В. Г. Жадан, Об одном варианте допустимого аффинно-масштабирующего метода для полуопределенного программирования, *Тр. ИММ УрО РАН*, 2014, том 20, номер 2, 145–160

ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 20 № 2 2014

УДК 519.856

ОБ ОДНОМ ВАРИАНТЕ ДОПУСТИМОГО АФФИННО-МАСШТАБИРУЮЩЕГО МЕТОДА ДЛЯ ПОЛУОПРЕДЕЛЕННОГО ПРОГРАММИРОВАНИЯ[1]

В. Г. Жадан

Рассматривается линейная задача полуопределенного программирования. Для ее решения предлагается прямой допустимый аффинно-масштабирующий метод, в котором точки итерационного процесса могут принадлежать границе допустимого множества.

Ключевые слова: линейная задача полуопределенного программирования, прямой аффинно-масштабирующий метод, наискорейший спуск.

V. G. Zhadan. On a variant of a feasible affine scaling method for semidefinite programming.

A linear problem of semidefinite programming is considered. For its solution, a primal feasible affine scaling method is proposed, in which the points of the iterative process may belong to the boundary of the feasible set.

Keywords: linear semidefinite programming problem, primal affine scaling method, steepest descent.

Введение

Линейные задачи полуопределенного программирования представляют собой важный раздел линейной оптимизации [1]. Данные задачи интенсивно изучаются начиная с 90-х годов прошлого века [2]. За это время были разработаны различные численные методы их решения, главным образом аффинно-масштабирующего типа (прямые, двойственные и прямо-двойственные). В частности, в [3] был предложен один из вариантов прямого метода внутренней точки, который также можно трактовать как аффинно-масштабирующий. Было показано, что метод обладает локальной сходимостью. Если стартовая точка в таком методе берется из относительной внутренности допустимого множества, то метод при надлежащем выборе шага перемещения ведет себя как релаксационный, т. е. все последующие точки принадлежат допустимому множеству и значение целевой функции убывает от итерации к итерации. Поэтому возникает идея брать максимально возможный шаг, чтобы увеличить эффективность метода и сделать его глобально сходящимся. Однако такая стратегия выбора шага приводит к тому, что на какой-то итерации точка может оказаться на границе допустимого множества и направление перемещения, подсчитанное в этой точке, выводит за его пределы. Выходом из этой ситуации может служить изменение правых частей в граничных точках, которое не позволяло бы траекториям покидать допустимое множество. Возможный подход к такому изменению был рассмотрен в [4], где направление перемещения строилось в виде суммы двух направлений, одно из которых принадлежит минимальной грани конуса положительно полуопределенных матриц, содержащей текущую точку, а второе — сопряженной грани. В настоящей работе также используется идея разбиения направления на две составляющие. Однако теперь в качестве второго направления берется собственный вектор двойственной невязки, подсчитанный в этой точке. Данное изменение правых частей является в некотором смысле обобщением подхода, применявшегося в [5] для задач линейного программирования.

[1]Работа выполнена при поддержке РФФИ (проект 14-07-00805), программы ведущих научных школ (НШ-4640.2014.1) и программы Президиума РАН П-18.

Работа состоит из пяти разделов. В разд. 1 дается постановка задачи и приводятся некоторые сведения о геометрических свойствах допустимого множества. Основной вариант допустимого метода описывается в разд. 2. В разд. 3 рассматривается вариант метода, когда направление перемещения принадлежит минимальной грани. Совместное направление, при котором возможен сход с минимальной грани, строится в разд. 4. Наконец, в разд. 5 показывается, что траектория при таком выборе направления перемещения, а также шага перемещения из условия наискорейшего спуска не покидает допустимое множество и сходится глобально.

1. Задача полуопределенного программирования

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

Скалярное (внутреннее) произведение между двумя квадратными матрицами L и M порядка 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 из $\mathbb{S}^n$, 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 $L \in \mathbb{S}^n$ положительно полуопределена тогда и только тогда, когда $L \bullet M \geq 0$ для всех $M \succeq 0$, т. е. конус $\mathbb{S}^n_+$ является *самосопряженным*.

Рассмотрим задачу полуопределенного программирования следующего вида:

$$\begin{gathered} \min\, C \bullet X, \\ A_i \bullet X = b^i, \qquad i = J^m, \qquad X \succeq 0. \end{gathered} \tag{1.1}$$

Здесь и ниже: J^k — множество индексов от 1 до k, матрицы C, X и A_i, $i \in J^m$, принадлежат пространству $\mathbb{S}^n$.

Двойственной к (1.1) является задача

$$\begin{gathered} \max\, \langle b, u \rangle, \\ \sum_{i=1}^{m} u^i A_i + V = C, \qquad V \succeq 0, \end{gathered} \tag{1.2}$$

в которой $b = (b^1, \ldots, b^m)^T \in \mathbb{R}^m$, $V \in \mathbb{S}^n$, угловые скобки обозначают обычное евклидово скалярное произведение в $\mathbb{R}^m$. Предполагается, что задачи (1.1) и (1.2) имеют решения. Кроме того, считаем, что матрицы A_i, $i \in J^m$, линейно независимы и C не принадлежит линейному подпространству, порожденному этими матрицами.

Обозначим допустимое множество в исходной задаче (1.1) через $\mathcal{F}_P$, т. е.

$$\mathcal{F}_P = \mathcal{F}_A \cap \mathbb{S}^n_+, \qquad \mathcal{F}_A = \left\{ X \in \mathbb{S}^n \colon A_i \bullet X = b^i,\ i \in J^m \right\}.$$

Его относительной внутренностью является множество $\mathcal{F}^0_P = \mathcal{F}_P \cap \mathbb{S}^n_{++}$.

Ниже нам потребуются некоторые известные результаты (см., например, [2]), касающиеся геометрических свойств допустимого множества $\mathcal{F}_P$. Для произвольной матрицы M обозначим через $\mathcal{R}(M)$ пространство ее столбцов, через $\mathcal{N}(M)$ — ее нуль-пространство (ядро).

Рассмотрим сначала конус $\mathbb{S}^n_+$. *Грани* конуса $\mathbb{S}^n_+$ тесно связаны с подпространствами $\mathcal{L}$ пространства $\mathbb{R}^n$, а именно $\mathcal{G}$ есть грань $\mathbb{S}^n_+$ тогда и только тогда, когда

$$\mathcal{G} = \mathcal{G}(\mathcal{L}) = \left\{X \in \mathbb{S}^n_+ \colon \mathcal{R}(X) \subseteq \mathcal{L}\right\}$$

для некоторого $\mathcal{L} \subset \mathbb{R}^n$. Если размерность $\mathcal{L}$ равна r, то $\operatorname{rank} X \leq r$ для всех элементов X из $\mathcal{G}(\mathcal{L})$, а размерность $\mathcal{G}(\mathcal{L})$ равняется $r_\triangle$. Сама матрица X может быть представлена в виде $X = Q\Lambda Q^T$, где Q — матрица полного ранга размера $n \times r$ и $\Lambda \in \mathbb{S}^r_+$. Для всех матриц $X \in \mathcal{G}(\mathcal{L})$ матрица Q одна и та же. *Сопряженная грань* $\mathcal{G}^\star(\mathcal{L})$ к грани $\mathcal{G}(\mathcal{L})$ определяется как грань, соответствующая ортогональному подпространству $\mathcal{L}^\perp$, т. е. $\mathcal{G}^\star(\mathcal{L}) = \mathcal{G}(\mathcal{L}^\perp)$.

