В. Г. Жадан, Прямой метод Ньютона для линейной задачи полуопределенного программирования, *Тр. ИММ УрО РАН*, 2008, том 14, номер 2, 67–80

ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ

Том 14 № 2 2008

УДК 519.854

ПРЯМОЙ МЕТОД НЬЮТОНА ДЛЯ ЛИНЕЙНОЙ ЗАДАЧИ ПОЛУОПРЕДЕЛЕННОГО ПРОГРАММИРОВАНИЯ[1]

В. Г. Жадан

Рассматривается линейная задача полуопределенного программирования. Для ее решения предлагается прямой метод Ньютона, который является обобщением прямого барьерно-ньютоновского метода для задач линейного программирования. Исследуются свойства метода и доказывается его локальная сходимость.

Введение

Задачи линейной оптимизации являются одними из основных в математическом программировании, и им постоянно уделяется огромное внимание. Большой вклад в теорию и методы решения таких задач внес И. И. Еремин [1]. В настоящей работе рассматривается линейная задача полуопределенного программирования, в которой переменными являются симметричные положительно полуопределенные матрицы. За последние годы был получен ряд значительных результатов, касающихся как существования и единственности решений линейных задач данного вида, так и численных методов нахождения их решений (см., напр., [2–4]).

В [5] для решения линейной задачи полуопределенного программирования (в дальнейшем просто задача полуопределенного программирования) был предложен метод внутренней точки, который является обобщением разработанного ранее в [6] барьерно-проективного метода для задач линейного программирования. Метод является прямым, двойственные переменные в нем выбираются из условия уменьшения невязки ограничений типа равенства по определенному закону. Была доказана локальная сходимость непрерывного и дискретного вариантов метода. Так как данный метод относится к классу градиентных методов, скорость его сходимости оказывается лишь линейной.

В настоящей работе рассматривается метод, который обладает более высокой, чем линейная, скоростью сходимости. В отличие от метода из [5], где по существу с помощью метода простой итерации решается уравнение, описывающее условие дополняющей нежесткости, здесь для этой цели используется метод Ньютона.

В разделе 1 настоящей работы приводится постановка задачи полуопределенного программирования. В разделе 2 рассматривается вопрос о выборе двойственных переменных. Итерационный процесс прямого метода Ньютона строится в разд. 3, в разд. 4 доказывается его локальная сходимость. Через $\mathrm{Diag}(a_1, \dots, a_n)$, или просто $D(a)$, в работе обозначается диагональная матрица с компонентами вектора $a = [a_1, \dots, a_n]$ на диагонали, через $\mathrm{diag}\, A$ — диагональ матрицы A. Символ $\otimes$ используется для обозначения произведения матриц по Кронекеру. Угловые скобки служат для обозначения обычного евклидова скалярного произведения. Вектор-столбец, состоящий из прямой суммы столбцов матрицы M, обозначается $\mathrm{vec}\, M$. Если матрица M — симметричная, то с ней также связывается вектор-столбец $\mathrm{vech}\, M$, в который помещаются последовательно сверху вниз столбцы матрицы M, но не полностью, а только их части, начинающиеся с диагонального элемента.

[1]Работа выполнена при поддержке РФФИ (проект 08-01-00608а), Программы государственной поддержки ведущих научных школ (НШ-5073.2008.1) и Программы фундаментальных исследований ОМН РАН № 3 (проект 3.14).

1. Задача полуопределенного программирования

Пусть $\mathcal{S}^n$ обозначает пространство симметричных матриц порядка n. Пусть, кроме того, $\mathcal{S}^n_+$ и $\mathcal{S}^n_{++}$ — подмножества из $\mathcal{S}^n$, состоящие соответственно из положительно полуопределенных и положительно определенных матриц. Множество $\mathcal{S}^n_+$ является конусом в $\mathcal{S}^n$, множество $\mathcal{S}^n_{++}$ — его внутренностью. Для указания на то, что матрица $M \in \mathcal{S}^n$ положительно полуопределена (положительно определена), будем пользоваться также неравенством $M \succeq 0$ ($M \succ 0$). Конус $\mathcal{S}^n_+$ не является полиэдральным, его размерность равняется так называемому "треугольному числу" $k_\triangle(n) = n(n+1)/2$.

Скалярное (внутреннее) произведение двух матриц L и M из $\mathcal{S}^n$ определяется как след матрицы L^TM и обозначается

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

где l_{ij} и m_{ij} — (ij)-элементы соответственно матриц L и M. Если L и M — две положительно полуопределенные матрицы, то $L \bullet M \geq 0$ и $L \bullet M = 0$ в том и только том случае, когда $LM = ML = 0_{nn}$. Более того, согласно теореме Фейера (о следе), матрица $M \in \mathcal{S}^n$ положительно полуопределена тогда и только тогда, когда $M \bullet L \geq 0$ для всех $L \succeq 0$, т.е. конус $\mathcal{S}^n_+$ является *самосопряженным.*

Рассмотрим задачу полуопределенного программирования следующего вида:

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

где матрицы C, X и A_i, $1 \leq i \leq m$, принадлежат множеству $\mathcal{S}^n$. Если от всех матриц дополнительно потребовать, чтобы они были диагональными, то (1) переходит в обычную задачу линейного программирования.

Двойственной к (1) является задача

$$\begin{array}{rcl} \max \ \langle b, u\rangle, & & \\ \sum\limits_{i=1}^{m} u^iA_i + V & = & C, \\ V & \succeq & 0, \end{array} \tag{2}$$

в которой $b = (b^1, \ldots, b^m)^T \in \mathbb{R}^m$, $V \in \mathcal{S}^n$. Предполагается, что задача (1) имеет решение и что матрицы A_i, $1 \leq i \leq m$, линейно независимы.

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

$$\begin{array}{rcl} \mathcal{F}_P & = & \left\{X \in \mathcal{S}^n_+ : \ A_i \bullet X = b^i, \quad i = 1, \ldots, m\right\}, \\ \mathcal{F}_D & = & \left\{u \in \mathbb{R}^m : \ C - \sum\limits_{i=1}^{m} u^iA_i \succeq 0\right\}. \end{array}$$

Обозначим также

$$f_* = \inf_{X \in \mathcal{F}_P} C \bullet X, \qquad f^* = \sup_{u \in \mathcal{F}_D} \langle b, u\rangle.$$

Для любых $X \in \mathcal{F}_P$ и $u \in \mathcal{F}_D$ выполняется неравенство *слабой двойственности*: $\langle b, u\rangle \leq C \bullet X$. Если $f_* = f^*$, то говорят, что задачи (1) и (2) находятся в *совершенной двойственности* (однако при этом одно из значений f_* или f^* может не достигаться). В случае, когда

$\langle b, u\rangle = C \bullet X$ для некоторых $X \in \mathcal{F}_P$ и $u \in \mathcal{F}_D$, имеет место *строгая двойственность*. Условия регулярности ограничений Слейтера позволяют гарантировать существование строгой двойственности (см., напр., [4]).

