В. Г. Жадан, А. А. Орлов, Прямо-двойственный метод Ньютона для линейной задачи полуопределенного программирования, *Тр. ИММ УрО РАН*, 2013, том 19, номер 2, 157–169

ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 19 № 2 2013

УДК 519.854

ПРЯМО-ДВОЙСТВЕННЫЙ МЕТОД НЬЮТОНА ДЛЯ ЛИНЕЙНОЙ ЗАДАЧИ ПОЛУОПРЕДЕЛЕННОГО ПРОГРАММИРОВАНИЯ[1]

В. Г. Жадан, А. А. Орлов

Рассматривается линейная задача полуопределенного программирования. Для ее решения предлагается прямо-двойственный метод Ньютона. Показывается, что при условии невырожденности решений прямой и двойственной задач и их строгой дополнительности метод обладает локальной сходимостью со сверхлинейной скоростью.

Ключевые слова: задача полуопределенного программирования, метод Ньютона, прямо-двойственный метод, локальная сходимость.

V. G. Zhadan, A. A. Orlov. Primal–dual Newton method for a linear problem of semidefinite programming.

A linear problem of semidefinite programming is considered, and the primal–dual Newton method is proposed for its solution. The superlinear local convergence of the method is established under the assumption that the primal and dual problems are nondegenerate and strictly complementary.

Keywords: semidefinite programming problem, Newton method, primal–dual method, local convergence.

Введение

Задачи линейной оптимизации являются одними из наиболее важных в математическом программировании. Большой вклад в развитие теории и методов решения таких задач внесли И. И. Еремин [1; 2] и его ученики. В настоящей работе рассматривается линейная задача полуопределенного программирования. Она заключается в нахождении минимума линейной целевой функции при линейных ограничениях типа равенства на конусе положительно полуопределенных симметричных матриц. Изучению таких задач в последние два десятилетия уделялось много внимания [3; 4]. Был предложен ряд численных методов для их решения, обобщающих главным образом методы внутренней точки, разработанные ранее для решения задач линейного программирования. Среди них наиболее эффективными оказались прямо-двойственные методы [5–7]. Однако имеются примеры, когда прямо-двойственные методы не обладают сходимостью [8].

В настоящей работе рассматривается один из вариантов прямо-двойственного метода, который строится на основе применения метода Ньютона для решения системы равенств, входящих в условия оптимальности. Показывается, что при выполнении условий невырожденности для обеих задач и условия строгой дополнительности решений метод обладает локальной сходимостью. Метод может рассматриваться также как обобщение на задачи полуопределенного программирования прямо-двойственного метода [9].

В разд. 1 даются постановки прямой и двойственной линейных задач полуопределенного программирования в стандартной форме, а также вводятся основные определения и обозначения. В разд. 2 приводятся достаточные условия оптимальности, имеющие симметричный вид. В качестве переменных в них используются исходные прямые переменные и слабые двойственные переменные (двойственные невязки). Явный вид ньютоновских направлений в неособых точках дается в разд. 3. Наконец, в разд. 4 описывается итерационный процесс, обосновывается его локальная сходимость. Раздел 5 посвящен результатам вычислительного эксперимента.

[1]Работа выполнена при поддержке РФФИ (проект 11-01-00786), программ государственной поддержки ведущих научных школ (НШ-4096.2010.1) и президиума РАН П-15.

1. Постановка задачи и обозначения

Пусть $\mathcal{S}^n$ — пространство симметричных матриц порядка n. Его размерность равна так называемому "треугольному числу" $k_\triangle(n) = n(n+1)/2$. Неравенство $M \succeq 0$ означает, что матрица M принадлежит конусу $\mathcal{S}^n_+ \subset \mathcal{S}^n$, состоящему из положительно полуопределенных матриц. Соответственно, неравенство $M \succ 0$ означает, что матрица M принадлежит конусу $\mathcal{S}^n_{++} \subset \mathcal{S}^n_+$ положительно определенных матриц.

Скалярное (внутреннее) произведение двух матриц L и M одного порядка определяется как след матрицы $L^T M$ и обозначается как

$$L \bullet M = \mathrm{tr}\,(L^T M) = \sum\nolimits_{i,j=1}^n l_{ij} m_{ij},$$

где l_{ij} и m_{ij} — (ij)-е элементы соответственно матриц L и M. Если $L \in \mathcal{S}^n_+$ и $M \in \mathcal{S}^n_+$, то обязательно $L \bullet M \geq 0$. Более того, $L \bullet M = 0$ в том и только в том случае, когда $LM = ML = 0_{nn}$.

Рассмотрим линейную задачу полуопределенного программирования: найти

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

где матрицы C, X и A_i, $1 \leq i \leq m$, принадлежат пространству $\mathcal{S}^n$. Предполагается, что матрицы A_i, $1 \leq i \leq m$, линейно независимы. Двойственной к (1) является задача

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

в которой $b = (b^1, \ldots, b^m)^T \in \mathbb{R}^m$, $u \in \mathbb{R}^m$ и $V \in \mathcal{S}^n$. Угловые скобки указывают на обычное евклидово скалярное произведение в векторном пространстве $\mathbb{R}^m$.

Обозначим допустимые множества в прямой и двойственной задачах соответственно $\mathcal{F}_P$ и $\mathcal{F}_D$. Через $\mathcal{F}_{D,V}$ обозначим проекцию множества $\mathcal{F}_D$ на конус $\mathcal{S}^n_+$:

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

Через $V(u)$ будем обозначать зависимость $V(u) = C - \sum_{i=1}^m u^i A^i$.

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 $X_* \bullet V_* = 0$. В этом случае матрицы X_* и V_* коммутируют между собой. Поэтому найдется такая ортогональная матрица Q, что

$$X_* = Q D(\eta_*) Q^T, \quad V_* = Q D(\theta_*) Q^T, \tag{3}$$

где $\eta_* = [\eta^1_*, \ldots, \eta^n_*]$ и $\theta_* = [\theta^1_*, \ldots, \theta^n_*]$ — собственные значения матриц X_* и V_* соответственно. Для η^i_* и θ^i_* выполняется *условие дополнительности* $\eta^i_* \theta^i_* = 0$, $1 \leq i \leq n$. *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.*

Ниже нам потребуется ряд обозначений. Единичная матрица порядка n обозначается как I_n. Символ $\otimes$ между матрицами указывает на 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 рассматривается также 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 $k_\triangle(n)$.

Для перехода от вектора $\mathrm{vec}\, M$ к вектору $\mathrm{vech}\, M$ и для обратного перехода используются специальные *элиминационные* и *дуплицирующие* матрицы (см. [10]). Элиминационная матрица $\mathcal{L}_n$ для каждой квадратной матрицы M порядка n совершает преобразование $\mathcal{L}_n \operatorname{vec} M = \operatorname{vech} M$. Напротив, дуплицирующая матрица $\mathcal{D}_n$ для каждой симметричной матрицы M порядка n осуществляет обратное преобразование $\mathcal{D}_n \operatorname{vech} M = \operatorname{vec} M$. Матрица $\mathcal{L}_n$ имеет порядок $k_\triangle(n) \times n^2$, матрица $\mathcal{D}_n$ — порядок $n^2 \times k_\triangle(n)$. Обе матрицы $\mathcal{L}_n$ и $\mathcal{D}_n$ являются матрицами полного ранга, равного $k_\triangle(n)$. Матрица $\mathcal{L}_n$ полуортогональна, т. е. $\mathcal{L}_n \mathcal{L}_n^T = I_{k_\triangle(n)}$. Кроме того, $\mathcal{L}_n \mathcal{D}_n = I_{k_\triangle(n)}$.

При работе с векторами $\operatorname{vech} M$ или $\operatorname{vecs} M$ удобно нумеровать их элементы не числами натурального ряда, а парами индексов из набора

$$J_\triangle = \{(1,1), \dots, (n,1), (2,2), \dots, (n,2), (3,3), \dots, (n,n)\},$$

содержащего $k_\triangle(n)$ пар индексов. Номера вида (i,i), в которых первый индекс совпадает со вторым, будем называть *диагональными*.