Для точки $X \in \mathbb{S}^n_+$ обозначим через $\mathcal{G}_{\min}(X; \mathbb{S}^n_+)$ *минимальную грань* конуса $\mathbb{S}^n_+$, содержащую X:

$$\mathcal{G}_{\min}(X; \mathbb{S}^n_+) = \{Y \in \mathbb{S}^n_+ \colon \mathcal{R}(Y) \subseteq \mathcal{R}(X)\}.$$

Таким образом, если матрица $X \in \mathbb{S}^n_+$ имеет ранг r, то грань $\mathcal{G}_{\min}(X; \mathbb{S}^n_+)$ изоморфна конусу $\mathbb{S}^r_+$ и, следовательно, имеет размерность $r_\triangle$. Сопряженная грань $\mathcal{G}^*_{\min}(X; \mathbb{S}^n_+)$ изоморфна конусу $\mathbb{S}^{n-r}_+$ и имеет размерность $(n-r)_\triangle$.

Обратимся теперь к допустимому множеству $\mathcal{F}_P$ в прямой задаче. Для матрицы $X \in \mathcal{F}_P$ обозначим через $\mathcal{G}_{\min}(X; \mathcal{F}_P)$ *минимальную грань* множества $\mathcal{F}_P$, содержащую X. Если для $X \in \mathcal{F}_P$ минимальная грань совпадает с ней самой, то такая матрица является *крайней точкой* множества $\mathcal{F}_P$.

Допустимое множество $\mathcal{F}_P$ есть пересечение конуса $\mathbb{S}^n_+$ с аффинным множеством $\mathcal{F}_A$. Так как пересечение двух выпуклых множеств является гранью тогда и только тогда, когда оно является пересечением двух граней, то грань множества $\mathcal{F}_P$ определяется как $\mathcal{G}(\mathcal{L}) \bigcap \mathcal{F}_A$, а минимальная грань для $X \in \mathcal{F}_P$ есть теперь

$$\mathcal{G}_{\min}(X; \mathcal{F}_P) = \mathcal{G}_{\min}(X; \mathbb{S}^n_+) \bigcap \mathcal{F}_A = \{Y \in \mathcal{F}_P \colon \mathcal{R}(Y) \subseteq \mathcal{R}(X)\}.$$

Пусть r есть ранг матрицы X и $X = QQ^T$, где Q — матрица полного ранга размера $n \times r$. Положим $A_i^Q = Q^T A_i Q$, $i \in J^m$. Размерность грани $\mathcal{G}_{\min}(X; \mathcal{F}_P)$ равняется величине

$$\dim \mathcal{G}_{\min}(X; \mathcal{F}_P) = r_\triangle - \operatorname{rank}\left[A_1^Q, \ldots, A_m^Q\right]. \tag{1.3}$$

Из (1.3), в частности, следует, что матрица X ранга r является крайней точкой множества $\mathcal{F}_P$ тогда и только тогда, когда $\operatorname{rank}\left[A_1^Q, \ldots, A_m^Q\right] = r_\triangle$. Для линейно независимых матриц A_1, $A_2, \ldots, A_m$ данное равенство может выполняться только в случае, когда $r_\triangle \leq m$.

Возьмем произвольную матрицу X из допустимого множества $\mathcal{F}_P$ ранга r. Предположим, что

$$X = Q \operatorname{Diag}(\eta_1, \ldots, \eta_r, 0, \ldots, 0)\, Q^T, \tag{1.4}$$

где Q — ортогональная матрица порядка n. Касательное пространство к $\mathbb{S}^n_+$ в X имеет следующий вид [6]:

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

Здесь $\mathbb{R}^{k \times l}$ — пространство $(k \times l)$-матриц. Размерность $\mathcal{T}_X$ определяется рангом матрицы X и вычисляется как

$$\dim \mathcal{T}_X = r_\triangle + r(n-r) = n_\triangle - (n-r)_\triangle.$$

Если матрица $\Delta X \in \mathcal{T}_X$ и $H \neq 0_{r(n-r)}$, то $X \pm \alpha\, \Delta X$ не содержится в $\mathbb{S}^n_+$ для $\alpha > 0$.

Обозначим через $\mathcal{N}_A$ подпространство в $\mathbb{S}^n$, параллельное аффинному множеству $\mathcal{F}_A$. Размерность $\mathcal{N}_A$ в силу сделанного предположения о линейной независимости матриц A_i, $i \in J^m$, равна $n_\triangle - m$. Дадим теперь определение невырожденной точки $X \in \mathcal{F}_P$ для прямой задачи (1.1), следуя [7].

Определение 1.1. Точка $X \in \mathcal{F}_P$ называется *невырожденной*, если $\mathcal{T}_X + \mathcal{N}_A = \mathbb{S}^n$.

Так как $\dim \mathbb{S}^n = n_\triangle$, то равенство $\mathcal{T}_X + \mathcal{N}_A = \mathbb{S}^n$ имеет место только тогда, когда

$$\dim \mathcal{T}_X + \dim \mathcal{N}_A \geq \dim \mathbb{S}^n,$$

откуда вытекает неравенство $m \leq n_\triangle - (n-r)_\triangle$. Данное неравенство является необходимым условием невырожденности допустимой точки X в прямой задаче (1.1).

Определение 1.2. Прямая задача полуопределенного программирования (1.1) является *невырожденной*, если все точки $X \in \mathcal{F}_P$ оказываются невырожденными.

Пусть X_* и $[u_*, V_*]$ — решения соответственно задач (1.1) и (1.2). Тогда матрицы X_* и V_* коммутируют между собой и выполняется условие дополнительности $\eta^i_* \theta^i_* = 0$, $i \in J^n$, для собственных чисел этих матриц. Условие строгой дополнительности означает, что одновременно $\eta^i_* + \theta^i_* > 0$ для всех $i \in J^n$.

2. Допустимый аффинно-масштабирующий метод

Введем сначала дополнительные обозначения. Если M — квадратная матрица порядка n, то символом $\operatorname{vec} M$ 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 $\operatorname{vec} M$ рассматривать вектор-столбец $\operatorname{hvec} M$. 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 $\operatorname{svec} M$. От $\operatorname{hvec} M$ он отличается только тем, что все элементы, не стоящие на диагонали матрицы M, при помещении в $\operatorname{svec} M$ умножаются на $\sqrt{2}$. Как вектор $\operatorname{hvec} M$, так и вектор $\operatorname{svec} M$ имеют длину $n_\triangle$.

Для перехода от вектора $\operatorname{vec} M$ к вектору $\operatorname{hvec} M$ и для обратного перехода используются специальные *элиминационные* и *дуплицирующие* матрицы (см. [8]). Элиминационная матрица $\mathcal{L}_n$ для каждой квадратной матрицы M порядка n совершает преобразование $\mathcal{L}_n \operatorname{vec} M = \operatorname{hvec} M$. Напротив, дуплицирующая матрица $\mathcal{D}_n$ для каждой симметричной матрицы M порядка n осуществляет обратное преобразование $\mathcal{D}_n \operatorname{hvec} M = \operatorname{vec} M$. Матрица $\mathcal{L}_n$ имеет размер $n_\triangle \times n^2$, матрица $\mathcal{D}_n$ — размер $n^2 \times n_\triangle$. Обе матрицы $\mathcal{L}_n$ и $\mathcal{D}_n$ являются матрицами полного ранга, равного $n_\triangle$. Матрица $\mathcal{L}_n$ — полуортогональная, т. е. $\mathcal{L}_n \mathcal{L}_n^T = I_{n_\triangle}$. Кроме того, $\mathcal{L}_n \mathcal{D}_n = I_{n_\triangle}$.