Теорема 1. *Пусть в задачах* (1) *и* (2) *выполнено условие регулярности ограничений Слейтера, т.е. множества*

$$\mathcal{F}_P^0 = \left\{X \in \mathcal{S}^n_{++} : \; A_i \bullet X = b^i, \quad i = 1, \dots, m\right\},$$

$$\mathcal{F}_D^0 = \left\{u \in \mathbb{R}^m : \; C - \sum_{i=1}^m u^i A_i \succ 0\right\}$$

непусты. Тогда $f_* = f^*$ и существуют решения обеих задач.

Если X_* и $[u_*, V_*]$ — оптимальные решения соответственно задач (1) и (2), то $X_* \bullet V_* = 0$. Но для матриц X_* и $V_* \in \mathcal{S}^n_+$ данное равенство выполняется тогда и только тогда, когда $X_* V_* = V_* X_* = 0_{nn}$. Отсюда следует, что оптимальные матрицы X_* и V_* коммутируют. Поэтому найдется ортогональная матрица Q такая, что

$$X_* = Q \operatorname{Diag}(\eta_*) Q^T, \qquad V_* = Q \operatorname{Diag}(\theta_*) Q^T, \tag{3}$$

где $\eta_* = [\eta_*^1, \dots, \eta_*^n]$ и $\theta_* = [\theta_*^1, \dots, \theta_*^n]$ — собственные значения матриц X_* и V_*, соответственно. Для самих собственных значений η_*^i и θ_*^i выполняется *условие дополнительности*: $\eta_*^i \theta_*^i = 0$, $1 \le i \le n$. *Условие строгой дополнительности* означает, что для каждого $1 \le i \le n$ одно из значений η_*^i или θ_*^i строго положительно. В этом случае решения X_* и V_* называют *строго комплементарными.*

В отличие от задач линейного программирования, для которых если существуют оптимальные решения, то существуют и строго комплементарные оптимальные решения, для задач полуопределенного программирования данное свойство не всегда выполняется.

2. Выбор двойственных переменных

Пусть X и V — симметричные матрицы. Обозначим через $X * V$ их симметризованное произведение

$$X * V = \frac{1}{2}\left(XV + VX\right).$$

Утверждение 1. *Для матриц $X \in \mathcal{S}^n_+$ и $V \in \mathcal{S}^n_+$ равенство $X * V = 0_{nn}$ возможно в том и только том случае, когда $XV = VX = 0_{nn}$.*

Доказательство. Достаточность очевидна. Покажем необходимость.
Пусть $X * V = 0$. 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 $V \succeq 0$, отсюда заключаем, что $XV = VX = 0_{nn}$. $\square$

Возьмем теперь в качестве V матрицу

$$V = V(u) = C - \sum_{i=1}^m u^i A_i, \tag{4}$$

где $u = [u^1, \dots, u^m]^T$ — вектор двойственных переменных. Выберем вектор u таким образом, чтобы выполнялось условие

$$A_i \bullet (X * V(u)) = \tau\left(A_i \bullet X - b^i\right), \qquad 1 \le i \le m, \tag{5}$$

где $\tau > 0$ — некоторый параметр. После подстановки (4) в (5) приходим к системе линейных алгебраических уравнений относительно вектора u:

$$A_i \bullet \left[X * \left(C - \sum_{j=1}^{m} u^j A_j\right)\right] = \tau \left(A_i \bullet X - b^i\right), \qquad 1 \le i \le m.$$

Данная система может быть переписана следующим образом:

$$\sum_{j=1}^{m} \left[A_i \bullet (X * \mathcal{A}_j)\right] u^j = A_i \bullet (X * C) + \tau \left(b^i - A_i \bullet X\right), \quad 1 \le i \le m. \tag{6}$$

Пусть $\mathcal{A}$ — $(mn \times n)$-матрица, составленная из матриц A_i, $1 \le i \le m$,

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

Пусть, кроме того,

$$\Gamma(X) = \mathcal{A} \bullet (X * \mathcal{A}^T).$$

Матрица $\Gamma(X)$ — квадратная симметричная порядка m, ее (i,j)-й элемент равняется

$$\left(\mathcal{A} \bullet (X * \mathcal{A}^T)\right)_{ij} = A_i \bullet (X * A_j) = A_j \bullet (X * A_i) = \left(\mathcal{A} \bullet (X * \mathcal{A}^T)\right)_{ji}.$$

Обозначим также через $\mathcal{A} \bullet X$ и $\mathcal{A} \bullet (X * C)$ m-мерные векторы. Их i-е элементы равны соответственно

$$(\mathcal{A} \bullet X)^i = A_i \bullet X, \qquad (\mathcal{A} \bullet (X * C))^i = A_i \bullet (X * C).$$

В этих обозначениях систему (6) можно представить в следующем матричном виде:

$$\Gamma(X) u = \mathcal{A} \bullet (X * C) + \tau \left(b - \mathcal{A} \bullet X\right).$$

Если у матрицы $\Gamma(X)$ существует обратная, то отсюда находим

$$u = u(X) = \Gamma^{-1}(X) \left[\mathcal{A} \bullet (X * C) + \tau \left(b - \mathcal{A} \bullet X\right)\right]. \tag{7}$$

Поэтому

$$V(X) = V(u(X)) = C - \mathcal{A}^T \left\{\left[\Gamma^{-1}(X) \left(\mathcal{A} \bullet (X * C) + \tau \left(b - \mathcal{A} \bullet X\right)\right)\right] \otimes I_n\right\}. \tag{8}$$

Рассмотрим условия, при которых матрица $\Gamma(X)$ оказывается невырожденной. Пусть A, B и X — произвольные матрицы из $\mathcal{S}^n$, причем матрица X положительно определена. Положим $X^{1/2}$ — корень квадратный из X, т.е. $X = X^{1/2} X^{1/2}$. Тогда, учитывая симметричность всех матриц,

$$\begin{aligned} A \bullet (X * B) &= A \bullet ((X^{1/2} X^{1/2}) * B) \\ &= \operatorname{tr}\left(A X^{1/2} X^{1/2} B + A B X^{1/2} X^{1/2}\right) / 2 \\ &= \left[\operatorname{tr}\left(X^{1/2} B A X^{1/2}\right) + \operatorname{tr}\left(X^{1/2} A B X^{1/2}\right)\right] / 2 \\ &= \left[\operatorname{tr}\left(\overline{B}^T \overline{A}\right) + \operatorname{tr}\left(\overline{A}^T \overline{B}\right)\right] / 2, \end{aligned}$$

где введены обозначения: $\overline{A} = A X^{1/2}$, $\overline{B} = B X^{1/2}$. Поэтому $A \bullet (X * B) = \operatorname{tr}\left(\overline{A}^T * \overline{B}\right)$ или $A \bullet (X * B) = (\overline{A} \bullet \overline{B} + \overline{B} \bullet \overline{A})/2$. Поскольку $\overline{A} \bullet \overline{B} = \overline{B} \bullet \overline{A}$, то

$$A \bullet (X * B) = \overline{A} \bullet \overline{B} = \overline{B} \bullet \overline{A}.$$

Согласно сказанному выше, (i,j)-й элемент матрицы $\Gamma(X)$ представим в виде $\overline{A}_i \bullet \overline{A}_j$, поэтому Γ является матрицей Грама. Для существования обратной к ней матрицы необходимо и достаточно, чтобы матрицы $\overline{A}_i$, $1 \le i \le m$, были линейно независимы.

