_nV^{\otimes}\mathcal{D}_nX_V^{\Delta}(u, V).$$

Введем следующие обозначения:

$$\tilde{X}^{\otimes}(u, V) = \mathcal{L}_nX^{\otimes}(u, V)\mathcal{D}_n, \quad \tilde{V}^{\otimes} = \mathcal{L}_nV^{\otimes}\mathcal{D}_n. \tag{72}$$

Обе матрицы $\tilde{X}^{\otimes}(u, V)$ и $\tilde{V}^{\otimes}$ являются квадратными порядка $k_{\Delta}(n)$.

После подстановки соответствующих выражений (61), (62), (67) и (64) для частных производных, приходим к следующей матрице Якоби для отображения $\mathcal{Q}(u, V)$:

$$\mathcal{Q}_{[u,\, \mathrm{vech}V]}(u, V) = \begin{bmatrix} \mathcal{A}_{\mathrm{vecs}}X_u^{\Delta} & \mathcal{A}_{\mathrm{vecs}}X_V^{\Delta} \\ \tilde{V}^{\otimes}X_u^{\Delta} & \tilde{X}^{\otimes} + \tilde{V}^{\otimes}X_V^{\Delta} \end{bmatrix}. \tag{73}$$

Здесь $\tilde{X}^{\otimes} = \tilde{X}^{\otimes}(u, V)$, $X_u^{\Delta} = X_u^{\Delta}(u, V)$, $X_V^{\Delta} = X_V^{\Delta}(u, V)$.

Определим матрицы $X_u^{\Delta}(u, V)$ и $X_V^{\Delta}(u, V)$. С этой целью продифференцируем равенство (27) по u и $\mathrm{vech}V$ соответственно. Дифференцирование по u дает

$$(\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vecs}} + \mathcal{L}_nV^{\otimes}\mathcal{D}_n)X_u^{\Delta}(u, V) = \tau\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}. \tag{74}$$

Дифференцирование по $\mathrm{vech}V$ приводит к соотношению

$$\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vecs}}X_V^{\Delta}(u, V) + \frac{\partial(\mathcal{L}_nV^{\otimes}\mathcal{D}_n\mathrm{vech}X(u, V))}{\partial\mathrm{vech}V} = \tau I_{k_{\Delta}(n)}. \tag{75}$$

Рассмотрим отдельно второе слагаемое в левой части (75). Имеем

$$\frac{\partial(\mathcal{L}_nV^{\otimes}\mathcal{D}_n\mathrm{vech}X(u, V))}{\partial\mathrm{vech}V} = \mathcal{L}_n\frac{\partial(V^{\otimes}\mathrm{vec}X(u, V))}{\partial\mathrm{vec}V}\mathcal{D}_n = \mathcal{L}_nF_V(u, V)\mathcal{D}_n.$$

Поэтому

$$\frac{\partial(\mathcal{L}_nV^{\otimes}\mathcal{D}_n\mathrm{vech}X(u, V))}{\partial\mathrm{vech}V} = \mathcal{L}_nX^{\otimes}(u, V)\mathcal{D}_n + \mathcal{L}_nV^{\otimes}\mathcal{D}_nX_V^{\Delta}(u, V). \tag{76}$$

После подстановки выражения (76) в (75) приходим к равенству

$$\mathcal{L}_nX^{\otimes}(u, V)\mathcal{D}_n + [\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vecs}} + \mathcal{L}_nV^{\otimes}\mathcal{D}_n]X_V^{\Delta}(u, V) = \tau I_{k_{\Delta}(n)}. \tag{77}$$

Тогда с учетом введенных обозначений (72) равенства (74) и (77) можно переписать в виде

$$\Phi(V)X_u^{\Delta}(u, V) = \tau\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}},$$

$$\tilde{X}^{\otimes} + \Phi(V)X_V^{\Delta}(u, V) = \tau I_{k_{\Delta}(n)}.$$

Если матрица $\Phi(V)$ неособая, то отсюда следует, что

$$X_u^\Delta(u, V) = \tau\Phi^{-1}(V)\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}},$$

$$X_V^\Delta(u, V) = \Phi^{-1}(V)[\tau I_{k_\Delta}(n) - \tilde{X}^{\otimes}(u, V)].$$