2. Условия оптимальности

Для существования решений у пары задач (1) и (2) достаточно, чтобы была разрешима следующая система равенств и неравенств:

$$\begin{aligned} X \bullet V &= 0, \\ A_i \bullet X &= b^i, \quad 1 \le i \le m, \\ V &= C - \textstyle\sum_{i=1}^m u^i A_i, \\ X \succeq 0, &\quad V \succeq 0. \end{aligned} \tag{4}$$

Ниже будем считать, что задача (4) имеет решение. Обозначим через $X * V$ симметризованное произведение матриц X и V из $\mathcal{S}^n$, 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 в том случае, когда $XV = VX = 0_{nn}$.*

Доказательство. Докажем только необходимость. Пусть $X * V = 0_{nn}$, тогда $XV = -(XV)^T$. 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 $V \succeq 0$, то $XV = VX = 0_{nn}$.

Утверждение доказано.

С учетом утверждения 1 и положительной полуопределенности матриц X и V первое равенство в (4) может быть переписано следующим образом: $X * V = 0_{nn}$. Заменяя теперь это и остальные матричные равенства из системы (4) на векторные аналоги, получаем:

$$\begin{aligned} \operatorname{vec}(X * V) &= 0_{nn}, \\ \mathcal{A}_{vec} \operatorname{vec} X &= b, \\ \operatorname{vec} V &= \operatorname{vec} C - \mathcal{A}_{vec}^T u, \\ X \succeq 0, &\quad V \succeq 0. \end{aligned} \tag{5}$$

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

Опустим пока в системе (5) неравенства $X \succeq 0$, $V \succeq 0$ и основное внимание уделим равенствам. Обратимся сначала к первому равенству. В силу известной формулы $\operatorname{vec}(ABC) = (C^T \otimes A) \operatorname{vec} B$, справедливой для любых матриц A, B и C, для которых определено произведение ABC, получаем, что $\operatorname{vec}(X * V) = X^{\otimes} \operatorname{vec} V = V^{\otimes} \operatorname{vec} X$, где $X^{\otimes} = [X \otimes I_n + I_n \otimes X]/2$, $V^{\otimes} = [V \otimes I_n + I_n \otimes V]/2$ — кронекеровские суммы матриц X и V соответственно. Таким образом, первое равенство из (5) может быть представлено в виде

$$X^{\otimes} \operatorname{vec} V = V^{\otimes} \operatorname{vec} X = 0_{n^2}. \tag{6}$$

Преобразуем теперь третье равенство из (5). Обозначим через $\mathcal{R}_A$ подпространство в $\mathcal{S}^n$, порожденное матрицами $A_1, \ldots, A_m$, через $\mathcal{R}_A^\perp$ — его ортогональное дополнение. Пусть $l = k_\triangle(n) - m$ и $\tilde{A}_1, \ldots, \tilde{A}_l$ — линейно независимые матрицы из $\mathcal{R}_A^\perp$. Они задают некоторый базис в $\mathcal{R}_A^\perp$. Составим из этих векторов матрицу $\tilde{\mathcal{A}}_{vec}$ порядка $l \times n^2$, помещая их в качестве строк. Символом $\tilde{b}$ обозначим l-мерный вектор с элементами $\tilde{b}_j = \tilde{A}_j \bullet C$, $1 \le j \le l$. Тогда после умножения третьего равенства из (5) слева на матрицу $\tilde{\mathcal{A}}_{vec}$ получим равенство $\tilde{\mathcal{A}}_{vec} V = \tilde{b}$. Объединяя его с (6) и вторым равенством из системы (5), получим:

$$\begin{aligned} X^{\otimes}\,\mathrm{vec}\; V = V^{\otimes}\,\mathrm{vec}\; X &= 0_{n^2},\\ \mathcal{A}_{vec}\,\mathrm{vec}\; X &= b,\\ \tilde{\mathcal{A}}_{vec}\,\mathrm{vec}\; V &= \tilde{b}. \end{aligned}$$

Если дополнительно учесть симметричность всех матриц, то имеем:

$$\begin{aligned} \mathcal{L}_n V^{\otimes}\mathcal{D}_n\,\mathrm{vech}\; X = \mathcal{L}_n X^{\otimes}\mathcal{D}_n\,\mathrm{vech}\; V &= 0_{k_\triangle(n)},\\ \mathcal{A}_{vecs}\,\mathrm{vech}\; X &= b,\\ \tilde{\mathcal{A}}_{vecs}\,\mathrm{vech}\; V &= \tilde{b}. \end{aligned} \tag{7}$$

Здесь $\mathcal{A}_{vecs}$ и $\tilde{\mathcal{A}}_{vecs}$ — матрицы, строками которых являются соответственно векторы $\mathrm{vecs}\,A_i$, $1 \le i \le m$, и $\mathrm{vecs}\,\tilde{A}_j$, $1 \le j \le l$. Любые положительно полуопределенные матрицы X и V, удовлетворяющие системе (7), являются одновременно решениями линейных задач полуопределенного программирования (1) и (2).

3. Ньютоновские направления

Система (7) является системой $2k_\triangle(n)$ уравнений относительно $2k_\triangle(n)$ переменных vech X и vech V. Если решать ее методом Ньютона, то ньютоновские направления vech ΔX и vech ΔV в паре $[X, V]$ должны удовлетворять следующей системе уравнений:

$$\begin{aligned} \tilde{V}^{\otimes}\,\mathrm{vech}\;\Delta X + \tilde{X}^{\otimes}\,\mathrm{vech}\;\Delta V &= -r_c(X, V),\\ \mathcal{A}_{vecs}\,\mathrm{vech}\;\Delta X &= -r_p(X),\\ \tilde{\mathcal{A}}_{vecs}\,\mathrm{vech}\;\Delta V &= -r_d(X). \end{aligned} \tag{8}$$

Здесь и далее $\tilde{X}^{\otimes} = \mathcal{L}_n X^{\otimes}\mathcal{D}_n$, $\tilde{V}^{\otimes} = \mathcal{L}_n V^{\otimes}\mathcal{D}_n$. Через $r_c(X,V)$, $r_p(X)$ и $r_d(V)$ обозначены невязки $r_c(X,V) = \mathrm{vech}\; X * V$, $r_p(X) = \mathcal{A}_{vecs}\,\mathrm{vech}\; X - b$, $r_d(V) = \tilde{\mathcal{A}}_{vecs}\,\mathrm{vech}\; V - \tilde{b}$, причем $r_c(X,V) = \mathrm{vech}\; X * V = \mathcal{L}_n\,\mathrm{vec}\; X * V = \tilde{X}^{\otimes}\,\mathrm{vech}\; V = \tilde{V}^{\otimes}\,\mathrm{vech}\; X$.