Обратимся к конусу положительно полуопределенных матриц $\mathcal{S}^n_+$. Положим

$$\mathcal{M}_r = \{X \in \mathcal{S}^n : \ \mathrm{rank}\ X = r\}, \qquad \mathcal{M}^+_r = \mathcal{S}^n_+ \bigcap \mathcal{M}_r.$$

Тогда границы конуса $\mathcal{S}^n_+$ и его внутренность могут быть представлены как

$$\partial\mathcal{S}^n_+ = \mathcal{M}^+_0 \bigcup \cdots \bigcup \mathcal{M}^+_{n-1}, \qquad \mathrm{int}\,\mathcal{S}^n_+ = \mathcal{M}^+_n.$$

Пусть X — произвольная допустимая матрица из $\mathcal{F}_P$ и $\mathrm{rank}\ X = r$. Предположим также, что для X имеет место разложение

$$X = Q\,\mathrm{Diag}\ (\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)\, Q^T, \tag{9}$$

где Q — ортогональная матрица. Касательное пространство к $\mathcal{M}_r$ в X определяется следующим образом [7]:

$$\mathcal{T}_X = \left\{ Q \begin{bmatrix} G & H \\ H^T & 0 \end{bmatrix} Q^T : \ G \in \mathcal{S}^r \right\}.$$

Обозначим также

$$\mathcal{N}_{\mathcal{A}} = \{Y \in \mathcal{S}^n : \ A_i \bullet Y = 0, \ 1 \le i \le m\}$$

и приведем определение невырожденной точки для прямой задачи (1), следуя [8].

О п р е д е л е н и е 1. Точка $X \in \mathcal{F}_P$ называется *невырожденной*, если $\mathcal{T}_X + \mathcal{N}_{\mathcal{A}} = \mathcal{S}^n$.

Можно привести характеристику того, что $X \in \mathcal{F}_P$ является невырожденной точкой, используя представление (9). А именно, пусть Q_1 и Q_2 — подматрицы матрицы Q, состоящие соответственно из первых r и последующих $n-r$ столбцов Q. Тогда X будет будет невырожденной точкой $\mathcal{F}_P$ в том и только том случае, когда матрицы

$$B_i = \begin{bmatrix} Q_1^T A_i Q_1 & Q_1^T A_i Q_2 \\ Q_2^T A_i Q_1 & 0 \end{bmatrix}, \qquad 1 \le i \le m, \tag{10}$$

линейно независимы. Если $X \in \mathcal{F}_P$ — невырожденная точка, причем $\mathrm{rank}\ X = r$, то необходимо, чтобы выполнялось неравенство $m \le k_\triangle(n) - k_\triangle(n-r)$.

Утверждение 2. *Пусть точка $X \in \mathcal{F}_P$ является невырожденной, тогда матрица* $\Gamma(X)$ *неособая.*

Д о к а з а т е л ь с т в о. Как отмечено выше, достаточно показать, что матрицы $\overline{A}_i = A_i X^{1/2}$, $1 \le i \le m$, линейно независимы.

Пусть, от противного, матрицы $\overline{A}_i$ линейно зависимы. Тогда найдутся числа $c_1, \ldots, c_m$, не все равные нулю и такие, что

$$\sum_{i=1}^m c_i A_i X^{1/2} = \Big(\sum_{i=1}^m c_i A_i\Big) X^{1/2} = 0_{nn}. \tag{11}$$

Предположим, что матрица X имеет ранг r, и пусть λ — вектор, составленный из собственных значений матрицы X, первые r из которых строго положительны, а последующие $n-r$ равны нулю. Пусть, кроме того, для X имеет место представление (9). Тогда $X^{1/2} = QD^{1/2}(\lambda)Q^T$.

Равенство (11) может быть переписано в виде

$$Q\Big(\sum_{i=1}^m c_i Q^T A_i Q\Big) D^{1/2}(\lambda) Q^T = 0_{nn}. \tag{12}$$

Так как Q — ортогональная матрица, то (12) выполняется в том и только том случае, когда

$$\Big(\sum_{i=1}^{m} c_i Q^T A_i Q\Big) D^{1/2}(\lambda) = 0_{nn}.$$

Отсюда следует, что существуют такие $c_1, \ldots, c_m$, не равные нулю одновременно, для которых левая $(n \times r)$-подматрица матрицы

$$\overline{Q} = \sum_{i=1}^{m} c_i Q^T A_i Q \tag{13}$$

является нулевой.

Используя представление (9), разобьем матрицы $Q^T A_i Q$ на блоки:

$$Q^T A_i Q = \begin{bmatrix} Q_1^T A_i Q_1 & Q_1^T A_i Q_2 \\ Q_2^T A_i Q_1 & Q_2^T A_i Q_2 \end{bmatrix}, \qquad 1 \le i \le m.$$

В силу невырожденности точки X и (10) матрицы

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

линейно независимы. Но тогда линейно независимыми должны быть $(n \times r)$-матрицы:

$$\begin{bmatrix} Q_1^T A_i Q_1 \\ Q_2^T A_i Q_1 \end{bmatrix}, \qquad 1 \le i \le m.$$

Поэтому не существует коэффициентов $c_1, \ldots, c_m$, обращающих левую $(n \times r)$-подматрицу матрицы (13) в нулевую матрицу. $\square$

Задачу (1) назовем *невырожденной*, если все точки X, принадлежащие множеству $\mathcal{F}_P$, не вырождены. Тогда в силу непрерывности существует некоторая окрестность множества $\mathcal{F}_P$, для точек из которой матрица $\Gamma(X)$ будет неособая и, следовательно, вектор $V(X)$ полностью определен. Ниже везде предполагается, что задача (1) является невырожденной.

3. Итерационный процесс

В [5] для решения задачи (1) предлагалось отыскивать предельные при $t \to +\infty$ точки решения задачи Коши

$$\frac{dX}{dt} = -X * V(X), \qquad X_0 = X(0) \in \mathcal{S}^n_{++}. \tag{14}$$

Матрица $X \in \mathcal{F}_P$ такая, что $X * V(X) = 0_{nn}$, является стационарной точкой для этой системы. В силу (5) для ограничений типа равенства

$$g^i(X) = b^i - A_i \bullet X = 0, \qquad 1 \le i \le m, \tag{15}$$

выполняется

$$\frac{dg^i(X)}{dt} = -\Big\langle \operatorname{vec} A_i,\, \operatorname{vec}\Big(\frac{dX}{dt}\Big)\Big\rangle = A_i \bullet (X * V(X)) = -\tau g^i(X), \qquad 1 \le i \le m. \tag{16}$$

Поэтому $g^i(X) \to 0$ при $t \to +\infty$ для каждого $1 \le i \le m$. Итерационный процесс для отыскания стационарных точек строился путем интегрирования системы (14) по схеме Эйлера.