Пусть E_n — квадратная матрица порядка n, все элементы которой равны единице. Пусть, кроме того, D_2 — диагональная матрица порядка $n_\triangle$, на диагонали которой располагается вектор $\operatorname{svec} E_n$. Наряду с матрицами $\mathcal{L}_n$ и $\mathcal{D}_n$ в дальнейшем будем пользоваться также матрицами $\tilde{\mathcal{L}}_n = D_2 \mathcal{L}_n$ и $\tilde{\mathcal{D}}_n = \mathcal{D}_n D_2^{-1}$.

Таким образом, если M — симметричная матрица порядка n, то, как можно проверить,

$$\operatorname{svec} M = \tilde{\mathcal{L}}_n \operatorname{vec} M = \tilde{\mathcal{D}}_n^T \operatorname{vec} M, \qquad \operatorname{vec} M = \tilde{\mathcal{D}}_n \operatorname{svec} M.$$

Для матриц $\tilde{\mathcal{L}}_n$ и $\tilde{D}_n$ сохраняется свойство $\tilde{\mathcal{L}}_n \tilde{\mathcal{D}}_n = I_{n_\triangle}$.

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

$$J^n_\square = \{(1,1), \ldots, (n,1), (1,2), \ldots, (n,2), \ldots, (1,n), \ldots, (n,n)\}. \tag{2.1}$$

Всего в таком наборе n^2 парных индексов. Номера вида (i,i), в которых первый индекс совпадает со вторым, будем называть *диагональными*. Набор $J^n_\square$ содержит n таких диагональных номеров, а именно $(1,1)$, $(2,2)$, $\ldots$, (n,n). Те номера (i,j) из $J^n_\square$, в которых индексы i и j отличны друг от друга, будем называть *внедиагональными*. При этом если $i > j$, то такой номер (i,j) называется *младшим внедиагональным* номером.

Следующий набор парных индексов удобно использовать при работе с векторами $\mathrm{hvec}\, M$ или $\mathrm{svec}\, M$. Это набор

$$J^n_{\triangle} = \{(1,1), \ldots, (n,1), (2,2), \ldots, (n,2), (3,3), \ldots, (n,n)\},$$

содержащий $n_{\triangle}$ парных индексов. В него входят только диагональные номера и младшие внедиагональные номера.

Обратимся теперь к задачам полуопределенного программирования (1.1) и (1.2). Мы предположили, что их решения существуют. Тогда в силу необходимых и достаточных условий оптимальности система равенств и неравенств

$$\begin{array}{rcl} X \bullet V & = & 0, \\ A_i \bullet X & = & b^i, \quad i \in J^m, \\ V & = & C - \sum_{i=1}^m u^i A_i, \\ X \succeq 0, & & V \succeq 0 \end{array} \tag{2.2}$$

обязательно имеет решение.

Обозначим через $X \circ V$ симметризованное произведение матриц X и V из $\mathbb{S}^n$, т. е. матрицу $X \circ V = (XV + VX)/2$. Данное произведение обладает полезным свойством, а именно для матриц $X \in \mathbb{S}^n_+$ и $V \in \mathbb{S}^n_+$ равенство $X \circ V = 0_{nn}$ возможно в том и только в том случае, когда $XV = VX = 0_{nn}$.

Воспользуемся также тем, что и равенство $X \bullet V = 0$ для X и V из $\mathbb{S}^n_+$ выполняется тогда и только тогда, когда $XV = VX = 0_{nn}$. Поэтому первое равенство из (2.2) может быть переписано в виде

$$X \circ V = 0_{nn}. \tag{2.3}$$

Заменим теперь в системе (2.2), (2.3) матричные равенства на их векторные аналоги. С учетом известной формулы $\mathrm{vec}\,(ABC) = (C^T \otimes A)\mathrm{vec}\, B$, справедливой для любых матриц A, B и C, для которых определено произведение ABC, получаем

$$X^{\otimes}\, \mathrm{vec}\, V = 0_{n^2}, \qquad \mathcal{A}_{vec}\, \mathrm{vec}\, X = b, \qquad \mathrm{vec}\, V = \mathrm{vec}\, C - \mathcal{A}^T_{vec} u. \tag{2.4}$$

Здесь $X^{\otimes}$ — кронекеровская сумма матрицы X, определяемая как

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

Через $\mathcal{A}_{vec}$ обозначена матрица размера $m \times n^2$ со строками $\mathrm{vec}\, A_i$, $i \in J^m$.

Учтем симметричность матриц. Тогда система (2.4) заменяется на следующую:

$$\tilde{X}^{\otimes} \mathrm{svec}\, V = 0_{n_{\triangle}}, \qquad \mathcal{A}_{svec}\, \mathrm{svec}\, X = b, \qquad \mathrm{svec}\, V = \mathrm{svec}\, C - \mathcal{A}^T_{svec} u. \tag{2.5}$$

В (2.5) и ниже $\tilde{X}^{\otimes} = \tilde{\mathcal{L}}_n X^{\otimes} \tilde{\mathcal{D}}_n$, $\mathcal{A}_{svec}$ — матрица размера $m \times n_{\triangle}$ со строками $\mathrm{svec}\, A_i$, $i \in J^m$.

Подставим вектор $\mathrm{svec}\, V$ из третьего равенства в первое и умножим его левую и правую части на матрицу $\mathcal{A}_{svec}$. В результате приходим к уравнению относительно вектора u:

$$\Gamma(\tilde{X}^{\otimes}) u = \mathcal{A}_{svec}\, \tilde{X}^{\otimes} \mathrm{svec}\, C, \tag{2.6}$$

где через $\Gamma(\tilde{X}^{\otimes})$ обозначена матрица $\Gamma(\tilde{X}^{\otimes}) = \mathcal{A}_{svec}\, \tilde{X}^{\otimes} \mathcal{A}^T_{svec}$.

Если матрица $\Gamma(\tilde{X}^{\otimes})$ неособая, то, разрешая уравнение (2.6) относительно u, получаем:

$$u = u(X) = \Gamma^{-1}(\tilde{X}^{\otimes}) \mathcal{A}_{svec}\, \tilde{X}^{\otimes} \mathrm{svec}\, C. \tag{2.7}$$

В [3] показано, что при предположении о невырожденности прямой задачи (1.1) матрица $\Gamma(\tilde{X}^{\otimes})$ оказывается неособой для всех точек X из некоторой области, содержащей допустимое множество $\mathcal{F}_P$. В дальнейшем нам потребуются обозначения

$$V(u) = C - \sum_{i=1}^m u^i A_i, \qquad V(X) = V(u(X)).$$

Утверждение 2.1. *Пусть точка* X_* *является решением прямой задачи* (1.1). *Тогда пара* $[u_*, V_*]$, *где* $u_* = u(X_*)$, $V_* = V(u_*)$, *есть решение двойственной задачи* (1.2).