Подставим найденные выражения в матрицу Якоби (73). Тогда данная матрица $\mathfrak{Q}_{[u,\,\mathrm{vech}\,V]}(u, V)$ преобразуется к виду

$$\begin{bmatrix} \tau\mathcal{A}_{\mathrm{vecs}}\Phi^{-1}\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} & \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(\tau I_{k_\Delta(n)} - \tilde{X}^{\otimes}) \\ \tau\tilde{V}^{\otimes}\Phi^{-1}\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} & (I_{k_\Delta(n)} - \tilde{V}^{\otimes}\Phi^{-1})\tilde{X}^{\otimes} + \tau\tilde{V}^{\otimes}\Phi^{-1} \end{bmatrix}, \tag{78}$$

где $\Phi^{-1} = \Phi^{-1}(V)$.

7.1. Локальная сходимость методов

Покажем, что методы (22), (31) обладают локальной сходимостью. На самом деле, достаточно доказать, что локальной сходимостью обладает более общий метод (31). Сходимость первого метода (22) будет следовать автоматически из сходимости (31).

Нам потребуется матрица (78) в решении $[u_*, V_*]$ двойственной задачи (2). Для этого следует знать матрицы $X_*^{\otimes}$ и $V_*^{\otimes}$, где $X_* = X(u_*, V_*)$ – решение прямой задачи (1). При этом предполагаем, что в оптимальных решениях задач (1) и (2) выполнено условие строгой дополнительности.

Поскольку матрицы X_* и V_* симметричные, обе матрицы $X_*^{\otimes}$ и $V_*^{\otimes}$ также являются симметричными. Более того, как несложно проверить (см., например, [13]), из-за того что матрицы X_* и V_* коммутируют, матрицы $X_*^{\otimes}$ и $V_*^{\otimes}$ также коммутируют. Поэтому найдется ортогональная матрица U порядка n^2 такая, что

$$X_*^{\otimes} = UD(\nu_*)U^{\mathrm{T}}, \quad V_*^{\otimes} = UD(\mu_*)U^{\mathrm{T}}, \tag{79}$$

где ν_* и μ_* суть n^2-мерные векторы, состоящие из собственных значений, соответственно, матриц $X_*^{\otimes}$ и $V_*^{\otimes}$. Все эти собственные значения в силу симметричности матриц $X_*^{\otimes}$ и $V_*^{\otimes}$ вещественные. На самом деле в качестве матрицы U может быть взята ортогональная матрица $Q \otimes Q$ и имеет место

Лемма 3 (см. [13]). *Пусть выполнено разложение* (5). *Тогда справедливы следующие разложения:*

$$X_*^{\otimes} = (Q \otimes Q)D(\eta_*^{\otimes})(Q^{\mathrm{T}} \otimes Q^{\mathrm{T}}), \tag{80}$$

$$V_*^{\otimes} = (Q \otimes Q)D(\theta_*^{\otimes})(Q^{\mathrm{T}} \otimes Q^{\mathrm{T}}), \tag{81}$$

где $\eta_*^{\otimes}$ *и* $\theta_*^{\otimes}$ *– диагонали, соответственно, матриц*

$$D^{\otimes}(\eta_*) = [D(\eta_*) \otimes I_n + I_n \otimes D(\eta_*)]/2,$$

$$D^{\otimes}(\theta_*) = [D(\theta_*) \otimes I_n + I_n \otimes D(\theta_*)]/2.$$

Предположим, что точка V_* в разложении (5) такова, что

$$\theta_*^1 > 0, \ldots, \theta_*^r > 0, \quad \theta_*^{r+1} = \ldots = \theta_*^n = 0. \tag{82}$$

Тогда из предположения о строгой дополнительности X_* и V_* следует, что для соответствующих собственных значений $\eta_*^1, \ldots, \eta_*^n$ матрицы X_* выполняется условие

$$\eta_*^1 = \ldots = \eta_*^r = 0, \quad \eta_*^{r+1} > 0, \ldots, \eta_*^n > 0.$$

Пусть, как и прежде, $n_2 = k_\Delta(n - r)$, $n_1 = k_\Delta(n) - n_2$. Пусть, кроме того, $n_3 = k_\Delta(r)$. В этом случае у вектора $\tilde{\theta}_*^\otimes$ первые компоненты n_1 положительные, а остальные – нулевые. Напротив, у вектора $\tilde{\eta}_*^\otimes = \mathcal{L}_n \eta_*^\otimes \mathcal{D}_n$ заведомо положительными являются только последние компоненты n_2, а среди первых компонент n_1 найдутся нулевые компоненты в количестве n_3 штук.