В настоящей работе вместо (14) предлагается искать решение матричного уравнения

$$F(X) = X * V(X) = 0_{nn} \tag{17}$$

методом Ньютона. Как следует из (5), любое решение данного уравнения удовлетворяет одновременно равенствам (15). Если найденное решение X будет таким, что $X \succeq 0$ и $V = V(X) \succeq 0$, то тем самым будет получено решение задачи (1).

Утверждение 3. *Пусть для задач* (1) *и* (2) *имеет место строгая двойственность. Пусть, кроме того, X_* — невырожденное решение задачи* (1), $[u_*, V_*]$ — *решение двойственной задачи* (2). *Тогда X_* удовлетворяет уравнению* (17) *и* $V_* = V(X_*)$.

Д о к а з а т е л ь с т в о. Из условия строгой двойственности следует, что для матриц X_* и

$$V_* = C - \sum_{i=1}^{m} u_*^i A_i$$

выполняется равенство $X_* \bullet V_* = 0$. Поэтому X_* и V_* коммутируют и $X_* V_* = V_* X_* = 0_{nn}$. Таким образом, выполняется равенство

$$X_* \Big(C - \sum_{i=1}^{m} u_*^i A_i \Big) = 0_{nn}. \tag{18}$$

С другой стороны, вектор $u = u(X_*)$ должен удовлетворять системе уравнений

$$\mathcal{A} \bullet \Big[X_* * \Big(C - \sum_{i=1}^{m} u^i A_i \Big) \Big] = 0. \tag{19}$$

Так как по предположению X_* — невырожденное решение задачи (1), то, согласно утверждению 2, данная система имеет единственное решение. Но точка $u = u_*$ в силу (18) удовлетворяет (19). Поэтому $u(X_*) = u_*$. Отсюда следует, что для матрицы $V(X_*)$, в которой $u(X_*) = u_*$, выполнено $V(X_*) * X_* = 0_{nn}$. Таким образом, X_* — решение уравнения (17). $\square$

Обозначим через $F_X(X)$ матрицу Якоби матричной функции $F(X) = X * V(X)$ в точке X. Согласно [9], ее удобно определять следующим образом:

$$F_X(X) = \frac{\partial \operatorname{vec} F(X)}{\partial (\operatorname{vec} X)^T}.$$

Матрица $F_X(X)$ является квадратной порядка n^2. Удобно нумеровать ее строки и столбцы не обычными цифрами из натурального ряда, а последовательностью пар индексов (i, j), соответствующих индексам элементов из vec X, т.е. набором

$$J = \{(1,1), \ldots, (n,1), (1,2), \ldots, (n,2), \ldots, (1,n), \ldots, (n,n)\}. \tag{20}$$

Тогда в матрице $F_X(X)$ в позиции $(i,j) \times (p,q)$ стоит производная $\partial F_{ij} / \partial X_{pq}$.

Так как $F(X)$ — симметричная матричная функция от симметричного матричного аргумента, то среди столбцов и строк матрицы $F_X(X)$ имеются одинаковые. Если взять матрицу Якоби

$$F_X^{\triangle}(X) = \frac{\partial \operatorname{vech} F(X)}{\partial (\operatorname{vech} X)^T},$$

то она связана с $F_X(X)$ соотношением [10]:

$$F_X^{\triangle}(X) = \mathcal{L}_n F_X(X) \mathcal{D}_n. \tag{21}$$

Здесь $\mathcal{L}_n$ и $\mathcal{D}_n$ — элиминационная и дуплицирующая матрицы, соответственно.

Матрица $\mathcal{L}_n$ имеет размер $k_\triangle(n) \times n^2$, матрица $\mathcal{D}_n$ — размер $n^2 \times k_\triangle(n)$. Матрица $\mathcal{D}_n$ для квадратной симметричной матрицы X порядка n осуществляет преобразование $\mathcal{D}_n \operatorname{vech} X = \operatorname{vec} X$. Матрица $\mathcal{L}_n$, напротив, действует на произвольную квадратную матрицу X порядка n таким образом, что $\mathcal{L}_n \operatorname{vec} X = \operatorname{vech} X$.

Обратимся теперь к набору из $k_\triangle(n)$ пар индексов, которые имеют вид, аналогичный (20), но соответствующий нумерации вектора vech X, т.е. набору

$$J_\triangle = \{(1,1), (2,1), \ldots, (n,1), (2,2), \ldots, (n,2), \ldots, (n,n-1), (n,n)\}.$$

Тогда единичную матрицу порядка $k_\triangle(n)$ можно представить как совокупность единичных ортов размера $k_\triangle(n)$, в которых единица располагается на месте (i,j)-го элемента, $(i,j) \in J_\triangle$. Такие единичные орты будем обозначать $e_{(i,j)}$. Пусть, кроме того, e_i обозначает i-й единичный орт матрицы I_n.

Матрица $\mathcal{L}_n$ определяется единственным образом и имеет полный ранг. Ее явное представление следующее:

$$\mathcal{L}_n = \sum_{(i,j)\in J_\triangle} e_{(i,j)} \otimes e_j^T \otimes e_i^T.$$

Таким образом, матрица $\mathcal{L}_n$ есть $(0,1)$-матрица и полуортогональная, т.е. $\mathcal{L}_n\mathcal{L}_n^T = I_{k_\triangle(n)}$. Общее число единиц в ней равно $k_\triangle(n)$, каждый ее столбец содержит ровно одну единицу, каждая строка — не более одной единицы.

Матрица $\mathcal{D}_n$ также имеет полный ранг, в явном виде она записывается следующим образом:

$$\mathcal{D}_n = \Big[\sum_{(i,j)\in J_\triangle} e_{(i,j)} \left(\mathrm{vec}\,(E_{ij})\right)^T \Big]^T,$$

где $E_{ij} = e_i e_j^T + e_j e_i^T$, если $i \neq j$, и $E_{ii} = e_i e_i^T$. Согласно определению матриц $\mathcal{L}_n$ и $\mathcal{D}_n$ имеет место равенство $\mathcal{L}_n\mathcal{D}_n = I_{k_\triangle(n)}$.