Если ввести квадратную матрицу $\mathcal{W}(X,V)$ и $k_\triangle(n)$-мерный вектор невязок $r_f(X,V)$, положив

$$\mathcal{W}(X,V) = \begin{bmatrix} \tilde{V}^{\otimes} & \tilde{X}^{\otimes} \\ \mathcal{A}_{vecs} & 0_{m\,k_\triangle(n)} \\ 0_{l\,k_\triangle(n)} & \tilde{\mathcal{A}}_{vecs} \end{bmatrix}, \quad r_f(X,V) = \begin{bmatrix} r_p(X) \\ r_d(V) \end{bmatrix}, \tag{9}$$

то система (8) может быть переписана в виде

$$\mathcal{W}(X,V) \begin{bmatrix} \mathrm{vech}\;\Delta X \\ \mathrm{vech}\;\Delta V \end{bmatrix} = -\begin{bmatrix} r_c(X,V) \\ r_f(X,V) \end{bmatrix}. \tag{10}$$

Найдем решение системы (8), предполагая, что пара $[X, V]$ является неособой, т. е. X и V — неособые матрицы. В этом случае $\tilde{X}^{\otimes}$ и $\tilde{V}^{\otimes}$, как впрочем и $\mathcal{W}(X,V)$, будут неособыми матрицами. Пусть $q \in \mathbb{R}^{k_\triangle(n)}$ и $p \in \mathbb{R}^{k_\triangle(n)}$— произвольные векторы, удовлетворяющие соответственно равенствам

$$\mathcal{A}_{vecs} q = b, \quad \tilde{\mathcal{A}}_{vecs} p = \tilde{b}. \tag{11}$$

Тогда согласно (8) $\mathcal{A}_{vecs}\,(\mathrm{vech}\;\Delta X + \mathrm{vech}\; X - q) = 0_m$, $\tilde{\mathcal{A}}_{vecs}\,(\mathrm{vech}\;\Delta V + \mathrm{vech}\; V - p) = 0_l$. Отсюда делаем вывод, что ньютоновские направления можно искать в виде

$$\mathrm{vech}\;\Delta X = q - \mathrm{vech}\; X + \tilde{\mathcal{A}}^T_{vech}\Delta w, \quad \mathrm{vech}\;\Delta V = p - \mathrm{vech}\; V + \mathcal{A}^T_{vech}\Delta u, \tag{12}$$

где $\Delta w \in \mathbb{R}^l$, $\Delta u \in \mathbb{R}^m$. Первое из равенств (8) после подстановки этих выражений переписывается как

$$\tilde{V}^{\otimes}\tilde{\mathcal{A}}^T_{vech}\Delta w + \tilde{X}^{\otimes}\mathcal{A}^T_{vech}\Delta u = r_c(X,V) - \tilde{V}^{\otimes}q - \tilde{X}^{\otimes}p. \quad (13)$$

Обозначим $\tilde{X}^{-\otimes} = \big(\tilde{X}^{\otimes}\big)^{-1}$, $\tilde{V}^{-\otimes} = \big(\tilde{V}^{\otimes}\big)^{-1}$. Если умножить равенство (13) слева на матрицу $\mathcal{A}_{vecs}\tilde{V}^{-\otimes}$, то приходим к системе линейных уравнений для определения вектора Δu:

$$\mathcal{A}_{vecs}\tilde{V}^{-\otimes}\tilde{X}^{\otimes}\mathcal{A}^T_{vech}\Delta u = \mathcal{A}_{vecs}\big[\,\mathrm{vech}\ X - q - \tilde{V}^{-\otimes}\tilde{X}^{\otimes}p\big]. \quad (14)$$

Аналогично, после умножения равенства (13) слева на матрицу $\tilde{\mathcal{A}}_{vecs}\tilde{X}^{-\otimes}$ получаем другую линейную систему для определения вектора Δw:

$$\tilde{\mathcal{A}}_{vecs}\tilde{X}^{-\otimes}\tilde{V}^{\otimes}\tilde{\mathcal{A}}^T_{vech}\Delta w = \tilde{\mathcal{A}}_{vecs}\big[\,\mathrm{vech}\ V - p - \tilde{X}^{-\otimes}\tilde{V}^{\otimes}q\big]. \quad (15)$$

Поскольку обе матрицы $\mathcal{A}_{vecs}\tilde{V}^{-\otimes}\tilde{X}^{\otimes}\mathcal{A}^T_{vech}$ и $\tilde{\mathcal{A}}_{vecs}\tilde{X}^{-\otimes}\tilde{V}^{\otimes}\tilde{\mathcal{A}}^T_{vech}$ неособые, то на основании (14) и (15) имеем:

$$\Delta u = \big(\mathcal{A}_{vecs}\tilde{V}^{-\otimes}\tilde{X}^{\otimes}\mathcal{A}^T_{vech}\big)^{-1}\mathcal{A}_{vecs}\big[\,\mathrm{vech}\ X - q - \tilde{V}^{-\otimes}\tilde{X}^{\otimes}p\big],$$
$$\Delta w = \big(\tilde{\mathcal{A}}_{vecs}\tilde{X}^{-\otimes}\tilde{V}^{\otimes}\tilde{\mathcal{A}}^T_{vech}\big)^{-1}\tilde{\mathcal{A}}_{vecs}\big[\,\mathrm{vech}\ V - p - \tilde{X}^{-\otimes}\tilde{V}^{\otimes}q\big].$$

После подстановки найденных векторов Δw и Δu в (12) приходим к следующим выражениям для векторов $\mathrm{vech}\ \Delta X$ и $\mathrm{vech}\ \Delta V$:

$$\mathrm{vech}\ \Delta X = q - \mathrm{vech}\ X + \tilde{\mathcal{A}}^T_{vech}\big(\tilde{\mathcal{A}}_{vecs}\tilde{X}^{-\otimes}\tilde{V}^{\otimes}\tilde{\mathcal{A}}^T_{vech}\big)^{-1}\tilde{\mathcal{A}}_{vecs}\big[\,\mathrm{vech}\ V - p - \tilde{X}^{-\otimes}\tilde{V}^{\otimes}q\big],$$
$$\mathrm{vech}\ \Delta V = p - \mathrm{vech}\ V + \mathcal{A}^T_{vech}\big(\mathcal{A}_{vecs}\tilde{V}^{-\otimes}\tilde{X}^{\otimes}\mathcal{A}^T_{vech}\big)^{-1}\mathcal{A}_{vecs}\big[\,\mathrm{vech}\ X - q - \tilde{V}^{-\otimes}\tilde{X}^{\otimes}p\big].$$

Из (11) видно, что в качестве вектора p может быть взят, в частности, вектор $\mathrm{vech}\ C$. Пусть, кроме того, $\tilde{C}$ — произвольная симметричная матрица, такая что $\mathcal{A}_{vecs}\,\mathrm{vech}\ \tilde{C} = b$. Положим для единообразия $q = \mathrm{vech}\ \tilde{C}$. Тогда

$$\mathrm{vech}\ \Delta X = \mathrm{vech}\ \tilde{C} - \mathrm{vech}\ X$$
$$+\tilde{\mathcal{A}}^T_{vech}\big(\tilde{\mathcal{A}}_{vecs}\tilde{X}^{-\otimes}\tilde{V}^{\otimes}\tilde{\mathcal{A}}^T_{vech}\big)^{-1}\tilde{\mathcal{A}}_{vecs}\big[\,\mathrm{vech}\ V - \mathrm{vech}\ C - \tilde{X}^{-\otimes}\tilde{V}^{\otimes}\,\mathrm{vech}\ \tilde{C}\big],$$
$$\mathrm{vech}\ \Delta V = \mathrm{vech}\ C - \mathrm{vech}\ V$$
$$+\mathcal{A}^T_{vech}\big(\mathcal{A}_{vecs}\tilde{V}^{-\otimes}\tilde{X}^{\otimes}\mathcal{A}^T_{vech}\big)^{-1}\mathcal{A}_{vecs}\big[\,\mathrm{vech}\ X - \mathrm{vech}\ \tilde{C} - \tilde{V}^{-\otimes}\tilde{X}^{\otimes}\,\mathrm{vech}\ C\big].$$

В строго внутренней допустимой паре $[X,V]$, т.е. когда $X \succ 0$, $V \succ 0$ и выполняются равенства $\mathcal{A}_{vecs}\,\mathrm{vech}\ X = b$, $\tilde{A}\,\mathrm{vech}\ V = \tilde{b}$, выражения для $\mathrm{vech}\ \Delta X$ и $\mathrm{vech}\ \Delta V$ упрощаются:

$$\mathrm{vech}\ \Delta X = \big[I_{k_\triangle(n)} - \tilde{\mathcal{A}}^T_{vech}\big(\tilde{\mathcal{A}}_{vecs}\tilde{X}^{-\otimes}\tilde{V}^{\otimes}\tilde{\mathcal{A}}^T_{vech}\big)^{-1}\tilde{\mathcal{A}}_{vecs}\tilde{X}^{-\otimes}\tilde{V}^{\otimes}\big]\,\mathrm{vech}\ \tilde{C} - \mathrm{vech}\ X.$$
$$\mathrm{vech}\ \Delta V = \big[I_{k_\triangle(n)} - \mathcal{A}^T_{vech}\big(\mathcal{A}_{vecs}\tilde{V}^{-\otimes}\tilde{X}^{\otimes}\mathcal{A}^T_{vech}\big)^{-1}\mathcal{A}_{vecs}\tilde{V}^{-\otimes}\tilde{X}^{\otimes}\big]\,\mathrm{vech}\ C - \mathrm{vech}\ V.$$