Теорема 1. *Пусть X_* и $[u_*, V_*]$ – решения прямой и двойственной задач* (1) *и* (2), *причем X_* и V_* строго комплементарны. Пусть, кроме того, точки X_* и V_* невырожденные. Тогда итерации, определяемые по формуле* (31), *в которой шаг α_k берется постоянным и достаточно малым, локально сходятся к $[u_*, V_*]$ с линейной скоростью.*

Доказательство. Воспользуемся представлением (60) для итерационного процесса (31). Пусть $\alpha_k = \alpha$, где $\alpha > 0$. Если ввести отображение $\mathcal{B}(Z) = Z + \alpha\mathcal{G}_2(Z)$, то точка $Z_* = [u_*, \operatorname{vech} V_*]$, согласно сказанному выше, является неподвижной точкой этого отображения для любого $\alpha > 0$.

Пусть $\mathcal{B}_Z(Z_*)$ – матрица Якоби отображения $\mathcal{B}(Z)$ в точке Z_*. Покажем, что найдется $\bar{\alpha} > 0$, для которого при $0 < \alpha < \bar{\alpha}$ спектральный радиус ρ этой матрицы удовлетворяет условию

$$\rho(\mathcal{B}_Z(Z_*)) < 1. \tag{83}$$

Тогда, по теореме Островского (см., например, [20]), вектор Z_* является точкой притяжения для итерационного процесса (60) и итерации, определяемые процессом (60), локально сходятся к Z_* с линейной скоростью.

Согласно сказанному в предыдущем разделе, квадратная матрица $\mathcal{B}_Z(Z_*)$ порядка $m + k_\Delta(n)$ имеет вид

$$\mathcal{B}_Z(Z_*) = I - \alpha\mathcal{Q}_Z(Z_*). \tag{84}$$

Поэтому, чтобы найти спектральный радиус матрицы $\mathcal{B}_Z(Z_*)$, фактически остается только найти собственные значения матрицы $\mathcal{Q}_Z(Z_*)$. Пусть λ – ее произвольное собственное значение. Тогда оно удовлетворяет следующему характеристическому уравнению:

$$\begin{vmatrix} \tau\mathcal{A}_{\mathrm{vecs}}\Phi^{-1}\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} - \lambda I_m & \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(\tau I_{k_\Delta(n)} - \tilde{X}_*^\otimes) \\ \tau\tilde{V}_*^\otimes\Phi^{-1}\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} & (I_{k_\Delta(n)} - \tilde{V}_*^\otimes\Phi^{-1})\tilde{X}_*^\otimes + \tau\tilde{V}_*^\otimes\Phi^{-1} - \lambda I_{k_\Delta(n)} \end{vmatrix} = 0,$$

в котором $\Phi^{-1} = \Phi^{-1}(V_*)$.

Умножим правый столбец справа на матрицу $\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}$ и вычтем его из левого столбца. Тогда получим

$$\begin{vmatrix} \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}\tilde{X}_*^\otimes\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} - \lambda I_m & \mathcal{A}_{\mathrm{vecs}}\Phi^{-1}(\tau I_{k_\Delta(n)} - \tilde{X}_*^\otimes) \\ \Lambda_1 & \Lambda_2 \end{vmatrix} = 0,$$

где введены обозначения

$$\Lambda_1 = (\tilde{V}_*^\otimes\Phi^{-1} - I_{k_\Delta(n)})\tilde{X}_*^\otimes\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}} + \lambda\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}},$$

$$\Lambda_2 = (I_{k_\Delta(n)} - \tilde{V}_*^\otimes\Phi^{-1})\tilde{X}_*^\otimes + \tau\tilde{V}_*^\otimes\Phi^{-1} - \lambda I_{k_\Delta(n)}.$$

Умножим теперь первую строку слева на матрицу $\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}$ и сложим ее со второй строкой. Учитывая, что

$$(\mathcal{A}_{\mathrm{vech}}^{\mathrm{T}}\mathcal{A}_{\mathrm{vecs}} + \tilde{V}_*^\otimes)\Phi^{-1} = I_{k_\Delta(n)},$$