Матричное уравнение (17) эквивалентно векторному уравнению

$$\mathrm{vech}\, F(X) = 0_{k_\triangle(n)}. \tag{22}$$

Применим теперь метод Ньютона для решения уравнения (22). Ньютоновское направление $\mathrm{vech}\,(X_{k+1} - X_k)$ на k-й итерации удовлетворяет условию

$$\mathrm{vech}\, F(X_k) + F_X^\triangle(X_k)\, \mathrm{vech}\,(X_{k+1} - X_k) = 0_{k_\triangle(n)}.$$

Откуда, если матрица $F_X^\triangle(X_k)$ невырожденная, получаем

$$\mathrm{vech}\, X_{k+1} = \mathrm{vech}\, X_k - [F_X^\triangle(X_k)]^{-1} \mathrm{vech}\, F(X_k)$$

или, с учетом равенства (21),

$$\mathrm{vech}\, X_{k+1} = \mathrm{vech}\, X_k - [\mathcal{L}_n F_X(X_k)\mathcal{D}_n]^{-1} \mathrm{vech}\, F(X_k). \tag{23}$$

Все матрицы X_k, X_{k+1} и $F(X_k)$ должны быть симметричными. Так как $\mathcal{L}_n\mathcal{D}_n = I_{k_\triangle(n)}$, то после умножения левой и правой части (23) на матрицу $\mathcal{D}_n$ получаем результат

$$\mathrm{vec}\, X_{k+1} = \mathrm{vec}\, X_k - \mathcal{D}_n [\mathcal{L}_n F_X(X_k)\mathcal{D}_n]^{-1} \mathcal{L}_n \mathrm{vec}\, F(X_k). \tag{24}$$

Пусть $\mathcal{W}$ и X — симметричные квадратные матрицы соответственно порядков n^2 и n. Обозначим через $\mathcal{W}_{(i,j)}$ строку матрицы $\mathcal{W}$ с номером $(i,j) \in J$. Пусть $W_{(i,j)}$ — соответствующая этой строке симметричная квадратная матрица порядка n, т.е. $\mathcal{W}_{(i,j)} = \mathrm{vec}\, W_{(i,j)}$. Квадратную матрицу порядка n такую, что ее (i,j)-й элемент есть $W_{(i,j)} \bullet X$, будем обозначать через $\mathcal{W} \bullet X$. Итерационный процесс (24) с использованием введенных обозначений может быть записан в следующей матричной форме:

$$X_{k+1} = X_k - \{\mathcal{D}_n [\mathcal{L}_n F_X(X_k)\mathcal{D}_n]^{-1} \mathcal{L}_n\} \bullet F(X_k). \tag{25}$$

Найдем явный вид матрицы Якоби $F_X(X)$. Для упрощения записи положим

$$X^\otimes = \frac{1}{2}\left[(X \otimes I_n) + (I_n \otimes X)\right], \qquad V^\otimes = \frac{1}{2}\left[(V \otimes I_n) + (I_n \otimes V)\right].$$

Имеет место следующий результат.

Утверждение 4. *Матрица Якоби $F_X(X)$ для матричной функции $F(X)$ имеет вид*

$$F_X(X) = V^{\otimes}(X) + X^{\otimes}V_X(X). \tag{26}$$

Д о к а з а т е л ь с т в о. Дифференциал произведения двух матричных функций $XV(X)$ равен

$$d(XV(X)) = (dX)V(X) + XdV(X).$$

В векторной форме это соотношение запишется в виде

$$d\,(\mathrm{vec}\,(XV(X)) = \mathrm{vec}\,(d(XV(X)) = \mathrm{vec}\,((dX)V(X)) + \mathrm{vec}\,(XdV(X)).$$

Тогда на основании формулы $\mathrm{vec}\,(ABC) = (C^T \otimes A)\,\mathrm{vec}\,B$, имеющей место для любых матриц A, B и C, для которых определено их произведение ABC, получаем

$$d\,(\mathrm{vec}\,(XV(X)) = (V^T(X) \otimes I_n)\,d\ \mathrm{vec}\ X + (I_n \otimes X)\,\mathrm{vec}\,(dV(X)). \tag{27}$$

Но по определению первого дифференциала

$$\mathrm{vec}\,(d(XV(X)) = (XV(X))_X\,d\,\mathrm{vec}\ X.$$

Подставляя данное равенство в (27), приходим к

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

Таким образом,

$$(XV(X))_X = (V(X) \otimes I_n) + (I_n \otimes X)V_X(X). \tag{28}$$

Здесь учтено, что матрица $V(X)$ симметричная.

Аналогичным образом может быть показано, что

$$(V(X)X)_X = (X \otimes I_n)V_X(X) + (I_n \otimes V(X)). \tag{29}$$

Из (28) и (29) следует (26). □

Согласно (4)

$$V_X(X) = -\sum_{i=1}^{m} (\mathrm{vec}\ A_i)\,u_X^i(X). \tag{30}$$

Каждый правый сомножитель $u_X^i(X)$ в (30) является n^2-мерной вектор-строкой, (p,q)-й элемент которой есть частная производная $\partial u^i/\partial X_{pq}$. Здесь $(p,q) \in J$. Пусть $u_X(X)$ — матрица Якоби вектор-функции $u(X)$ размером $m \times n^2$, составленная из строк $u_X^i(X)$, $1 \le i \le m$. Пусть, кроме того, $\mathcal{A}_{\mathrm{vec}}$ есть $m \times n^2$ матрица, i-й строкой которой является вектор $\mathrm{vec}\,A_i$. Тогда матрицу (30) можно записать в виде

$$V_X(X) = -\mathcal{A}_{\mathrm{vec}}^T\,u_X(X). \tag{31}$$

Таким образом, вычисление $V_X(X)$ сводится к вычислению матрицы Якоби $u_X(X)$.

Утверждение 5. *Пусть точка X является невырожденной. Тогда*

$$u_X(X) = (\mathcal{A}_{\mathrm{vec}}\,X^{\otimes}\mathcal{A}_{\mathrm{vec}}^T)^{-1}(\mathcal{A}_{\mathrm{vec}}V^{\otimes} - \tau\mathcal{A}_{\mathrm{vec}}). \tag{32}$$

Д о к а з а т е л ь с т в о. Дифференцируя равенства (5), получаем

$$d\,[A_i \bullet (X * V(X))] = \tau\, d\,(A_i \bullet X)\,, \qquad 1 \le i \le m, \tag{33}$$

или, учитывая симметричность матриц A_i, X и V,

$$\frac{1}{2}\left\{d\ \mathrm{tr}\,(A_i X V) + d\ \mathrm{tr}\,(A_i V X)\right\} = \tau\, d\ \mathrm{tr}\,(A_i X), \qquad 1 \le i \le m.$$

Согласно правилам дифференцирования

$$d\ \mathrm{tr}\,(A_i X V) = \mathrm{tr}\ d\,(A_i X V) = \mathrm{tr}\,(A_i\, dX\, V) + \mathrm{tr}\,(A_i X\, dV). \tag{34}$$