Доказательство. Так как, по предположению, обе задачи (1.1) и (1.2) имеют решение, то в силу условий оптимальности (2.5) найдутся $\bar{u}$ и $\bar{V} \succeq 0$ такие, что $\bar{V} = V(\bar{u})$ и выполняется равенство $\tilde{X}_*^{\otimes} \operatorname{svec} \bar{V} = 0_{n_\triangle}$. Пара $[\bar{u}, \bar{V}]$ является решением задачи (1.2). Если теперь умножить это равенство слева на матрицу $\mathcal{A}_{svec}$ и подставить в него выражение $\operatorname{svec} \bar{V} = \operatorname{svec} C - \mathcal{A}_{svec}^T \bar{u}$, то оно перейдет в следующее равенство:

$$\Gamma(\tilde{X}_*^{\otimes})\bar{u} = \mathcal{A}_{svec} \tilde{X}_*^{\otimes} \operatorname{svec} C. \tag{2.8}$$

Матрица $\Gamma(\tilde{X}_*^{\otimes})$ из-за предположения о невырожденности задачи (1.1) неособая. Поэтому равенство (2.8), если его рассматривать как уравнение относительно $\bar{u}$, может иметь только единственное решение.

С другой стороны, согласно формуле (2.7) $u_* = \Gamma^{-1}(\tilde{X}_*^{\otimes})\mathcal{A}_{svec} \tilde{X}_*^{\otimes} \operatorname{svec} C$. Подставляя данное значение u_* в (2.8), получаем, что оно тоже удовлетворяет этому уравнению. Отсюда приходим к выводу, что u_* и V_* совпадают соответственно с $\bar{u}$ и $\bar{V}$. Таким образом, пара $[u_*, V_*]$ есть решение двойственной задачи (1.2).

Утверждение доказано.

Пусть $\operatorname{svec} V(X) = \operatorname{svec} C - \mathcal{A}_{svec}^T u(X)$. Заменим в первом равенстве из (2.5) вектор $\operatorname{svec} V$ на найденный $\operatorname{svec} V(X)$:

$$\tilde{X}^{\otimes}(I_{n_\triangle} - \mathcal{A}_{svec}^T \Gamma^{-1}(\tilde{X}^{\otimes})\mathcal{A}_{svec} \tilde{X}^{\otimes}) \operatorname{svec} C = 0_{n_\triangle}.$$

Это система $n_\triangle$ нелинейных уравнений относительно $n_\triangle$ переменных — компонент вектора $\operatorname{svec} X$. Для ее решения можно применять различные численные методы решения систем нелинейных уравнений, в частности метод простой итерации.

Обозначим

$$\Delta_{svec}(X) = \tilde{X}^{\otimes}(I_{n_\triangle} - \mathcal{A}_{svec}^T \Gamma^{-1}(\tilde{X}^{\otimes})\mathcal{A}_{svec} \tilde{X}^{\otimes}) \operatorname{svec} C. \tag{2.9}$$

Взяв $X_0 \in \mathcal{F}_P^0$ и применяя метод простой итерации, приходим к итерационному процессу

$$\operatorname{svec} X_{k+1} = \operatorname{svec} X_k - \alpha_k \Delta_{svec}(X_k). \tag{2.10}$$

Здесь $\alpha_k > 0$ — шаг спуска. В матричном представлении процесс (2.10) запишется как

$$X_{k+1} = X_k - \alpha_k \Delta(X_k), \qquad \Delta(X) = X \circ V(X). \tag{2.11}$$

Укажем на простейшие свойства такого итерационного процесса.

Утверждение 2.2 [3]. *Для любой точки* $X \in \mathcal{F}_P$ *выполняются неравенство* $C \bullet \Delta(X) \geq 0$ *и равенства* $A_i \bullet \Delta(X) = 0$, *где* $i \in J^m$.

Пусть теперь $X \in \mathbb{S}_+^n$ и

$$X = QD(\eta)Q^T, \tag{2.12}$$

где Q — ортогональная матрица порядка n, $D(\eta)$ — диагональная матрица с вектором собственных значений $\eta = [\eta^1, \ldots, \eta^n]$ на диагонали. Тогда

$$X^{\otimes} = (Q \otimes Q)D(\eta^{\otimes})(Q^T \otimes Q^T).$$

Здесь $\eta^{\otimes}$ — диагональ диагональной матрицы $D^{\otimes}(\eta)$. Матрица $Q \otimes Q$ порядка n^2 также будет ортогональной, и $(Q \otimes Q)^{-1} = Q^T \otimes Q^T$. Если предположить, что матрица $X \in \mathcal{F}_P$ не является положительно определенной, а лишь положительно полуопределенной, то она принадлежит относительной границе множества $\mathcal{F}_P$ и среди собственных чисел η^i имеются нулевые.

Пусть $V^Q = Q^T V Q$ — матрица, являющаяся представлением матрицы V в ортонормированном базисе, задаваемом столбцами матрицы Q.

Утверждение 2.3. *Пусть точка X принадлежит относительной границе допустимого множества $\mathcal{F}_P$. Тогда направление $\Delta(X)$ принадлежит касательному подпространству $\mathcal{T}_X$ к конусу $\mathbb{S}^n_+$ в этой точке.*

Доказательство. После подстановки разложения (2.12) в выражение (2.9) для вектора $\Delta_{svec}(X)$ и перехода к вектору $\Delta_{vec}(X) = \tilde{\mathcal{D}}_n \Delta_{svec}(X)$ получаем

$$\Delta_{vec}(X) = (Q \otimes Q) D(\eta^{\otimes})(Q^T \otimes Q^T) \operatorname{vec} V(X) = (Q \otimes Q) D(\eta^{\otimes}) \operatorname{vec} V^Q(X).$$

Отсюда видно, что $\Delta_{vec}(X) = \operatorname{vec} QY(X)Q^T$, где $Y(X)$ — матрица с прямой суммой столбцов $D(\eta^{\otimes}) \operatorname{vec} V^Q(X)$, т. е. $Y(X) = D(\eta) \circ V^Q(X)$.

Считаем, не умаляя общности, что матрица X имеет ранг $r < n$ и $\eta^i > 0$, $i \in J^r$. Тогда у вектора $\eta^{\otimes}$ все элементы с парными номерами $(i,j) \in J^n_{\square}$ такими, что $r < i, j \leq n$, — нулевые. Поэтому у матрицы $Y(X)$ правая нижняя квадратная подматрица порядка $n - r$ — нулевая. Таким образом, направление $\Delta(X)$ принадлежит касательному подпространству к конусу $\mathbb{S}^n_+$ в точке X.

Утверждение доказано.

Согласно утверждению 2.2 итерационный процесс (2.11) является релаксационным. Как показано в [3], он локально сходится к решению задачи (1.1), если шаг α_k берется постоянным и достаточно малым, а для решений задач (1.1) и (1.2) выполнено условие строгой дополнительности. Однако для того чтобы добиться наибольшего убывания значения целевой функции на текущей итерации, следует взять шаг α_k максимально возможным при условии, что следующая точка X_{k+1} удовлетворяет неравенству $X_{k+1} \succeq 0$. Но при таком выборе шага α_k точка X_{k+1} может оказаться на границе множества $\mathcal{F}_P$. В силу утверждения 2.3 в этом случае направление $\Delta(X)$ принадлежит касательному подпространству к конусу $\mathbb{S}^n_+$ в этой точке и, в принципе, может выводить за пределы данного конуса и, стало быть, за пределы $\mathcal{F}_P$. Поэтому необходимо изменить подход к определению направления $\Delta(X)$ в граничных точках допустимого множества.

3. Направление из минимальной грани

Предположим теперь, что точка X принадлежит относительной границе множества $\mathcal{F}_P$, т. е. среди собственных чисел матрицы X имеются нулевые. Рассмотрим, как изменится направление $\Delta(X)$, если дополнительно потребовать, чтобы оно принадлежало минимальной грани, содержащей X. С этой целью обратимся к "сужению" исходной задачи (1.1) на эту грань.