приходим к уравнению

$$\begin{vmatrix} \mathcal{A}_{\text{vecs}}\Phi^{-1}\tilde{X}_*^{\otimes}\mathcal{A}_{\text{vech}}^{\text{т}} - \lambda I_m & \mathcal{A}_{\text{vecs}}\Phi^{-1}(\tau I_{k_\Delta(n)} - \tilde{X}_*^{\otimes}) \\ 0_{k_\Delta(n)m} & (\tau-\lambda)I_{k_\Delta(n)} \end{vmatrix} = 0.$$

Отсюда сразу следует, что у матрицы $\mathfrak{Q}_Z(Z_*)$ имеются собственные значения в количестве $k_\Delta(n)$ штук, равные τ. Остальные собственные значения совпадают с собственными значениями матрицы

$$\mathfrak{Q}_1 = \mathcal{A}_{\text{vecs}}\Phi^{-1}(V_*)\tilde{X}_*^{\otimes}\mathcal{A}_{\text{vech}}^{\text{т}}.$$

Найдем конкретный вид матрицы $\mathfrak{Q}_1$. Используя (80), аналогично тому как это делалось для матрицы $\tilde{V}_*$, получаем

$$\tilde{X}_*^{\otimes} = \mathcal{L}_n(Q\otimes Q)D(\eta_*^{\otimes})(Q^{\text{т}}\otimes Q^{\text{т}})\mathfrak{D}_n = \mathcal{H}\mathcal{L}_nD(\eta_*^{\otimes})\mathfrak{D}_n\mathcal{H}^{-1} = \mathcal{H}D(\tilde{\eta}_*^{\otimes})\mathcal{H}^{-1}.$$

Тогда согласно (47)–(49) имеем

$$\mathfrak{Q}_1 = \mathcal{A}_{\text{vec2}}^{Q}\mathcal{W}_1^{-1}(\tilde{\theta}_*^{\otimes})D(\tilde{\eta}_*^{\otimes})(\mathcal{A}_{\text{vec2}}^{Q})^{\text{т}}.$$

Подставляя выражение (55) для матрицы $\mathcal{W}_1^{-1}(\tilde{\theta}_*^{\otimes})$, приходим к разложению

$$\mathfrak{Q}_1 = [NB]\begin{bmatrix} \mathcal{W}_{11}^{-1} & \mathcal{W}_{12}^{-1} \\ \mathcal{W}_{21}^{-1} & \mathcal{W}_{22}^{-1} \end{bmatrix}\begin{bmatrix} D^N(\tilde{\eta}_*^{\otimes}) & 0 \\ 0 & D^B(\tilde{\eta}_*^{\otimes}) \end{bmatrix}\begin{bmatrix} N^{\text{т}} \\ B^{\text{т}} \end{bmatrix},$$

где через $D^N(\tilde{\eta}_*^{\otimes})$ и $D^B(\tilde{\eta}_*^{\otimes})$ обозначены левая верхняя и правая нижняя дагональные подматрицы матрицы $D(\tilde{\eta}_*^{\otimes})$ порядков, соответственно, n_1 и n_2. Таким образом, имеем

$$\mathfrak{Q}_1 = [NB]\begin{bmatrix} \mathcal{W}_{11}^{-1}D^N(\tilde{\eta}_*^{\otimes}) & \mathcal{W}_{12}^{-1}D^B(\tilde{\eta}_*^{\otimes}) \\ \mathcal{W}_{21}^{-1}D^N(\tilde{\eta}_*^{\otimes}) & \mathcal{W}_{22}^{-1}D^B(\tilde{\eta}_*^{\otimes}) \end{bmatrix}\begin{bmatrix} N^{\text{т}} \\ B^{\text{т}} \end{bmatrix},$$

откуда следует, что

$$\mathfrak{Q}_1 = N\mathcal{W}_{11}^{-1}D^N(\tilde{\eta}_*^{\otimes})N^{\text{т}} + B\mathcal{W}_{21}^{-1}D^N(\tilde{\eta}_*^{\otimes})N^{\text{т}} + N\mathcal{W}_{12}^{-1}D^B(\tilde{\eta}_*^{\otimes})B^{\text{т}} + B\mathcal{W}_{22}^{-1}D^B(\tilde{\eta}_*^{\otimes})B^{\text{т}}.$$