Но

$$\begin{aligned} \mathrm{tr}\,(A_i\, dX\, V) &= (\mathrm{vec}\ A_i)^T\, \mathrm{vec}\,(d\,X\,V) \\ &= (\mathrm{vec}\ A_i)^T (V^T \otimes I_n)\, d\ \mathrm{vec}\ X \\ &= (\mathrm{vec}\ A_i)^T (V \otimes I_n)\, d\ \mathrm{vec}\ X. \end{aligned} \tag{35}$$

Для второго слагаемого в правой части (34) получаем таким же образом

$$\mathrm{tr}\,(A_i\, X\, dV) = (\mathrm{vec}\ A_i)^T (I_n \otimes X) V_X(X)\, d\ \mathrm{vec}\ X. \tag{36}$$

Если обозначить через $\phi(X)$ скалярную функцию от матричного аргумента $\phi(X) = \mathrm{tr}\ (A_i X V(X))$, то из (34)–(36) приходим к ее матрице Якоби:

$$\phi_X(X) = (\mathrm{vec}\ A_i)^T \left[(V \otimes I_n) + (I_n \otimes X) V_X(X)\right].$$

Аналогично, если взять функцию $\psi(X) = \mathrm{tr}\,(A_i V X)$, то ее матрица Якоби имеет вид

$$\psi_X(X) = (\mathrm{vec}\ A_i)^T \left[(I_n \otimes V) + (X \otimes I_n)\, V_X(X)\right].$$

Для правой части равенства (33) находим

$$d\ \mathrm{tr}\,(A_i X) = \mathrm{tr}\ d\,(A_i X) = \mathrm{tr}\,(A_i dX) = (\mathrm{vec}\ A_i)^T d\ \mathrm{vec}\ X,$$

откуда следует, что матрица Якоби для скалярной функции $\varphi(X) = \mathrm{tr}\,(A_i X)$ есть просто $\varphi_X(X) = (\mathrm{vec}\ A_i)^T$.

Приравнивая теперь матрицы Якоби от левой и правой части (5), получаем

$$(\mathrm{vec}\ A_i)^T \left[(V^{\otimes} + X^{\otimes} V_X(X)\right] = \tau\, (\mathrm{vec}\ A_i)^T,$$

где $1 \le i \le m$. Объединим все эти равенства в одно с помощью матрицы $\mathcal{A}_{\mathrm{vec}}$. Имеем

$$\mathcal{A}_{\mathrm{vec}} \left[V^{\otimes} + X^{\otimes} V_X(X)\right] = \tau \mathcal{A}_{\mathrm{vec}}. \tag{37}$$

Согласно (30) $V_X(X) = -\mathcal{A}_{\mathrm{vec}}^T\, u_X(X)$. Подставляя данное выражение в (37), получаем

$$\Gamma(X)\, u_X(X) = \mathcal{A}_{\mathrm{vec}} V^{\otimes} - \tau\, \mathcal{A}_{\mathrm{vec}}, \tag{38}$$

где через $\Gamma(X)$ обозначена матрица $\Gamma(X) = \mathcal{A}_{\mathrm{vec}}\, X^{\otimes} \mathcal{A}_{\mathrm{vec}}^T$. Если матрица $\Gamma(X)$ неособая, то из (38) приходим к (32). □

В формуле (38) матрица $\Gamma(X)$ квадратная порядка m. Она может быть записана в виде (7). Действительно, так как

$$(I_n \otimes X)\, \mathrm{vec}\ A_i = \mathrm{vec}\,(X A_i), \qquad (X \otimes I_n)\, \mathrm{vec}\ A_i = \mathrm{vec}\,(A_i X),$$

то ее (i, j)-й элемент равен

$$\Gamma_{i,j} = (\mathrm{vec}\ A_i)^T\, \mathrm{vec}\,(X * A_j) = A_i \bullet (X * A_j).$$

Отсюда следует (7). Поэтому в невырожденной допустимой точке $X \in \mathcal{F}_P$ согласно утверждению 2 матрица $\Gamma(X)$ должна быть неособой.

Для сокращения записи введем обозначение

$$\mathcal{P}(X^{\otimes}) = X^{\otimes}\mathcal{A}_{\mathrm{vec}}^T \left(\mathcal{A}_{\mathrm{vec}}\, X^{\otimes}\mathcal{A}_{\mathrm{vec}}^T\right)^{-1}\mathcal{A}_{\mathrm{vec}}.$$

Тогда на основании утверждений 4 и 5 и формулы (31) приходим к выводу, что матрица Якоби функции $F_X(X)$ имеет вид

$$F_X(X) = \left[I_{n^2} - \mathcal{P}(X^{\otimes})\right] V^{\otimes} + \tau\, \mathcal{P}(X^{\otimes}). \tag{39}$$

Объединяя формулы (8), (25) и (39), приходим к следующему итерационному процессу метода Ньютона:

$$X_{k+1} = X_k - \left[\mathcal{D}_n \Lambda^{-1}(X_k)\mathcal{L}_n\right] \bullet F(X_k), \tag{40}$$

где $X_0 \in S^n$ и

$$\Lambda(X) = \mathcal{L}_n \left\{\left[I_{n^2} - \mathcal{P}(X^{\otimes})\right] V^{\otimes} + \tau\, \mathcal{P}(X^{\otimes})\right\}\mathcal{D}_n, \tag{41}$$

$$F(X) = X * \Big\{C - \mathcal{A}^T\big\{\big[\left(\mathcal{A}_{\mathrm{vec}}\, X^{\otimes}\mathcal{A}_{\mathrm{vec}}^T\right)^{-1}\left(\mathcal{A}\bullet(X*C) + \tau\,(b - \mathcal{A}\bullet X)\right)\big] \otimes I_n\big\}\Big\}.$$

Непрерывный вариант метода Ньютона описывается следующей системой обыкновенных дифференциальных уравнений:

$$\frac{dX}{dt} = -\gamma \left[\mathcal{D}_n \Lambda^{-1}(X)\mathcal{L}_n\right] \bullet F(X). \tag{42}$$

Здесь $X(0) \in S^n$, $\gamma > 0$ — некоторый параметр.

4. Свойства метода и локальная сходимость

Введем в рассмотрение матрицу $\mathcal{N}_n = \left(I_{k_\triangle(n)} + \mathcal{K}_n\right)/2$, где $\mathcal{K}_n$ — коммутационная матрица. Для каждой квадратной матрицы M порядка n матрица $\mathcal{K}_n$ совершает преобразование $\mathcal{K}_n \operatorname{vec} M = \operatorname{vec} M^T$. Матрица $\mathcal{N}_n$ в явном виде может быть записана как

$$\mathcal{N}_n = \frac{1}{2}\Big(I_{n^2} + \sum_{i=1}^{n}\sum_{j=1}^{n}(e_i e_j^T)\otimes(e_j e_i^T)\Big).$$

Данная матрица является симметричной и идемпотентной. На каждую квадратную матрицу M порядка n она действует следующим образом:

$$\mathcal{N}_n \operatorname{vec} M = \operatorname{vec} M^S,$$

где $M^S = \left(M + M^T\right)/2$ — симметричная часть матрицы M. Таким образом, если матрица M симметричная, то $\operatorname{vec} M$ не меняется под действием матрицы $\mathcal{N}_n$. Имеет место равенство $\mathcal{D}_n\mathcal{L}_n\mathcal{N}_n = \mathcal{N}_n$ (см. [10]).