Пусть $X \in \mathbb{S}^n_+$ и ранг X равен $r < n$. Считаем по-прежнему, что для матрицы X имеет место разложение (2.12), причем положительные собственные числа находятся в начале вектора η (см. (1.4)). Пусть Q_B — левая $n \times r$ подматрица ортогональной матрицы Q, а $D(\eta_B)$ — левая верхняя диагональная подматрица матрицы $D(\eta)$ порядка r, т. е. $X = Q_B D(\eta_B) Q_B^T$. Тогда указанное "сужение" задачи (1.1) на минимальную грань $\mathcal{G}_{\min}(X; \mathbb{S}^n_+)$ запишется в виде

$$\begin{gathered} \min\ C \bullet X, \\ A^i \bullet X = b^i, \qquad i \in J^m, \qquad X \in \mathcal{G}_{\min}(X; \mathbb{S}^n_+). \end{gathered} \tag{3.1}$$

Учтем далее, что точки $X \in \mathcal{G}_{\min}(X; \mathbb{S}^n_+)$ представимы в виде $X = Q_B Z Q_B^T$, где $Z \in \mathbb{S}^r_+$. Подставляя в (3.1) вместо X данное разложение, приходим к системе

$$\begin{gathered} \min\ C^{Q_B} \bullet Z, \\ A_i^{Q_B} \bullet Z = b^i, \qquad i \in J^m, \qquad Z \succeq 0, \end{gathered}$$

где $C^{Q_B} = Q_B^T C Q_B$ и $A_i^{Q_B} = Q_B^T A_i Q_B$, $i \in J^m$. Двойственной к данной задаче будет система

$$\max \langle b, u \rangle,$$
$$\sum_{i=1}^m u^i A_i^{Q_B} + V^{Q_B} = C^{Q_B}, \qquad V^{Q_B} \succeq 0.$$

Соответствующие условия оптимальности (2.2), (2.3) теперь могут быть записаны в виде

$$\begin{array}{rcl} Z \circ V^{Q_B} & = & 0, \\ A_i^{Q_B} \bullet Z & = & b^i, \quad i \in J^m, \\ V^{Q_B} & = & C^{Q_B} - \sum\limits_{i=1}^m u^i A_i^{Q_B}, \\ Z \succeq 0, & & V^{Q_B} \succeq 0. \end{array}$$

Переходя к векторной форме представления равенств, аналогичной (2.4), получаем

$$\tilde{Z}^{\otimes} \operatorname{svec} V^{Q_B} = 0_{r_\triangle}, \qquad \mathcal{A}_{svec}^{Q_B} \operatorname{svec} Z = b, \qquad \operatorname{svec} V^{Q_B} = \operatorname{svec} C^{Q_B} - (\mathcal{A}_{svec}^{Q_B})^T u.$$

Здесь и ниже $\tilde{Z}^{\otimes} = \tilde{\mathcal{L}}_r Z^{\otimes} \tilde{\mathcal{D}}_r$. Матрица $\mathcal{A}_{svec}^{Q_B}$ имеет размер $m \times r_\triangle$, ее строками являются векторы $\operatorname{svec} A_i^{Q_B}$, $i \in J^m$.

Далее поступаем полностью аналогично тому, как это делалось ранее для нахождения направления $\Delta(X)$ в основном варианте допустимого метода. С этой целью выведем сначала уравнение для определения вектора u:

$$\Gamma_{Q_B}(\tilde{Z}^{\otimes}) u = \mathcal{A}_{svec}^{Q_B} \tilde{Z}^{\otimes} \operatorname{svec} C^{Q_B}, \tag{3.2}$$

в котором $\Gamma_{Q_B}(\tilde{Z}^{\otimes}) = \mathcal{A}_{svec}^{Q_B} \tilde{Z}^{\otimes} (\mathcal{A}_{svec}^{Q_B})^T$.

О п р е д е л е н и е 3.1. Точка $X = Q_B D(\eta_B) Q_B^T$ является *регулярной*, если матрицы $A_i^{Q_B}$, $i \in J^m$, линейно независимы. В противном случае эта точка называется *нерегулярной*.

Таким образом, любая невырожденная крайняя точка X ранга r, удовлетворяющего равенству $r_\triangle = m$, является регулярной. Крайняя точка X ранга r, в которой $r_\triangle < m$, заведомо нерегулярная. Верно и обратное: все регулярные точки являются невырожденными. Это следует из достаточных условий невырожденности (см. [7]).

Предположим, что точка $X = Q_B Z Q_B^T$, в которой $Z = D(\eta_B) > 0_r$, является регулярной. Разрешая уравнение (3.2), получаем

$$u(Z) = \Gamma_{Q_B}^{-1}(\tilde{Z}^{\otimes}) \mathcal{A}_{svec}^{Q_B} \tilde{Z}^{\otimes} \operatorname{svec} C^{Q_B}, \quad \operatorname{svec} V^{Q_B} = \operatorname{svec} C^{Q_B} - (\mathcal{A}_{svec}^{Q_B})^T \Gamma_{Q_B}^{-1}(\tilde{Z}^{\otimes}) \mathcal{A}_{svec}^{Q_B} \tilde{Z}^{\otimes} \operatorname{svec} C^{Q_B}.$$

Имеем также для соответствующего направления $\Delta(Z)$ в векторном представлении

$$\Delta_{svec}(Z) = \tilde{Z}^{\otimes} [I_{r_\triangle} - (\mathcal{A}_{svec}^{Q_B})^T \Gamma_{Q_B}^{-1}(\tilde{Z}^{\otimes}) \mathcal{A}_{svec}^{Q_B} \tilde{Z}^{\otimes}] \operatorname{svec} C^{Q_B}. \tag{3.3}$$

Теперь наша цель состоит в том, чтобы найти выражение для направления $\Delta^{Q_B}(X)$ в исходном пространстве $\mathbb{S}^n$ опять же в векторном виде. Имеем

$$\Delta_{svec}^{Q_B}(X) = \tilde{\mathcal{L}}_n \operatorname{vec}(Q_B \Delta(Z) Q_B^T) = \tilde{\mathcal{L}}_n (Q_B \otimes Q_B) \tilde{\mathcal{D}}_r \Delta_{svec}(Z). \tag{3.4}$$

Таким образом, надо выразить $\Delta_{svec}(Z)$ через исходную матрицу X.

Обозначим через $\eta_B^{\otimes}$ диагональ матрицы $D^{\otimes}(\eta_B)$ и через $\tilde{\eta}_B^{\otimes}$ вектор $\tilde{\eta}_B^{\otimes} = \mathcal{L}_r \eta_B^{\otimes}$. Имеет место следующий результат.