После подстановки соответствующих выражений (56), (57) и перегруппировки слагаемых получаем $\mathfrak{Q}_1 = \mathfrak{Q}_1^{(1)} + \mathfrak{Q}_1^{(2)}$, где

$$\mathfrak{Q}_1^{(1)} = \mathcal{P}_\perp N\mathcal{K}^{-1}D^N(\tilde{\eta}_*^{\otimes})N^{\text{т}},$$

$$\mathfrak{Q}_1^{(2)} = (1 - \mathcal{P}_\perp N\mathcal{K}^{-1}N^{\text{т}})B(B^{\text{т}}B)^{-1}D^B(\tilde{\eta}_*^{\otimes})B^{\text{т}}.$$

Определим собственные значения матрицы $\mathfrak{Q}_1$. Пусть λ – произвольное собственное значение, и пусть y – соответствующий ему собственный вектор матрицы $\mathfrak{Q}_1$. Тогда имеем

$$\mathfrak{Q}_1 y = \lambda y. \tag{85}$$

Предположим сначала, что $B^{\text{т}}y \neq 0_{n_2}$ и, стало быть, $\mathcal{P}y \neq 0_m$, $y \notin \mathcal{L}_B^{\perp}$. Тогда, умножая равенство (85) слева на $B^{\text{т}}$, приходим к соотношению

$$D^B(\tilde{\eta}_*^{\otimes})B^{\text{т}}y = \lambda B^{\text{т}}y. \tag{86}$$

Отсюда видно, что в случае, когда $\mathcal{P}y \neq 0_n$, для того чтобы быть собственным, вектор y должен быть ортогональным ко всем столбцам матрицы B, кроме одного. Тогда из (86) следует, что λ совпадает с соответствующим диагональным элементом матрицы $D^B(\tilde{\eta}_*^{\otimes})$, который, как уже отмечалось, положительный. Вектор y может иметь ненулевые проекции и на большее число столбцов матрицы B, только в этом случае всем им должны соответствовать одинаковые диагональные элементы $D^B(\tilde{\eta}_*^{\otimes})$. Итак, мы имеем собственные векторы в количестве n_2 штук, не принадлежащие подпространству $\mathcal{L}_B^{\perp}$, всем им отвечают действительные положительные собственные значения.

Рассмотрим теперь случай, когда $B^{\mathrm{T}}y = 0_{n_2}$, т.е. $\mathcal{P}y = 0_m$, $y \in \mathcal{L}_B^{\perp}$. Тогда равенство (85) сводится к следующему:

$$\mathcal{P}_{\perp}N\mathcal{K}^{-1}D^N(\tilde{\eta}_*^{\otimes})N^{\mathrm{T}}y = \lambda y.$$

Подставляя выражение (58) для матрицы $\mathcal{K}^{-1}$, получаем после проведения соответствующих выкладок соотношение

$$\Omega y = \lambda y, \quad \Omega = (I + \mathcal{Y})^{-1}\mathcal{P}_{\perp}N\tilde{\Lambda}_N N^{\mathrm{T}}\mathcal{P}_{\perp}, \tag{87}$$

где использованы обозначения $\mathcal{Y} = \mathcal{P}_{\perp}N\tilde{\Theta}_N^{-1}N^{\mathrm{T}}\mathcal{P}_{\perp}$ и

$$\tilde{\Lambda}_N = \Theta_N^{-1}D^N(\tilde{\eta}_*^{\otimes}) = (D^N(\tilde{\theta}_*^{\otimes}))^{-1}D^N(\tilde{\eta}_*^{\otimes}),$$

а также учтено, что $y = \mathcal{P}_{\perp}y$. Таким образом, y является собственным вектором матрицы Ω, соответствующим тому же самому собственному значению λ.