Разумеется, достаточно определить одно из двух ньютоновских направлений $\mathrm{vech}\ \Delta X$ или $\mathrm{vech}\ \Delta X$, так как они связаны между собой соотношением (13). Например, если известно $\mathrm{vech}\ \Delta V$, то для $\mathrm{vech}\ \Delta X$ получаем

$$\mathrm{vech}\ \Delta X = -\tilde{V}^{-\otimes}\big[\tilde{X}^{\otimes}\,\mathrm{vech}\ \Delta V + r_c(X,V)\big] = -\big[\,\mathrm{vech}\ X + \tilde{V}^{-\otimes}\tilde{X}^{\otimes}\,\mathrm{vech}\ \Delta V\big].$$

Аналогично, если известно $\mathrm{vech}\ \Delta X$, то для $\mathrm{vech}\ \Delta V$ имеем

$$\mathrm{vech}\ \Delta V = -\tilde{X}^{-\otimes}\big[\tilde{V}^{\otimes}\,\mathrm{vech}\ \Delta X + r_c(X,V)\big] = -\big[\,\mathrm{vech}\ V + \tilde{X}^{-\otimes}\tilde{V}^{\otimes}\,\mathrm{vech}\ \Delta X\big].$$

Какую из систем (14) или (15) целесообразно решать для нахождения ньютоновских направлений, зависит главным образом от размерности этих систем.

4. Итерационный процесс

Рассмотрим теперь прямо-двойственный итерационный процесс для решения пары задач (1), (2), описываемый следующими рекуррентными соотношениями:

$$X_{k+1} = X_k + \Delta X_k, \quad V_{k+1} = V_k + \Delta V_k, \quad (16)$$

где $X_0 \in \mathcal{S}^n$, $V_0 \in \mathcal{S}^n$, в качестве матриц ΔX_k и ΔV_k берутся ньютоновские направления из $\mathcal{S}^n$, удовлетворяющие уравнению (10).

Если X_* и $V_* = V(u_*)$ — решения соответственно задач (1) и (2), то локальная сходимость итерационного процесса (16) к этим точкам будет следовать из общих утверждений о сходимости метода Ньютона. Например, достаточно показать, что матрица $\mathcal{W}(X_*, V_*)$ из (9) является неособой. Рассмотрим сначала квадратную матрицу порядка $k_\triangle(n)$

$$\Phi(V) = \mathcal{A}_{vech}^T \mathcal{A}_{vecs} + \tilde{V}^{\otimes}. \tag{17}$$

Приведем условия, при которых она является неособой. Нам потребуется понятие невырожденной точки. Предположим, что ранг матрицы $Z \in \mathcal{S}^n_+$ равен r. Тогда она представима в виде

$$Z = Q \operatorname{Diag}\left(\lambda^1, \ldots, \lambda^r, 0, \ldots, 0\right) Q^T, \tag{18}$$

где Q — ортогональная матрица, $\lambda^i > 0$, $1 \le i \le r$. Если $r < n$, то матрица Z принадлежит границе конуса $\mathcal{S}^n_+$. Касательное пространство к $\mathcal{S}^n_+$ в этой точке имеет следующий вид (см. [11]):

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

Здесь $\mathbb{R}^{r\times s}$ — пространство матриц порядка $r \times s$. Размерность $\mathcal{T}_Z$ равна $k_\triangle(n) - k_\triangle(n-r)$.

Следуя [12], дадим определение невырожденной точки X в прямой задаче (1) и невырожденной точки V в двойственной задаче (2). Пусть по-прежнему $\mathcal{R}_A^\perp$ — ортогональное дополнение линейного подпространства $\mathcal{R}_A$ в $\mathcal{S}^n$, порожденного матрицами A_i, $1 \le i \le m$. Ортогональное дополнение $\mathcal{R}_A^\perp$ совпадает с l-мерным линейным подпространством в $\mathcal{S}^n$, порожденным матрицами $\tilde{A}_j$, $1 \le j \le l$.

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

Если матрица $X \in \mathcal{F}_P$ имеет ранг r, то для того, чтобы она была невырожденной точкой в прямой задаче (1), необходимо и достаточно, чтобы выполнялось неравенство $m \le k_\triangle(n) - k_\triangle(n-r)$. Аналогично, матрица $V \in \mathcal{F}_{D,V}$ ранга r будет невырожденной точкой в двойственной задаче (2) в том и только в том случае, когда $k_\triangle(n-r) \le m$.

В [12] приведены условия невырожденности точки $Z \in \mathcal{S}^n_+$, как в прямой, так и в двойственной задачах. Пусть Q_1 и Q_2 — подматрицы матрицы Q в разложении (18), состоящие соответственно из первых r и последующих $n-r$ столбцов.

Лемма 1. *Точка $X \in \mathcal{F}_P$ будет невырожденной в прямой задаче* (1) *тогда и только тогда, когда матрицы*

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

линейно независимы. Соответственно, точка $V \in \mathcal{F}_{D,V}$ будет невырожденной в двойственной задаче (2) *тогда и только тогда, когда матрицы $Q_2^T A_i Q_2$, $1 \le i \le m$, порождают пространство $\mathcal{S}^{n-r}$.*

В дальнейшем предполагается, что по крайней мере решения задач (1) и (2) являются невырожденными точками.

Лемма 2. *В невырожденной точке $V_* \in \mathcal{F}_{D,V}$ матрица $\Phi(V_*)$ неособая.*

Доказательство. Пусть для V_* справедливо разложение $V_* = QD(\theta_*)Q^T$, где Q — ортогональная матрица, θ_* — n-мерный вектор, составленный из собственных чисел матрицы V_*. Тогда

$$V_*^{\otimes} = (Q \otimes Q) D(\theta_*^{\otimes})(Q^T \otimes Q^T). \tag{19}$$

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

Используя (19), представим матрицу $\Phi(V_*)$ в виде

$$\Phi(V_*) = \mathcal{A}_{vech}^T \mathcal{A}_{vecs} + \mathcal{L}_n (Q \otimes Q) D(\theta_*^{\otimes})(Q^T \otimes Q^T)\mathcal{D}_n. \tag{20}$$

Но для любой квадратной матрицы M порядка n справедливы формулы (см. [10])

$$\mathcal{D}_n \mathcal{L}_n (M \otimes M)\mathcal{D}_n = (M \otimes M)\mathcal{D}_n, \quad \mathcal{D}_n \mathcal{L}_n M^{\otimes} \mathcal{D}_n = M^{\otimes}\mathcal{D}_n.$$

С помощью этих формул выражение (20) приводится к виду

$$\Phi(V_*) = \mathcal{A}_{vech}^T \mathcal{A}_{vecs} + \mathcal{H}\mathcal{L}_n D(\theta_*^{\otimes})\mathcal{D}_n \mathcal{H}^{-1},$$

где $\mathcal{H} = \mathcal{L}_n(Q \otimes Q)\mathcal{D}_n$ — квадратная матрица порядка $k_{\triangle}(n)$. Так как $[\mathcal{L}_n(M \otimes M)\mathcal{D}_n]^{-1} = \mathcal{L}_n(M^{-1} \otimes M^{-1})\mathcal{D}_n$ для любой неособой матрицы M (см. [10]), то $\mathcal{H}^{-1} = \mathcal{L}_n(Q^T \otimes Q^T)\mathcal{D}_n$.