Лемма 3.1. *Пусть* $Z = D(\eta_B)$ *и*

$$\tilde{X}^{\otimes}_{Q_B} = \tilde{\mathcal{D}}^T_n (Q_B \otimes Q_B) \tilde{\mathcal{L}}^T_r D(\tilde{\eta}^{\otimes}_B) \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \tilde{\mathcal{D}}_n. \tag{3.5}$$

Пусть, кроме того, $\Gamma(\tilde{X}^{\otimes}_{Q_B}) = \mathcal{A}_{svec} \tilde{X}^{\otimes}_{Q_B} \mathcal{A}^T_{svec}$ *и*

$$\mathcal{P}(\tilde{X}^{\otimes}_{Q_B}) = I_{n_\triangle} - \mathcal{A}^T_{svec} \Gamma^{-1}(\tilde{X}^{\otimes}_{Q_B}) \mathcal{A}_{svec} \tilde{X}^{\otimes}_{Q_B}. \tag{3.6}$$

Тогда

$$\Delta_{svec}(Z) = D(\tilde{\eta}^{\otimes}_B) \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \tilde{\mathcal{D}}_n \mathcal{P}(\tilde{X}^{\otimes}_{Q_B}) \operatorname{svec} C. \tag{3.7}$$

Доказательство. Выразим сначала матрицу $\Gamma_{Q_B}(Z)$ через матрицу $\tilde{X}^{\otimes}_{Q_B}$. Проводя выкладки, получаем

$$\begin{aligned}
\Gamma_{Q_B}(Z) &= \mathcal{A}^{Q_B}_{svec} \tilde{Z}^{\otimes} (\mathcal{A}^{Q_B}_{svec})^T = \mathcal{A}^{Q_B}_{svec} \tilde{\mathcal{L}}_r D^{\otimes}(\eta_B) \tilde{\mathcal{D}}_r (\mathcal{A}^{Q_B}_{svec})^T = \mathcal{A}^{Q_B}_{vec} \tilde{\mathcal{L}}^T_r \tilde{\mathcal{L}}_r D(\eta^{\otimes}_B) (\mathcal{A}^{Q_B}_{vec})^T \\
&= \mathcal{A}_{vec} (Q_B \otimes Q_B) \tilde{\mathcal{L}}^T_r \tilde{\mathcal{L}}_r D(\eta^{\otimes}_B) (Q^T_B \otimes Q^T_B) \mathcal{A}^T_{vec} \\
&= \mathcal{A}_{svec} \tilde{\mathcal{D}}^T_n (Q_B \otimes Q_B) \tilde{\mathcal{L}}^T_r \tilde{\mathcal{L}}_r D(\eta^{\otimes}_B) \tilde{\mathcal{D}}_r \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \tilde{\mathcal{D}}_n \mathcal{A}^T_{svec} \\
&= \mathcal{A}_{svec} \tilde{\mathcal{D}}^T_n (Q_B \otimes Q_B) \tilde{\mathcal{L}}^T_r D(\tilde{\eta}^{\otimes}_B) \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \tilde{\mathcal{D}}_n \mathcal{A}^T_{svec} = \mathcal{A}_{svec} \tilde{X}^{\otimes}_{Q_B} \mathcal{A}^T_{svec} = \Gamma(\tilde{X}^{\otimes}_{Q_B}).
\end{aligned} \tag{3.8}$$

Кроме того, учтем, что

$$\begin{aligned}
\tilde{Z}^{\otimes} \operatorname{svec} C^{Q_B}(\eta_B) &= \tilde{\mathcal{L}}_r D^{\otimes}(\eta_B) \tilde{\mathcal{D}}_r \operatorname{svec} C^{Q_B} = \tilde{\mathcal{L}}_r D(\eta^{\otimes}_B) \operatorname{vec} C^{Q_B} \\
&= \tilde{\mathcal{L}}_r D(\eta^{\otimes}_B) \tilde{\mathcal{D}}_r \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \operatorname{vec} C = D(\tilde{\eta}^{\otimes}_B) \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \tilde{\mathcal{D}}_n \operatorname{svec} C,
\end{aligned}$$

а также что

$$(\mathcal{A}^{Q_B}_{svec})^T = \tilde{\mathcal{L}}_r (\mathcal{A}^{Q_B}_{vec})^T = \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \mathcal{A}^T_{vec} = \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \tilde{\mathcal{D}}_n \mathcal{A}^T_{svec}.$$

Тогда

$$\begin{aligned}
\mathcal{A}^{Q_B}_{svec} \tilde{Z}^{\otimes} \operatorname{svec} C^{Q_B} &= \mathcal{A}_{svec} \tilde{\mathcal{D}}^T_n (Q_B \otimes Q_B) \tilde{\mathcal{L}}^T_r D(\tilde{\eta}^{\otimes}_B) \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \tilde{\mathcal{D}}_n \operatorname{svec} C \\
&= \mathcal{A}_{svec} \tilde{X}^{\otimes}_{Q_B} \operatorname{svec} C.
\end{aligned} \tag{3.9}$$

Используя (3.8) и (3.9), получаем, что выражение (3.3) для $\Delta_{svec}(Z)$ может быть переписано в виде

$$\Delta_{svec}(Z) = D(\tilde{\eta}^{\otimes}_B) \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \tilde{\mathcal{D}}_n [I_{n_\triangle} - \mathcal{A}^T_{svec} \Gamma^{-1}(\tilde{X}^{\otimes}_{Q_B}) \mathcal{A}_{svec} \tilde{X}^{\otimes}_{Q_B}] \operatorname{svec} C,$$

из чего следует (3.7).

Лемма доказана.

На основании утверждения леммы 3.1 и формулы (3.4) приходим к выводу, что

$$\Delta^{Q_B}_{svec}(X) = \tilde{\mathcal{L}}_n (Q_B \otimes Q_B) \tilde{\mathcal{D}}_r D(\tilde{\eta}^{\otimes}_B) \tilde{\mathcal{L}}_r (Q^T_B \otimes Q^T_B) \tilde{\mathcal{D}}_n \mathcal{P}(\tilde{X}^{\otimes}_{Q_B}) \operatorname{svec} C. \tag{3.10}$$

Недостатком выражения (3.10) является несимметричный вид матрицы, стоящей перед вектором $\mathcal{P}(\tilde{X}^{\otimes}_{Q_B}) \operatorname{svec} C$. На самом деле ее можно представить и в симметричном виде.

Обозначим через $\mathcal{Q}_B$ матрицу $\mathcal{Q}_B = Q_B \otimes Q_B$ размера $n^2 \times r^2$. Пусть q_j — j-й столбец матрицы Q_B, где $j \in J^r$. Пусть, кроме того, $q^{(i)}_j$ — i-й элемент вектора q_j, где $i \in J^n$. Согласно определению произведения матриц по Кронекеру получаем, что элементами матрицы $\mathcal{Q}_B$ являются следующие величины: $\mathcal{Q}^{(s,t)(k,l)}_B = q^s_k q^t_l$, где $(s,t) \in J^n_\square$ и $(k,l) \in J^r_\square$. Здесь использованы парные номера из наборов (2.1).

Введем еще одно обозначение. Для произвольной квадратной матрицы M порядка r обозначим через $\widehat{M}$ квадратную матрицу того же порядка r, которая получается из M следующим образом. Все наддиагональные элементы матрицы $\widehat{M}$ — нулевые, а все диагональные и поддиагональные элементы $\widehat{M}$ совпадают с соответствующими элементами матрицы M, причем поддиагональные элементы удваиваются. Непосредственной проверкой можно убедиться, что $\widehat{M} = \tilde{\mathcal{L}}^T_r \tilde{\mathcal{L}}_r M$ (см. также [8]).