Но поскольку симметричная матрица $I + \mathcal{Y}$ является положительно-определенной, то $\Omega = (I + \mathcal{Y})^{-1/2}\Omega_1(I + \mathcal{Y})^{1/2}$, где

$$\Omega_1 = (I + \mathcal{Y})^{-1/2}\mathcal{P}_{\perp}N\tilde{\Lambda}_N N^{\mathrm{T}}\mathcal{P}_{\perp}(I + \mathcal{Y})^{-1/2}.$$

Отсюда следует, что матрица Ω подобна матрице Ω_1, которая является симметричной и положительно-полуопределенной. Следовательно, все собственные значения матрицы Ω действительные и неотрицательные, причем число положительных собственных значений совпадает с рангом матрицы $\Omega_2 = \mathcal{P}_{\perp}N\tilde{\Lambda}_N N^{\mathrm{T}}\mathcal{P}_{\perp}$.

У диагональной матрицы $D^N(\tilde{\theta}_*^{\otimes})$ все диагональные элементы положительные. У диагональной матрицы $D^N(\tilde{\eta}_*^{\otimes})$ имеется $k_{\Delta}(r)$ нулевых диагональных элементов. Это – элементы, стоящие на следующих парных номерах из J_{Δ}:

$$(1, 1), (2, 1), \ldots, (r, 1), (2, 2), \ldots, (r, 2), \ldots, (r, r). \tag{88}$$

Остальные диагональные элементы в количестве $n_1 - k_{\Delta}(r)$ штук строго положительны. Таким образом, у диагональной матрицы $\tilde{\Lambda}_N$ все диагональные элементы положительны, за исключением тех, которые расположены на парных номерах из (88). Эти диагональные элементы нулевые.

Так как, по предположению, точка X_* невырожденная, то в силу (38) имеет место неравенство $m \leq n_1 - k_{\Delta}(r)$. Кроме того, согласно необходимым и достаточным условиям невырожденности в прямой задаче, ранг матрицы $\mathcal{A}_{\mathrm{vec}2}^{Q}$, из которой удалены столбцы с парными номерами из (88), равен m.

Пусть N_1 – подматрица матрицы N, которая получается, если из N удалить столбцы с парными номерами из (88). Согласно сказанному выше, среди столбцов матрицы N_1 найдутся $m - n_2$ линейно независимых столбцов таких, что проекции этих столбцов на подпространство $\mathcal{L}_B^{\perp}$ порождают все это подпространство.

Пусть, кроме того, $\tilde{\Lambda}_{N_1}$ – диагональная матрица порядка $n_1 - k_\Delta(r)$, которая получается из матрицы $\tilde{\Lambda}_{N_1}$ путем удаления строк и столбцов с парными номерами (88). Матрица Ω_2 есть матрица Грама для системы векторов, являющихся столбцами матрицы $\tilde{\Lambda}_{N_1}^{1/2} N_1^{\mathrm{T}} \mathcal{P}_\perp$. Ранг такой матрицы равняется $m - n_2$. Поэтому и ранг матрицы Ω_2 равен $m - n_2$. В силу сказанного выше, отсюда заключаем, что у матрицы Ω имеется $m - n_2$ положительных собственных значений. Всем им соответствуют собственные векторы из подпространства $\mathcal{L}_B^\perp$, размерность которого также равна $m - n_2$.

Мы показали, что все собственные значения λ матрицы $\mathfrak{Q}_1$ положительны. Следовательно, все собственные значения λ общей матрицы $\mathfrak{Q}_Z(Z_*)$ также положительны. Пусть $\lambda_{\max}$ – максимальное собственное значение $\mathfrak{Q}_Z(Z_*)$. Так как, согласно (84), собственные значения матрицы Якоби $\mathcal{B}_Z(Z_*)$ равны $1 - \alpha\lambda$, то, беря $\bar{\alpha} < 2/\lambda_{\max}$, получаем, что при $0 < \alpha < \bar{\alpha}$ выполнено условие (83). Поэтому процесс (60) локально сходится с линейной скоростью к точке $Z_* = [u_*, V_*]$. Теорема доказана.

Следствие. Пусть выполнены предположения теоремы 1. Тогда итерации, определяемые по формуле (21), в которой шаг α_k берется постоянным и достаточно малым, локально сходятся к u_* с линейной скоростью.

Доказательство. Справедливость данного результата следует из уже упоминавшегося свойства итерационного процесса (31), заключающегося в том, что он становится процессом (21), когда в качестве начальной берется точка $[u_0, V_0] \in \mathcal{F}_D$.