Пусть $\tilde{\theta}^{\otimes} = \mathcal{L}_n \theta^{\otimes}$. Легко проверить, что $\mathcal{L}_n D(\theta^{\otimes})\mathcal{D}_n = D(\tilde{\theta}^{\otimes})$. Далее, поскольку $Q \otimes Q$ — ортогональная матрица, то

$$\mathcal{A}_{vech}^T = \mathcal{L}_n(Q \times Q)(Q^T \otimes Q^T)\mathcal{A}_{vec}^T = \mathcal{H}\mathcal{L}_n(Q^T \otimes Q^T)\mathcal{A}_{vec}^T = \mathcal{H}(\mathcal{A}_{vech}^Q)^T,$$

$$\mathcal{A}_{vecs} = \mathcal{A}_{vec}(Q \times Q)(Q^T \otimes Q^T)\mathcal{D}_n = \mathcal{A}_{vec}(Q \times Q)\mathcal{D}_n \mathcal{H}^{-1} = \mathcal{A}_{vecs}^Q \mathcal{H}^{-1}.$$

Здесь и далее через $\mathcal{A}_{vech}^Q$ и $\mathcal{A}_{vecs}^Q$ обозначены матрицы порядка $m \times k_{\triangle}(n)$, строками которых являются соответственно векторы $\mathrm{vech}\,(Q^T A_i Q)$ и $\mathrm{vecs}\,(Q^T A_i Q)$, $1 \le i \le m$. Таким образом, для матрицы (20) справедливо представление

$$\Phi(V_*) = \mathcal{H}\mathcal{F}(\tilde{\theta}_*^{\otimes})\mathcal{H}^{-1}, \quad \mathcal{F}(\tilde{\theta}^{\otimes}) = (\mathcal{A}_{vech}^Q)^T \mathcal{A}_{vecs}^Q + D(\tilde{\theta}^{\otimes}). \tag{21}$$

Обозначим через D_2 диагональную матрицу порядка $k_{\triangle}(n)$. В ней компоненты вектора, стоящего на диагонали, равны единице, если их парный номер из $J_{\triangle}$ 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 очевидное равенство $D(\tilde{\theta}^{\otimes}) = D_2^{-1/2} D(\tilde{\theta}^{\otimes}) D_2^{1/2}$. Кроме того, $\mathcal{A}_{vecs}^Q = \mathcal{A}_{vech}^Q D_2$. Положим $\mathcal{A}_{vec2}^Q = \mathcal{A}_{vech}^Q D_2^{1/2}$. Тогда

$$\mathcal{F}(\tilde{\theta}^{\otimes}) = D_2^{-1/2}\mathcal{F}_1(\tilde{\theta}^{\otimes})D_2^{1/2}, \quad \mathcal{F}_1(\tilde{\theta}^{\otimes}) = (\mathcal{A}_{vec2}^Q)^T \mathcal{A}_{vec2}^Q + D(\tilde{\theta}^{\otimes}). \tag{22}$$

Матрица $\mathcal{F}_1(\tilde{\theta}_*^{\otimes})$ является симметричной положительно полуопределенной. Она будет положительно определенной, если линейная однородная система уравнений

$$\mathcal{F}_1(\tilde{\theta}_*^{\otimes})x = 0_{k_{\triangle}(n)} \tag{23}$$

имеет только тривиальное решение $x = 0_{k_{\triangle}(n)}$. Убедимся в этом. После умножения левой и правой части равенства (23) на x^T получаем с учетом положительной полуопределенности обеих матриц $(\mathcal{A}_{vec2}^Q)^T \mathcal{A}_{vec2}^Q$ и $D(\tilde{\theta}^{\otimes})$:

$$\langle x, (\mathcal{A}_{vec2}^Q)^T \mathcal{A}_{vec2}^Q x\rangle = 0, \quad \langle x, D(\tilde{\theta}_*^{\otimes})x\rangle = 0. \tag{24}$$

Если все собственные числа матрицы V_* строго положительны, то $\tilde{\theta}_*^{\otimes} > 0_{k_{\triangle}(n)}$, и из второго равенства (24) следует, что $x = 0_{k_{\triangle}(n)}$.

Далее считаем, что матрица V_* имеет ранг $r < n$, причем

$$\theta_*^1 > 0, \dots, \theta_*^r > 0, \quad \theta_*^{r+1} = \dots = \theta_*^n = 0. \tag{25}$$

Положим $n_1 = k_{\triangle}(n) - k_{\triangle}(n-r)$, $n_2 = k_{\triangle}(n-r)$. Из невырожденности точки $V_* \in \mathcal{F}_{D,V}$ следует неравенство $n_2 \le m$. Для вектора $\tilde{\theta}_*^{\otimes}$ справедливо разбиение

$$\tilde{\theta}_*^{\otimes} = \begin{bmatrix} (\tilde{\theta}_*^{\otimes})^N \\ (\tilde{\theta}_*^{\otimes})^B \end{bmatrix}, \quad (\tilde{\theta}_*^{\otimes})^N > 0_{n_1}, \quad (\tilde{\theta}_*^{\otimes})^B = 0_{n_2}. \tag{26}$$

Матрицу $\mathcal{A}_{vec2}^Q$ разобьем также на две подматрицы в соответствии с разбиением (26):

$$\mathcal{A}_{vec2}^Q = [N \mid B], \quad N = (\mathcal{A}_{vec2}^Q)^N \in \mathbb{R}^{m \times n_1}, \quad B = (\mathcal{A}_{vec2}^Q)^B \in \mathbb{R}^{m \times n_2}.$$

Положим $\tilde{\Theta}_*^N = D\big((\tilde{\theta}_*^{\otimes})^N\big)$. Тогда

$$\mathcal{F}_1(\tilde{\theta}_*^{\otimes}) = \begin{bmatrix} N^T N + \tilde{\Theta}_*^N & N^T B \\ B^T N & B^T B \end{bmatrix}. \qquad (27)$$

Разобьем вектор x, являющийся решением уравнения (23), на две части $x = [x^N, x^B]$ в соответствии с разбиением вектора $\tilde{\theta}^{\otimes}$. На основании второго равенства (24) заключаем, что $x^N = 0_{n_1}$. Поэтому первое равенство (24) сводится к $\langle x^B, B^T B x^B\rangle = 0$. Отсюда следует, что $Bx^B = 0_{n_2}$. Но в невырожденных точках $V \in \mathcal{F}_{D,V}$ согласно лемме 1 матрица B имеет полный ранг по столбцам, равный n_2. Таким образом, $x^B = 0$, и уравнение (23) имеет только тривиальное решение $x = 0_{k_\triangle(n)}$. Это означает невырожденность матрицы $\mathcal{F}_1(\tilde{\theta}_*^{\otimes})$ и, следовательно, всей матрицы $\Phi(V_*)$.

Лемма доказана.

Обратимся теперь к матрице

$$\Lambda(X,V) = \mathcal{A}_{vecs}\Phi^{-1}(V)\tilde{X}^{\otimes}\mathcal{A}_{vech}^T \quad (\tilde{X}^{\otimes} = \mathcal{L}_n X^{\otimes}\mathcal{D}_n). \qquad (28)$$

Лемма 3. *Пусть X_* и $V_* = V(u_*)$ — невырожденные строго комплементарные решения соответственно задач* (1) *и* (2). *Тогда матрица $\Lambda_* = \Lambda(X_*, V_*)$ неособая.*

Д о к а з а т е л ь с т в о. Так как X_* и V_* — решения задач (1) и (2), то их можно представить в виде (3). Тогда, как уже отмечалось, справедливы следующие разложения:

$$X_*^{\otimes} = (Q\otimes Q)D(\eta_*^{\otimes})(Q^T\otimes Q^T), \quad V_*^{\otimes} = (Q\otimes Q)D(\theta_*^{\otimes})(Q^T\otimes Q^T),$$

где $\eta_*^{\otimes}$ и $\theta_*^{\otimes}$ — диагонали матриц $D^{\otimes}(\eta_*)$ и $D^{\otimes}(\theta_*)$ соответственно.