Лемма 3.2. *Пусть $M \in \mathbb{S}^r$. Тогда имеет место равенство*

$$\tilde{\mathcal{L}}_n \mathcal{Q}_B \operatorname{vec} M = \tilde{\mathcal{D}}_n^T \mathcal{Q}_B \operatorname{vec} \widehat{M}. \tag{3.11}$$

Доказательство. Умножая матрицу $\mathcal{Q}_B$ слева на матрицу $\tilde{\mathcal{L}}_n$, получаем результирующую матрицу $\tilde{\mathcal{L}}_n \mathcal{Q}_B$ размера $n_\triangle \times r^2$, в которой все строки $\mathcal{Q}_B$ с парными номерами $(s,t) \notin J^n_\triangle$ удаляются. Остальные строки $\mathcal{Q}_B$ сохраняются. Однако если их парные номера (s,t) из $J^n_\triangle$ внедиагональные, то эти строки умножается на коэффициент $\sqrt{2}$. Таким образом, для этих номеров выполняется условие

$$(\tilde{\mathcal{L}}_n \mathcal{Q}_B)^{(s,t)(k,l)} = \sqrt{2}\, q_k^{(s)} q_l^{(t)}, \qquad (k,l) \in J^r_\square. \tag{3.12}$$

С другой стороны, если умножить матрицу $\mathcal{Q}_B$ слева на матрицу $\tilde{\mathcal{D}}_n^T$, то получим опять матрицу размера $n_\triangle \times r^2$. У нее строки с парными диагональными номерами остаются прежними, т. е. такими же, как у матрицы $\mathcal{Q}_B$. Для строк с внедиагональными номерами (s,t) из $J^n_\triangle$ (младшими внедиагональными номерами) получаем

$$(\tilde{\mathcal{D}}_n^T \mathcal{Q}_B)^{(s,t)(k,l)} = 1/\sqrt{2}\, [q_k^{(s)} q_l^{(t)} + q_k^{(t)} q_l^{(s)}], \qquad (k,l) \in J^r_\square. \tag{3.13}$$

Пусть $J^n_{\triangle,d} = \{(s,s) \in J^n_\triangle \colon s \in J^n\}$, $J^n_{\triangle,nd} = J^n_\triangle \setminus J^n_{\triangle,d}$. Множества $J^n_{\triangle,d}$ и $J^n_{\triangle,nd}$ — подмножества индексного множества $J^n_\triangle$, состоящие соответственно из диагональных и младших внедиагональных номеров.

Для диагональных номеров (s,s), в которых $s \in J^n$, имеем

$$\begin{aligned}\sum_{(k,l)\in J^r_\square} q_k^{(s)} q_l^{(s)} (\operatorname{vec} M)^{(k,l)} &= \sum_{(k,l)\in J^r_{\triangle,dg}} q_k^{(s)} q_l^{(s)} (\operatorname{vec} M)^{(k,l)} + 2 \sum_{(k,l)\in J^r_{\triangle,ndg}} q_k^{(s)} q_l^{(s)} (\operatorname{vec} M)^{(k,l)} \\ &= \sum_{(k,l)\in J^r_\triangle} q_k^{(s)} q_l^{(s)} (\operatorname{vec} \widehat{M}\,)^{(k,l)}. \end{aligned} \tag{3.14}$$

Здесь $(\operatorname{vec} M)^{(k,l)}$ — (k,l)-й элемент вектора $\operatorname{vec} M$. Для внедиагональных номеров $(s,t) \in J^n_\triangle$ получаем соответственно

$$\begin{aligned}\sum_{(k,l)\in J^r_\square} [q_k^{(s)} q_l^{(t)} + q_k^{(t)} q_l^{(s)}] (\operatorname{vec} \widehat{M}\,)^{(k,l)} &= \sum_{(k,l)\in J^r_\triangle} [q_k^{(s)} q_l^{(t)} + q_k^{(t)} q_l^{(s)}] (\operatorname{vec} \widehat{M}\,)^{(k,l)} \\ &= 2 \sum_{(k,l)\in J^r_\square} q_k^{(s)} q_l^{(t)} (\operatorname{vec} M)^{(k,l)}. \end{aligned} \tag{3.15}$$

Из (3.14) и (3.15) с учетом (3.12), (3.13) приходим к равенству (3.11).

Лемма доказана.

Утверждение 3.1. *Для вектора $\Delta^{Q_B}_{svec}(X)$ справедливо выражение*

$$\Delta^{Q_B}_{svec}(X) = \tilde{X}^{\otimes}_{Q_B} \mathcal{P}(\tilde{X}^{\otimes}_{Q_B}) \operatorname{svec} C. \tag{3.16}$$

Доказательство. Поскольку вектор $\tilde{\mathcal{D}}_r D(\tilde{\eta}^{\otimes}_B) \tilde{\mathcal{L}}_r (Q_B^T \otimes Q_B^T) \tilde{\mathcal{D}}_n \mathcal{P}(X_{Q_B}) \operatorname{svec} C$ есть прямая сумма столбцов некоторой симметричной матрицы M порядка r, то, заменяя матрицу $\tilde{\mathcal{D}}_r$ на матрицу $\tilde{\mathcal{L}}_r^T$, вместо $\operatorname{vec} M$ будем иметь $\operatorname{vec} \widehat{M}$. Поэтому согласно утверждению леммы 3.2 наряду с (3.10) справедливо представление

$$\Delta^{Q_B}_{svec}(X) = \tilde{\mathcal{D}}_n^T (Q_B \otimes Q_B) \tilde{\mathcal{L}}_r^T D(\tilde{\eta}^{\otimes}_B) \tilde{\mathcal{L}}_r (Q_B^T \otimes Q_B^T) \tilde{\mathcal{D}}_n \mathcal{P}(X_{Q_B}) \operatorname{svec} C,$$

которое с учетом введенного обозначения (3.5) записывается как (3.16).

Утверждение доказано.

Отметим, что для направления $\Delta^{Q_B}(X)$ в пространстве $\mathbb{S}^n$ сохраняются равенства и неравенство из утверждения 2.2. Кроме того, из приведенных выкладок следует, что зависимость $u(Z)$ также можно свести к зависимости $u(X)$, а именно

$$u = u(X) = \Gamma^{-1}(\tilde{X}^{\otimes}_{Q_B})\mathcal{A}_{svec}\,\tilde{X}^{\otimes}_{Q_B}\,\mathrm{svec}\,C. \tag{3.17}$$

Недостатком направления $\Delta^{Q_B}(X)$ является невозможность, двигаясь вдоль него, покинуть минимальную грань $\mathcal{G}_{\min}(X;\mathcal{F}_P)$. Поэтому, чтобы преодолеть этот недостаток, следует несколько изменить подход к выбору направления перемещения.

4. Возмущенное направление

Пусть по-прежнему точка X принадлежит границе $\mathcal{F}_P$ и $V(X)$ не есть положительно полуопределенная матрица. Для $V(X)$ также имеет место разложение $V(X) = HD(\theta)H^T$, где H — ортогональная матрица со столбцами h_i, $i \in J^n$, $\theta \in \mathbb{R}^n$ — вектор собственных значений $V(X)$. Считаем, что ранг матрицы $V(X)$ равен $s \le n$ и первым s столбцам $h_1, h_2, \ldots, h_s$ матрицы H соответствуют ненулевые собственные значения $V(X)$. Тогда

$$V(X) = \sum_{j=1}^{s} \theta_j H_j, \qquad H_j = h_j h_j^T. \tag{4.1}$$

В силу сделанного предположения относительно матрицы $V(X)$ среди ее собственных значений найдется хотя бы одно отрицательное. Пусть это будет значение θ^{j_*}. Ему соответствует собственный вектор h_{j_*}. Предположим, что h_{j_*} не принадлежит подпространству $\mathcal{L}(Q_B)$, порожденному столбцами матрицы Q_B. Тогда согласно (4.1) $V^{h_{j_*}}(X) = h_{j_*}^T V(X) h_{j_*} = \theta^{j_*} < 0$.