Так как $F(X)$ является симметричной матричной функцией, то в матрице Якоби $F_X(X)$ строки с номерами (i,j) и (j,i), $i \neq j$, совпадают. Это означает, что каждому столбцу матрицы $F_X(X)$ соответствует симметричная матрица, т.е. каждый столбец есть прямая сумма столбцов симметричной матрицы. Поэтому $F_X(X) = \mathcal{N}_n F_X(X)$ и, следовательно, для матрицы $\Lambda(X)$ наряду с (41) справедливо представление

$$\Lambda(X) = \mathcal{L}_n\mathcal{N}_n\Big\{[I_{n^2} - \mathcal{P}(X^{\otimes})]V^{\otimes} + \tau\,\mathcal{P}(X^{\otimes})\Big\}\mathcal{D}_n.$$

Если теперь умножить матрицу $\mathcal{A}_{\rm vec}$ на матрицу $\mathcal{D}_n\Lambda(X)$ справа, то в силу вышесказанного получаем

$$\begin{aligned} \mathcal{A}_{\rm vec}\,\mathcal{D}_n\Lambda(X) &= \mathcal{A}_{\rm vec}\,\mathcal{D}_n\mathcal{L}_n\mathcal{N}_n\big\{[I_{n^2}-\mathcal{P}(X^{\otimes})]\,V^{\otimes}+\tau\,\mathcal{P}(X^{\otimes})\big\}\mathcal{D}_n \\ &= \mathcal{A}_{\rm vec}\,\mathcal{N}_n\big\{[I_{n^2}-\mathcal{P}(X^{\otimes})]\,V^{\otimes}+\tau\,\mathcal{P}(X^{\otimes})\big\}\mathcal{D}_n \\ &= \mathcal{A}_{\rm vec}\big\{[I_{n^2}-\mathcal{P}(X^{\otimes})]\,V^{\otimes}+\tau\,\mathcal{P}(X^{\otimes})\big\}\mathcal{D}_n \\ &= \tau\,\mathcal{A}_{\rm vec}\,\mathcal{D}_n. \end{aligned}$$

Отсюда следует равенство

$$\mathcal{A}_{\rm vec}\,\mathcal{D}_n = \tau\,\mathcal{A}_{\rm vec}\,\mathcal{D}_n\Lambda^{-1}(X). \tag{43}$$

На основании (43) приходим к заключению, что

$$\begin{aligned} \mathcal{A}_{\rm vec}\,\mathcal{D}_n\Lambda^{-1}(X)\,\mathcal{L}_n{\rm vec}\,F(X) &= \mathcal{A}_{\rm vec}\,\mathcal{D}_n\mathcal{L}_n{\rm vec}\,F(X)\,/\,\tau \\ &= \mathcal{A}_{\rm vec}\,\mathcal{D}_n\mathcal{L}_n\,\mathcal{N}_n{\rm vec}\,F(X)\,/\,\tau \\ &= \mathcal{A}_{\rm vec}\,{\rm vec}\,F(X)\,/\,\tau. \end{aligned}$$

Но согласно (5)

$$\mathcal{A}_{\rm vec}\,{\rm vec}\,F(X) = \tau\,(\mathcal{A}\bullet X - b)\,.$$

Тем самым нами получен следующий результат.

Утверждение 6. *Вдоль траекторий системы* (42) *выполняются равенства*:

$$\frac{dg^i(X)}{dt} = -\left\langle {\rm vec}\,A_i, {\rm vec}\left(\frac{dX}{dt}\right)\right\rangle = \gamma\,\langle {\rm vec}\,A_i, \mathcal{D}_n\Lambda^{-1}(X)\,\mathcal{L}_n{\rm vec}\,F(X)\rangle = -\gamma\,g^i(X), \tag{44}$$

где $1\le i\le m$.

Таким образом, производные по времени ограничений (44), вычисленные в силу системы (42), совпадают с (16), если взять $\gamma=\tau$.

Итерационный процесс (40) строился в предположении, что матрица $\Lambda(X)=F_X^{\triangle}(X)$ невырожденная. Поэтому для того, чтобы итерационный процесс был определен, необходимо, чтобы она была невырожденной в некоторой области, содержащей решение задачи (1). Как показано в [5], имеет место следующий результат.

Лемма 1. *Пусть для прямой и двойственной задач* (1) *и* (2) *имеет место строгая двойственность, причем их решения* X_* *и* V_* *строго комплементарны. Пусть, кроме того, точка* X_* *есть вершина множества* $\mathcal{F}_P$. *Тогда матрица* $F_X^{\triangle}(X_*)$ *невырожденная.*

Если матрица $F_X^{\triangle}(X)$ невырожденная в решении задачи X_*, то в силу непрерывности это свойство сохранится и в некоторой окрестности X_*. Тогда итерационный процесс (40) полностью определен и обладает локальной сходимостью.

Теорема 2. *Пусть выполнены предположения леммы* 1. *Тогда метод* (40) *локально сходится к* X_* *со сверхлинейной скоростью. Если отображение* $F_X^{\triangle}(X)$ *удовлетворяет условию Липшица в некоторой окрестности* X_*, *то скорость сходимости квадратичная.*

Д о к а з а т е л ь с т в о. Для обоснования локальной сходимости процесса (40) фактически следует доказать, что локально сходится процесс (23). Сходимость же процесса (23) следует из утверждения леммы 1 и из [11, теорема 10.2.2]. □

Вид матрицы $\Lambda(X)$ при $X=X_*$ упрощается. Действительно, так как симметричные матрицы X_* и V_* коммутируют, то, как несложно убедиться с помощью равенства

$$(A\otimes B)\,(C\otimes D) = (AC)\otimes(BD)\,, \tag{45}$$

имеющего место для любых матриц A, B, C и D, для которых существуют произведения AC и BD, матрицы $X_*^{\otimes}$ и $V_*^{\otimes}$ также коммутируют. Поэтому найдется ортогональная матрица H порядка n^2 такая, что

$$X_*^{\otimes} = HD(\lambda_*)H^T, \qquad V_*^{\otimes} = HD(\mu_*)H^T, \tag{46}$$

где λ_* и μ_* — 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 $V_*^{\otimes}$ вещественны.