Предположим, что матрица V_* имеет ранг r, причем для ее собственных значений в разложении (3) имеет место (25). Из предположения о строгой комплементарности X_* и V_* следует, что для соответствующих собственных значений $\eta_*^1, \dots, \eta_*^n$ матрицы X_* $\eta_*^1 = \ldots = \eta_*^r = 0$, $\eta_*^{r+1} > 0, \dots, \eta_*^n > 0$.

Воспользуемся числами n_1 и n_2, определенными ранее. Пусть дополнительно $n_3 = k_\triangle(r)$. В этом случае у вектора $\tilde{\theta}_*^{\otimes}$ первые n_1 компоненты положительные, а остальные — нулевые. Напротив, у вектора $\tilde{\eta}_*^{\otimes} = \mathcal{L}_n\eta_*^{\otimes}\mathcal{D}_n$ заведомо положительными являются только последние n_2 компоненты, а среди первых n_1 найдутся нулевые компоненты в количестве n_3 штук.

Приведем матрицу Λ_* к более удобному для анализа виду. Прежде всего, аналогично тому, как это делалось для матрицы $\tilde{V}_*^{\otimes}$ при доказательстве леммы 2, получаем

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

Отсюда и из (21), (22) следует $\Lambda_* = \mathcal{A}_{vec2}^Q\mathcal{F}_1^{-1}(\tilde{\theta}_*^{\otimes})D(\tilde{\eta}_*^{\otimes})(\mathcal{A}_{vec2}^Q)^T$. Из-за невырожденности точки V_* правая нижняя подматрица B^TB матрицы (27) неособая. Обозначим через $\mathcal{K}$ неособую матрицу $\mathcal{K} = N^TN + \tilde{\Theta}_*^N - N^TB(B^TB)^{-1}B^TN$. Тогда с помощью формулы Фробениуса имеем

$$\mathcal{F}_1^{-1}(\tilde{\theta}_*) = \begin{bmatrix} \mathcal{F}_{11}^{-1} & \mathcal{F}_{12}^{-1} \\ \mathcal{F}_{21}^{-1} & \mathcal{F}_{22}^{-1} \end{bmatrix}, \qquad (29)$$

где

$$\mathcal{F}_{11}^{-1} = \mathcal{K}^{-1}, \quad \mathcal{F}_{21}^{-1} = (\mathcal{F}_{12}^{-1})^T = -\big(B^TB\big)^{-1}B^TN\mathcal{K}^{-1},$$
$$\mathcal{F}_{22}^{-1} = \big(B^TB\big)^{-1} + \big(B^TB\big)^{-1}B^TN\mathcal{K}^{-1}N^TB\big(B^TB\big)^{-1}.$$

После подстановки выражения (29) для матрицы $\mathcal{F}_1^{-1}(\tilde{\theta}_*^{\otimes})$ и проведения соответствующих умножений имеем

$$\Lambda_* = N\mathcal{F}_{11}^{-1}D^N(\tilde{\eta}_*^{\otimes})N^T + B\mathcal{F}_{21}^{-1}D^N(\tilde{\eta}_*^{\otimes})N^T + N\mathcal{F}_{12}^{-1}D^B(\tilde{\eta}_*^{\otimes})B^T + B\mathcal{F}_{22}^{-1}D^B(\tilde{\eta}_*^{\otimes})B^T, \qquad (30)$$

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

Пусть $\mathcal{L}_B$ — пространство столбцов матрицы B, $\mathcal{L}_B^{\perp}$ — его ортогональное дополнение. Тогда матрицу $\mathcal{K}$ можно записать как $\mathcal{K} = \tilde{\Theta}_*^N + N^T\left[I_m - \mathcal{P}\right]N = \tilde{\Theta}_*^N + N^T\mathcal{P}_\perp N$, где $\mathcal{P} =$

$B\left(B^TB\right)^{-1}B^T$ и $\mathcal{P}_\perp = I_m - \mathcal{P}$ — матрицы ортогонального проектирования на подпространства $\mathcal{L}_B$ и $\mathcal{L}_B^\perp$ соответственно. Матрица $\mathcal{P}_\perp$ является симметрической и идемпотентной, т. е. $\mathcal{P}_\perp = \mathcal{P}_\perp^2$. Поэтому $\mathcal{K} = \tilde{\Theta}_*^N + N^T\mathcal{P}_\perp^2 N$. Так как $\tilde{\Theta}_*^N$ — положительно определенная матрица, то согласно формуле Шермана — Моррисона — Вудберри имеем

$$\mathcal{K}^{-1} = \tilde{\Theta}_*^{-N} - \tilde{\Theta}_*^{-N}N^T\mathcal{P}_\perp\left(I + \mathcal{P}_\perp N\tilde{\Theta}_*^{-N}N^T\mathcal{P}_\perp\right)^{-1}\mathcal{P}_\perp N\tilde{\Theta}_*^{-N}, \quad \tilde{\Theta}_*^{-N} = \left(\tilde{\Theta}_*^N\right)^{-1}. \tag{31}$$

Перегруппируем слагаемые в представлении (30) матрицы Λ_* с учетом конкретного вида всех подматриц из $\mathcal{F}_1^{-1}(\tilde{\theta}_*^\otimes)$:

$$\Lambda_* = \mathcal{P}_\perp N\mathcal{K}^{-1}D^N(\tilde{\eta}_*^\otimes)N^T + \left(I - \mathcal{P}_\perp N\mathcal{K}^{-1}N^T\right)B(B^TB)^{-1}D^B(\tilde{\eta}_*^\otimes)B^T.$$

Определим собственные значения матрицы Λ_*. Пусть λ — произвольное собственное значение и y — соответствующий ему собственный вектор матрицы Λ_*. Тогда

$$\Lambda_* y = \lambda y. \tag{32}$$

Предположим сначала, что $B^Ty \neq 0_{n_2}$, и, значит, $\mathcal{P}y \neq 0_m$, $y \notin \mathcal{L}_B^\perp$. Тогда, умножая равенство (32) слева на B^T, приходим к соотношению

$$D^B(\tilde{\eta}_*^\otimes)B^Ty = \lambda B^Ty. \tag{33}$$

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 (33) 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 положительные собственные значения.

Рассмотрим теперь случай, когда $B^Ty = 0_{n_2}$, т. е. $\mathcal{P}y = 0_m$, $y \in \mathcal{L}_B^\perp$. Тогда равенство (32) сводится к $\mathcal{P}_\perp N\mathcal{K}^{-1}D^N(\tilde{\eta}_*^\otimes)N^Ty = \lambda y$. Подставляя выражение (31) для матрицы $\mathcal{K}^{-1}$, получаем $\Omega y = \lambda y$, $\Omega = (I + \mathcal{Y})^{-1}\mathcal{P}_\perp N\mathcal{G}_N N^T\mathcal{P}_\perp$. Здесь $\mathcal{Y} = \mathcal{P}_\perp N\tilde{\Theta}_*^{-N}N^T\mathcal{P}_\perp$, $\mathcal{G}_N = \tilde{\Theta}_*^{-N}D^N(\tilde{\eta}_*^\otimes)$, а также учтено, что $y = \mathcal{P}_\perp y$. Таким образом, y является собственным вектором матрицы Ω, соответствующим тому же самому собственному значению λ.

Так как $I + \mathcal{Y}$ — симметричная положительно определенная матрица, то

$$\Omega = (I + \mathcal{Y})^{-1/2}\,\Omega_1\,(I + \mathcal{Y})^{1/2},$$

где $\Omega_1 = (I + \mathcal{Y})^{-1/2}\,\mathcal{P}_\perp N\mathcal{G}_N N^T\mathcal{P}_\perp\,(I + \mathcal{Y})^{-1/2}$. 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 чисел совпадает с рангом матрицы $\Omega_2 = \mathcal{P}_\perp N\mathcal{G}_N N^T\mathcal{P}_\perp$.