Обратим внимание, что из утверждений теоремы и следствия вытекает локальная сходимость для любых достаточно близких к решениям начальных точек, в том числе и для недопустимых. Другими словами, матрица V_0 может и не быть положительно-определенной, что является полезным свойством с точки зрения устойчивости метода.

Стоит также отметить, что условия (14) не изменятся, если в первом равенстве $V^{\otimes}\mathrm{vec}X = 0_{n^2}$ вместо матрицы $V^{\otimes}$ взять ее квадрат. Нуль-пространства у матрицы $V^{\otimes}$ и ее квадрата совпадают. Используя эти измененные условия и повторяя рассуждения, можно было бы построить итерационные процессы, в правых частях которых вместо $V_k^{\otimes}$ стояли бы квадраты этой матрицы.

СПИСОК ЛИТЕРАТУРЫ

1. *Nesterov Yu.E., Nemirovski A.S.* Interior point polynomial algorithms in convex programming // SIAM Publications. Philadelphia: SIAM, 1994.
2. *Vandenberghe L., Boyd S.* Semidefinite programming // SIAM Rev. 1996. V. 38. P. 49–95.
3. *Todd M.J.* Semidefinite optimization // Acta Numer. 2001. V. 10. P. 515–560.
4. *Laurent M., Rendle F.* Semidefinite programming and integer programming // Discrete optimization. Handbookd Operat. Res. and Management. Sci. Amsterdam: Elsevier, 2002.
5. H. Wolkowicz, R. Saigal, L. Vandenberghe. Handbook of semidefinite programming / Eds Boston: Kluwer Acad. Publs, 2000.
6. *De Klerk E.* Aspects of semidefinite programming. Interior point algorithsm and selected applications. Dordrecth: Kluwer Acad. Publs, 2004.
7. *Alizadeh F.* Interior point methods in semidefinite programming with applications to combinatiorial optimization // SIAM J. Optimizat. 1995. V. 5. P. 13–51.
8. *Monteiro R.D.C.* Primal-dual path-following algorithms for semidefinite programming // SIAM J. Optimizat. 1997. V. 7. P. 663–678.
9. *Monteiro R.D.C.* First- and second-order methods for semidefinite programming // Math. Program. Ser. B. 2003. V. 97. № 1–2. P. 209–244.
10. *Дикин И.И.* Сходимость последовательности векторов двойственных переменных при исследовании одной задачи полуопределенного программирования // Кибернетика и системный анализ. 2003. № 2. С. 156–163.
11. *Поляк Б.Т., Щербаков П.С.* Рандомизированный метод решения задачи полуопределенного программирования // Стохастич. оптимизация в информатике. 2006. Т. 2. № 1. С. 38–70.
12. *Евтушенко Ю.Г., Жадан В.Г.* Двойственные барьерно-проективные и барьерно-ньютоновские методы для линейного программирования // Ж. вычисл. матем. и матем. физ. 1996. Т. 36. № 7. С. 30–45.

13. *Бабынин М.С., Жадан В.Г.* Прямой метод внутренней точки для линейной задачи полуопределенного программирования // Ж. вычисл. матем. и матем. физ. 2008. Т. 48. № 10. С. 1780–1801.
14. *Muramatsu M.* Affine scaling algorithm fails for semidefinite programming // Math. Program. 1998. V. 83. P. 393–406.
15. *Magnus J.R., Neudecker H.* The elimination matrix: some lemmas and applications // SIAM J. Alg. Disc. Meth. 1980. V. 1. № 4. P. 422–449.
16. *Магнус Я.Р., Нейдеккер Ч.* Матричное дифференциальное исчисление с приложениями к статистике и эконометрике. М.: Физматлит, 2002.
17. *Тыртышников Е.Е.* Матричный анализ и линейная алгебра. М.: Физматлит, 2007.
18. *Арнольд В.И.* О матрицах, зависящих от параметров // Успехи матем. наук. 1971. Т. 26. Вып. 2(158). С. 101–114.
19. *Alizadeh F., Haeberly J.-P.F., Overton M.L.* Complementarity and nondegeneracy in semidefinite programming // Math. Program. Series B. 1997. V. 7. № 2. P. 129–162.
20. *Ортега Дж., Рейнболдт В.* Итерационные методы решения систем нелинейных уравнений со многими неизвестными. М.: Мир, 1975.