Используя представление (3), получаем на основании формулы (45)

$$\begin{aligned} X_* \otimes I_n &= \left(QD(\eta_*)Q^T\right) \otimes I_n = ((QD(\eta_*)) \otimes I_n)\left(Q^T \otimes I_n\right) \\ &= (Q \otimes I_n)(D(\eta_*) \otimes I_n)\left(Q^T \otimes I_n\right). \end{aligned}$$

Аналогичным образом находим

$$I_n \otimes X_* = (I_n \otimes Q)(I_n \otimes D(\eta_*))\left(I_n \otimes Q^T\right).$$

Поэтому

$$X_*^{\otimes} = \left[(Q \otimes I_n)(D(\eta_*) \otimes I_n)\left(Q^T \otimes I_n\right) + (I_n \otimes Q)(I_n \otimes D(\eta_*))\left(I_n \otimes Q^T\right)\right]/2.$$

Отметим, что матрица $Q^T \otimes I_n$ является обратной к матрице $Q \otimes I_n$, а матрица $I_n \otimes Q^T$ — обратной к матрице $I_n \otimes Q$.

Обозначим

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

и покажем, что компоненты вектора $\eta_*^{\otimes}$ являются собственными значениями матрицы $X_*^{\otimes}$.

Утверждение 7. *Для матрицы $X_*^{\otimes}$ справедливо разложение*

$$X_*^{\otimes} = (Q \otimes Q)D(\eta_*^{\otimes})(Q^T \otimes Q^T), \tag{47}$$

где $Q \otimes Q$ — ортогональная матрица.

Доказательство. Так как симметричные матрицы $X_* \otimes I_n$ и $I_n \otimes X_*$ коммутируют, то ортогональную матрицу H в разложении (46) можно взять такую, что

$$X_* \otimes I_n = HD(\lambda_1)H^T, \qquad I_n \otimes X_* = HD(\lambda_2)H^T,$$

где λ_1 и λ_2 — n^2-мерные векторы, состоящие из собственных значений соответственно матриц $X_* \otimes I_n$ и $I_n \otimes X_*$. Отсюда следует представление

$$X_*^{\otimes} = HD\left(\frac{\lambda_1 + \lambda_2}{2}\right)H^T,$$

т.е. $\lambda_* = (\lambda_1 + \lambda_2)/2$.

Если положить $H = Q \otimes Q$, то $H^T = Q^T \otimes Q^T$ и

$$HH^T = (Q \otimes Q)\left(Q^T \otimes Q^T\right) = (QQ^T) \otimes (QQ^T) = I_n \otimes I_n = I_{n^2}.$$

Следовательно, матрица H — ортогональная. Имеем:

$$\begin{aligned} H^T(I_n \otimes X_*)H &= (Q^T \otimes Q^T)(I_n \otimes X_*)(Q \otimes Q) \\ &= \left(Q^T \otimes (Q^T X_*)\right)(Q \otimes Q) \\ &= (Q^T Q) \otimes (Q^T X_* Q) \\ &= I_n \otimes D(\eta_*). \end{aligned}$$

Аналогично получаем: $H^T(X_* \otimes I_n)H = D(\eta_*) \otimes I_n$.

Матрицы $D(\eta_*) \otimes I_n$ и $I_n \otimes D(\eta_*)$ — диагональные. Поэтому данную ортогональную матрицу $Q \otimes Q$ можно взять в качестве H. Отсюда следует справедливость разложения (47). □

Аналогичным образом доказывается следующее утверждение.

Утверждение 8. *Для матрицы* $V_*^{\otimes}$ *справедливо разложение*

$$V_*^{\otimes} = (Q \otimes Q) D(\theta_*^{\otimes})(Q^T \otimes Q^T), \tag{48}$$

где $\theta_*^{\otimes} = \mathrm{Diag}\,(D^{\otimes}(\eta_*))$, $D^{\otimes}(\theta_*) = [D(\theta_*) \otimes I_n + I_n \otimes D(\theta_*)]\,/\,2$.

После подстановки выражений (47) и (48) в $F_X^{\triangle}(X_*)$ получаем

$$F_X^{\triangle}(X_*) = \mathcal{L}_n \mathcal{N}_n (Q \otimes Q) \left\{ \left[I_{n^2} - \mathcal{G}(\eta_*^{\otimes}) \right] D(\theta_*^{\otimes}) + \tau \mathcal{G}(\eta_*^{\otimes}) \right\} (Q^T \otimes Q^T) \mathcal{D}_n. \tag{49}$$

В (49) матрица $\mathcal{G}(\eta_*^{\otimes})$ имеет вид

$$\mathcal{G}(\eta_*^{\otimes}) = D(\eta_*^{\otimes})(Q^T \otimes Q^T)\, \mathcal{A}_{\mathrm{vec}}^T \left[\mathcal{A}_{\mathrm{vec}}\,(Q \otimes Q) D(\eta_*^{\otimes})(Q^T \otimes Q^T) \mathcal{A}_{\mathrm{vec}}^T \right]^{-1} \mathcal{A}_{\mathrm{vec}}\,(Q \otimes Q).$$

В заключение отметим, что предположение о невырожденности задачи можно ослабить, потребовав, чтобы условие невырожденности выполнялось лишь в некоторой окрестности решения (в том числе и в недопустимых точках).

СПИСОК ЛИТЕРАТУРЫ

1. **Еремин И.И.** Теория линейной оптимизации. Екатеринбург: Изд-во “Екатеринбург”, 1999.
2. Handbook of semidefinite programming /eds. H. Wolkowicz, R. Saigal, L. Vandenberghe. Dordrecht: Kluwer Acad. Publ., 2000.
3. **Laurent M., Rendle F.** Semidefinite programming and integer programming // Discrete optimization. Handbooks in operations research and management science / eds. K. Aardal, G. Nemhauser, R. Weismantel. Amsterdam: Elsevier, 2005.
4. **Klerk E., de.** Aspects of semidefinite programming. Interior point algorithms and selected applications. Dordrecht: Kluwer Acad. Publ., 2004.
5. **Бабынин М.С., Жадан В.Г.** Барьерно-проективный метод для полуопределенного программирования. Сообщения по прикладной математике. М.: ВЦ РАН, 2007.
6. **Evtushenko Yu., Zhadan V.** Stable barrier-projection and barrier-newton methods in linear programming // Comput. Optimiz. and Appl. 1994. Vol. 3. P. 289–304.
7. **Арнольд В.И.** О матрицах, зависящих от параметров // Успехи мат. наук. 1971. Т. 26. Вып. 2(158). С. 101–114.
8. **Alizadeh F., Haeberly J.-P.F., Overton M.L.** Complementarity and nondegeneracy in semidefinite programming // Math. Programming. Ser. B. 1997. Vol. 7, no. 2. P. 129–162.
9. **Магнус Я.Р., Нейдеккер Ч.** Матричное дифференциальное исчисление с приложениями к статистике и эконометрике. М.: Физматлит, 2002.
10. **Magnus J.R., Neudecker H.** The elimination matrix: some lemmas and applications // SIAM J. Alg. Disc. Meth. 1980. Vol. 1, no. 4. P. 422–449.
11. **Ортега Дж., Рейнболдт В.** Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М.: Мир, 1975.

Поступила 10.01.2008