У диагональной матрицы $D^N(\tilde{\theta}_*^\otimes)$ все диагональные элементы положительные, а диагональная матрица $D^N(\tilde{\eta}_*^\otimes)$ имеет $k_\triangle(r)$ нулевых диагональных элементов. Это элементы из $J_\triangle$ стоят на следующих парных номерах:

$$(1,1), (2,1), \ldots, (r,1), (2,2), \ldots, (r,2), \ldots, (r,r). \tag{34}$$

Остальные диагональные элементы в количестве $n_1 - k_\triangle(r)$ штук строго положительны. Таким образом, у диагональной матрицы $\mathcal{G}_N$ все диагональные элементы положительны, за исключением тех, которые расположены на парных номерах из (34). Эти диагональные элементы нулевые.

Так как, по предположению, точка X_* невырожденная, то имеет место неравенство $m \le k_\triangle(n) - k_\triangle(r)$. Кроме того, согласно необходимым и достаточным условиям невырожденности в прямой задаче, ранг матрицы $\mathcal{A}_{vec2}^Q$, из которой удалены столбцы с парными номерами из (34), равен m, т. е. число столбцов в такой усеченной матрице не меньше чем m, и они в совокупности порождают пространство $\mathbb{R}^m$.

Пусть N_1 — подматрица матрицы N, которая получается, если из N удалить столбцы с парными номерами из (34). Согласно вышесказанному столбцы матрицы N_1 совместно со столбцами матрицы B порождают $\mathbb{R}^m$. Поэтому среди столбцов матрицы N_1 найдутся $m-n_2$ линейно независимых столбцов, таких что их проекции на подпространство $\mathcal{L}_B^{\perp}$ порождают все это подпространство.

Пусть, кроме того, $\mathcal{G}_{N_1}$ — диагональная матрица порядка $n_1 - k_{\triangle}(r)$, которая получается из матрицы $\mathcal{G}_N$ путем удаления строк и столбцов с парными номерами (34). Матрица Ω_2 есть матрица Грама для системы векторов, являющихся столбцами матрицы $\mathcal{G}_{N_1}^{1/2} N_1^T \mathcal{P}_{\perp}$. Ранг такой матрицы равен $m-n_2$, поэтому и ранг матрицы Ω_2 равен $m-n_2$. В силу вышесказанного 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 $m-n_2$.

Таким образом, все собственные значения матрицы Λ_* положительны. Поэтому Λ_* — неособая матрица.

Лемма доказана.

Проанализируем матрицу $\mathcal{W}(X,V)$ в точке $[X_*,V_*]$, являющейся решением прямой и двойственной задач. Имеет место следующее утверждение.

Лемма 4. *Пусть для прямой и двойственной задач* (1), (2) *решения* X_* *и* $V_* = V(u_*)$ *невырожденные и строго комплементарные. Тогда матрица* $\mathcal{W}(X_*,V_*)$ *неособая.*

Д о к а з а т е л ь с т в о. Рассмотрим систему линейных уравнений $\mathcal{W}(X_*,V_*)z = 0_{2k_{\triangle}(n)}$ и покажем, что она имеет только тривиальное решение $z = 0_{2k_{\triangle}(n)}$. Запишем данную систему в развернутом виде:

$$\tilde{V}_*^{\otimes} z_1 + \tilde{X}_*^{\otimes} z_2 = 0, \quad \mathcal{A}_{vecs} z_1 = 0, \quad \tilde{\mathcal{A}}_{vecs} z_2 = 0, \tag{35}$$

где $z_1, z_2 \in \mathbb{R}^{k_{\triangle}(n)}$. Из третьего уравнения следует, что вектор z_2 принадлежит нуль-пространству матрицы $\tilde{\mathcal{A}}_{vecs}$, поэтому $z_2 = \mathcal{A}_{vech}^T \mu,\ \mu \in \mathbb{R}^m$. Следовательно, (35) сводится к системе

$$\tilde{V}_*^{\otimes} z_1 + \tilde{X}_*^{\otimes} \mathcal{A}_{vech}^T \mu = 0, \quad \mathcal{A}_{vecs} z_1 = 0. \tag{36}$$

Умножая второе уравнение на $\mathcal{A}_{vech}^T$ и складывая с первым, получим $\Phi(V_*) z_1 = -\tilde{X}_*^{\otimes} \mathcal{A}_{vech}^T \mu$, где матрица $\Phi(V)$ имеет вид (17).

В силу утверждения леммы 2 матрица $\Phi(V_*)$ неособая, поэтому $z_1 = -\Phi^{-1}(V_*) \tilde{X}_*^{\otimes} \mathcal{A}_{vech}^T \mu$. Подставляя данную зависимость во второе уравнение из (36), приходим к системе уравнений относительно переменной μ

$$\mathcal{A}_{vecs} \Phi^{-1}(V_*) \tilde{X}_*^{\otimes} \mathcal{A}_{vech}^T \mu = 0. \tag{37}$$

Матрица данной системы является матрицей $\Lambda(X_*,V_*)$ вида (28). При сделанных предположениях согласно лемме 3 она неособая. Следовательно, система уравнений (37) имеет единственное решение $\mu = 0$. Единственное решение $z = 0$ имеет и исходная система (35). Таким образом, $\mathcal{W}(X_*,V_*)$ — неособая матрица.

Лемма доказана.

З а м е ч а н и е. Из невырожденности матрицы $\mathcal{W}(X_*,V_*)$ и непрерывности $\mathcal{W}(X,V)$ следует, что найдется окрестность точки $[X_*,V_*]$ такая, что для всех $[X,V]$ из этой окрестности матрица $\mathcal{W}(X,V)$ будет неособой. Следовательно, решение системы (8) существует и единственно.

Теорема 1. *Пусть для прямой и двойственной задач* (1), (2) *решения* X_* *и* V_*, *где* $V_* = V(u_*)$, *невырожденные и строго комплементарные. Тогда итерационный процесс* (16) *локально сходится к* $[X_*,V_*]$ *со сверхлинейной скоростью.*

Д о к а з а т е л ь с т в о следует из общих утверждений о сходимости метода Ньютона [13] и леммы 4.

5. Итерационный процесс в пространстве точек (X, u)

Наряду с итерационным процессом (16) можно рассмотреть его вариант в пространстве точек X и u. Учтем явный вид зависимости $V(u) = C - \sum_{i=1}^{m} u^i A_i$. Тогда достаточные условия упрощаются, и система (7) сводится к двум равенствам:

$$\begin{gathered} \mathrm{vech}(X * V(u)) = 0_{k_\triangle(n)}, \\ \mathcal{A}_{vecs}\mathrm{vech}X = b. \end{gathered} \tag{38}$$

Кроме того, если существуют точки $X_* \in \mathcal{S}^n$ и $u_* \in \mathbb{R}^m$, удовлетворяющие (38) такие, что матрицы X_* и $V_* = V(u_*)$ будут положительно полуопределенными, то X_* и $[u_*, V_*]$ — решения соответственно задач (1) и (2). Введем вектор-функцию

$$F_2(\mathrm{vech}\ X, u) = \begin{pmatrix} \mathrm{vech}(X * V(u)) \\ \mathcal{A}_{vecs}\mathrm{vech}X - b \end{pmatrix}$$

и рассмотрим систему

$$F_2(\mathrm{vech}\ X, u) = 0, \tag{39}$$

содержащую $k_\triangle(n) + m$ уравнений и $k_\triangle(n) + m$ неизвестных. Как и раньше будем решать систему (39) методом Ньютона:

$$\begin{bmatrix} \mathrm{vech}\ X_{k+1} \\ u_{k+1} \end{bmatrix} = \begin{bmatrix} \mathrm{vech}\ X_k \\ u_k \end{bmatrix} - \mathcal{W}_2^{-1}(X_k, u_k) \cdot F_2(\mathrm{vech}\ X_k, u_k), \tag{40}$$

где $\mathcal{W}_2(X, u)$ — матрица Якоби вектор-функции $F_2(\mathrm{vech}\ X, u)$.

Найдем явный вид матрицы $W_2(X, u)$. Согласно ее определению,

$$W_2(X, u) = \begin{pmatrix} \dfrac{\partial\, \mathrm{vech}\,(X * V(u))}{\partial(\mathrm{vech}\ X)^T} & \dfrac{\partial\, \mathrm{vech}\,(X * V(u))}{\partial u^T} \\ \mathcal{A}_{vecs} & 0 \end{pmatrix}.$$

Имеем

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

Далее,

$$\frac{\partial\, \mathrm{vech}\,(X * V(u))}{\partial u^T} = \frac{\partial\, \mathrm{vech}\,(X * V)}{\partial(\mathrm{vech}\ V)^T}\frac{\partial\, \mathrm{vech}\ V(u)}{\partial u^T} = -\mathcal{L}_n X^{\otimes}\mathcal{D}_n\mathcal{A}_{vech}^T,$$

поэтому,

$$W_2(X, u) = \begin{pmatrix} \mathcal{L}_n V^{\otimes}(u)\mathcal{D}_n & -\mathcal{L}_n X^{\otimes}\mathcal{D}_n\mathcal{A}_{vech}^T \\ \mathcal{A}_{vecs} & 0 \end{pmatrix}.$$

Лемма 5. *Пусть выполнены предположения леммы* 4. *Тогда матрица* $\mathcal{W}_2(X_*, V_*)$ *неособая.*

Доказательство аналогично доказательству леммы 4.

Теорема 2. *Пусть для прямой и двойственной задач* (1), (2) *решения* X_* *и* V_*, *где* $V_* = V(u_*)$, *невырожденные и строго комплементарные. Тогда итерационный процесс* (40) *локально сходится к* (X_*, u_*) *со сверхлинейной скоростью.*

Доказательство следует из общих утверждений о сходимости метода Ньютона (см. [13]) и леммы 5.

Помимо итерационных ньютоновских процессов (16) и (40), можно также рассмотреть итерационный процесс в пространстве переменных X, u и V.

6. Вычислительный эксперимент

Численные эксперименты проводились с использованием среды MATLAB на задачах малой и средней размерности. Для получения тестов входные данные генерировались “обратным ходом” по следующим правилам:

1) задавались параметры n, m, r, удовлетворяющие необходимому условию невырожденности точек X_* и V_*;

2) случайным образом генерировались положительные числа $\eta_*^1, \dots, \eta_*^r$, $\theta_*^{r+1}, \dots, \theta_*^n$ и ортогональная матрица Q, задающие решения X_* и V_* в задачах (1) и (2);

3) генерировались набор случайных симметричных матриц $A_1, \dots, A_m$ и случайный вектор $u_* = [u_*^1, \dots, u_*^m]$;

4) по X_*, V_* и u_* вычислялись вектор b и матрица C.

Начальные точки X_0, u_0 и V_0 генерировались случайным образом в окрестности решений, причем делалось это таким образом, чтобы получить “равномерное” отклонение от решения по норме всех переменных X, u и V.

В приведенной таблице представлены параметры построенных тестовых задач и количество итераций, потребовавшихся для решения. Результаты, полученные предложенным в данной работе методом (метод 2), сравнивалась с результатами работы прямо-двойственного метода Ньютона, в котором итерации проводились в пространстве переменных X и u (метод 3), а также с результатами работы двойственного метода Ньютона из [14] (метод 1). Все результаты являются усредненными по начальным точкам. Условием остановки являлось неравенство $X_k \bullet V_k < 10^{-5}$.

Параметры задачи							Число итераций		
n	m	r	$\frac{\|\|u_0-u_*\|\|}{\|\|u_*\|\|}$	$\frac{\|\|X_0-X_*\|\|}{\|\|X_*\|\|}$	$\frac{\|\|V_0-V_*\|\|}{\|\|V_*\|\|}$	$X_0 \bullet V_0$	метод 1	метод 2	метод 3
2	1	1	0.15	0.080	0.063	-1.903	1	1	1
5	7	3	0.15	0.145	0.082	2.961	3	3	2
10	25	5	0.15	0.106	0.109	-2.981	2	4	3
15	60	8	0.15	0.091	0.124	1.728	11	9	7
20	95	10	0.15	0.095	0.132	-4.146	23	13	12
25	165	13	0.15	0.103	0.062	-3.544	28	18	20
30	215	15	0.1	0.076	0.049	2.612	46	31	30
35	315	18	0.15	0.145	0.124	1.408	49	39	41
40	415	20	0.07	0.040	0.025	3.272	46	51	49
45	515	23	0.15	0.104	0.127	0.553	90	58	55
50	615	25	0.07	0.050	0.069	-3.796	92	88	80

Как видно из приведенной таблицы, прямо-двойственные методы в среднем показывают более лучшие результаты, чем двойственный метод.

СПИСОК ЛИТЕРАТУРЫ

1. **Еремин И.И.** Теория линейной оптимизации. Екатеринбург: Изд-во “Екатеринбург” , 1999. 312 с.
2. **Еремин И.И., Мазуров В.Д., Астафьев Н.Н.** Несобственные задачи линейного и выпуклого программирования. М.: Наука, 1983. 336 с.
3. Handbook of semidefinite programming / eds. H. Wolkowicz, R. Saigal, L. Vandenberghe. Dordrecht: Kluwer Acad. Publ., 2000. 656 p.
4. **Klerk E. de.** Aspects of semidefinite programming. Interior point algorithms and selected applications. Dordrecht: Kluwer Acad. Publ., 2004. 300 p.
5. **Alizadeh F., Haeberly J.-P.F., Overton M.L.** Primal-dual interior point methods for semidefinite programming. Convergence rates, stability and numerical results // SIAM J. Optim. 1998. Vol. 8, no. 3. P. 746–768.

6. **Nesterov Y.E., Todd M.J.** Primal-dual interior point methods for self-scaled cones // SIAM J. Optim. 1998. Vol. 8, no. 2. P. 324–364.
7. **Monteiro R.D.C.** Primal-dual path-following algorithms for semidefinite programming // SIAM J. Optim. 1997. Vol. 7, no. 3. P. 663–678.
8. **Muramatsu M., Vanderbei R.J.** Primal-dual affine scaling algorithms fails for semidefinite programming // Math. Oper. Res. 1999. Vol. 24, no. 1. P. 149–175.
9. **Евтушенко Ю.Г., Жадан В.Г., Черенков А.П.** Применение метода Ньютона к решению задач линейного программирования // Журн. вычисл. математики и мат. физики. 1995. Т. 35, № 6. С. 850–866.
10. **Magnus J.R., Neudecker H.** The elimination matrix: some lemmas and applications // SIAM J. Alg. Disc. Methods. 1980. Vol. 1, no. 4. P. 422–449.
11. **Арнольд В.И.** О матрицах, зависящих от параметров // Успехи мат. наук. 1971. Т. 26, вып. 2 (158). С. 101–114.
12. **Alizadeh F., Haeberly J.-P.F., Overton M.L.** Complementarity and nondegeneracy in semidefinite programming // Math. Programming. Ser. B. 1997. Vol. 77, no. 2. P. 111–128.
13. **Ортега Дж., Рейнболдт В.** Итерационные методы решения систем нелинейных уравнений со многими неизвестными. М.: Мир, 1975. 560 с.
14. **Жадан В.Г., Орлов А.А.** О сходимости двойственного метода Ньютона для линейной задачи полуопределенного программирования // Изв. Иркут. гос. ун-та. 2011. Т. 4, № 2. С. 75–90. (Математика.)

Поступила 01.02.2013

Жадан Виталий Григорьевич
д-р физ.-мат. наук, профессор
зав. отделом
ВЦ РАН им. А. А. Дородницына
e-mail: zhadan@ccas.ru

Орлов Александр Алексеевич
Московский физико-технический институт (ГУ)
e-mail: sashaorlov@gmail.ru