Возьмем далее некоторое $\varepsilon > 0$. Теперь вместо направления $\Delta^{Q_B}(X)$, принадлежащего минимальной грани, нас будет интересовать направление

$$\Delta^{Q_B,\,h_{j_*}}(X) = \Delta^{Q_B}(X) + \varepsilon\,\Delta^{h_{j_*}}(X), \tag{4.2}$$

где

$$\Delta^{Q_B}(X) = Q_B \Lambda_B Q_B^T, \qquad \Lambda_B = D(\eta_B) \circ V^{Q_B}(\tilde{u}), \qquad \Delta^{h_{j_*}}(X) = V^{h_{j_*}}(\tilde{u}) H_{j_*}. \tag{4.3}$$

Здесь первое направление $\Delta^{Q_B}(X)$ принадлежит грани $\mathcal{G}_{\min}(X;\mathbb{S}^n_+)$, содержащей точку X. Поскольку $h_{j_*} \notin \mathcal{L}(Q_B)$, второе направление $\Delta_{h_{j_*}}(X)$ выводит за пределы этой грани. Вектор двойственных переменных $\tilde{u}$ теперь будем искать, исходя из условий, что для возмущенного вектора перемещений $\Delta(X) = \Delta^{Q_B,h_{j_*}}(X)$ сохраняются равенства $A_i \bullet \Delta(X) = 0$, $i \in J^m$.

Пусть $\tilde{G}_{h_{j_*}} = (h_{j_*}^T \otimes h_{j_*}^T)\tilde{\mathcal{D}}_n$ и $\tilde{E}^{\otimes}_{h_{j_*}} = \tilde{G}^T_{h_{j_*}} \tilde{G}_{h_{j_*}}$. Введем обозначения:

$$\psi = \tilde{G}_{h_{j_*}} \mathcal{A}^T_{svec} \Gamma^{-1}(\tilde{X}^{\otimes}_{Q_B})\mathcal{A}_{svec}\,\tilde{G}^T_{h_{j_*}} > 0, \qquad c(\varepsilon,\psi) = (1+\varepsilon\psi)^{-1} > 0.$$

Матрица $\tilde{E}^{\otimes}_{h_{j_*}}$ — симметричная порядка $n_\triangle$.

Лемма 4.1. *Пусть X — регулярная граничная точка $\mathcal{F}_P$ и собственный вектор h_{j_*} матрицы $V(X)$ не принадлежит $\mathcal{L}(Q_B)$. Тогда для вектора $\tilde{u}$ справедливо разложение $\tilde{u} = u + \varepsilon\Delta u$, где u определяется согласно* (3.17) *и*

$$\Delta u = c(\varepsilon,\psi)\,\Gamma^{-1}(\tilde{X}^{\otimes}_{Q_B})\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}}\,\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\,\mathrm{svec}\,C. \tag{4.4}$$

Д о к а з а т е л ь с т в о. После подстановки выражения для $\Delta^{Q_B,h_{j_*}}(X)$ из (4.2) имеем

$$A_i \bullet [Q_B(D(\eta_B) \circ V^{Q_B}(\tilde{u}))Q_B^T + \varepsilon\, V^{h_{j_*}}(\tilde{u}) H_{j_*}] = 0, \qquad i \in J^m. \tag{4.5}$$

Представим равенства (4.5) в векторном виде. Прежде всего выразим в векторном виде первое слагаемое из левой части (4.5):

$$\langle \mathrm{svec}\, A_i^{Q_B}, \mathrm{svec}\,(D(\eta_B) \circ V^{Q_B}(\tilde{u}))\rangle = \langle \mathrm{svec}\, A_i^{Q_B}, D(\tilde{\eta}_B^{\otimes})\, \mathrm{svec}\, V^{Q_B}(\tilde{u})\rangle. \tag{4.6}$$

Если воспользоваться обозначением (3.5) для матрицы $\tilde{X}^{\otimes}_{Q_B}$, то, объединяя все правые части в (4.6), получаем, что эти правые части можно записать в виде

$$\mathcal{A}^{Q_B}_{svec}\, D^{\otimes}(\tilde{\eta}_B)\, \mathrm{svec}\, V^{Q_B}(\tilde{u}) = \mathcal{A}_{svec}\, \tilde{X}^{\otimes}_{Q_B}\, \mathrm{svec}\, C - \mathcal{A}_{svec}\, \tilde{X}^{\otimes}_{Q_B} (\mathcal{A}_{svec})^T \tilde{u}.$$

Далее, поскольку

$$V^{h_{j*}}(\tilde{u}) = C^{h_{j*}} - \sum_{i=1}^{m} \tilde{u}^i A_i^{h_{j*}} = h_{j*}^T C h_{j*} - \sum_{i=1}^{m} \tilde{u}^i h_{j*}^T A_i h_{j*},$$

то $V^{h_{j*}}(\tilde{u}) = (h_{j*}^T \otimes h_{j*}^T)(\mathrm{vec}\, C - \mathcal{A}_{vec}^T \tilde{u})$, или, с учетом симметричности матриц,

$$V^{h_{j*}}(\tilde{u}) = (h_{j*}^T \otimes h_{j*}^T)\tilde{\mathcal{D}}_n(\mathrm{svec}\, C - \mathcal{A}_{svec}^T \tilde{u}).$$

Учтем также, что

$$A_i \bullet H_{j*} = \mathrm{tr}\,(A_i\, H_{j*}) = \mathrm{tr}(h_{j*}^T A_i h_{j*}) = A_i^{h_{j*}} = (h_{j*}^T \otimes h_{j*}^T)\tilde{\mathcal{D}}_n\, \mathrm{svec}\, A_i,$$

откуда $\mathcal{A}^{h_{j*}}_{vec} = \mathcal{A}_{svec}\, \tilde{\mathcal{D}}_n^T (h_{j*} \otimes h_{j*})$. Следовательно,

$$\mathcal{A}^{h_{j*}}_{svec}\, \mathrm{svec}\, V^{h_{j*}}(\tilde{u}) = \mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}} (\mathrm{svec}\, C - \mathcal{A}_{svec}^T \tilde{u}).$$

Таким образом, система уравнений (4.5) относительно вектора $\tilde{u}$ принимает вид

$$[\mathcal{A}_{svec}\, \tilde{X}^{\otimes}_{Q_B} \mathcal{A}^T_{svec} + \varepsilon\, \mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}} \mathcal{A}^T_{svec}]\tilde{u} = [\mathcal{A}_{svec}\, \tilde{X}^{\otimes}_{Q_B} + \varepsilon\, \mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}}]\mathrm{svec}\, C. \tag{4.7}$$

Обозначим: $\tilde{\Gamma}(\tilde{X}^{\otimes}_{Q_B}, \tilde{E}^{\otimes}_{h_{j*}}) = \Gamma(\tilde{X}^{\otimes}_{Q_B}) + \Gamma_p(\tilde{E}^{\otimes}_{h_{j*}})$, где матрица $\Gamma(\tilde{X}^{\otimes}_{Q_B})$ определена в (3.8), а вторая матрица $\Gamma_p(\tilde{E}^{\otimes}_{h_{j*}})$ имеет вид $\Gamma_p(\tilde{E}^{\otimes}_{h_{j*}}) = \varepsilon\, \mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}} \mathcal{A}^T_{svec}$. Матрица $\tilde{\Gamma}(\tilde{X}^{\otimes}_{Q_B}, \tilde{E}_{h^{\otimes}_{j*}})$ неособая, поскольку в силу регулярности точки X матрица $\Gamma(\tilde{X}^{\otimes}_{Q_B})$ положительно определенная. Тогда, разрешая систему уравнений (4.7), находим

$$\tilde{u} = \tilde{\Gamma}^{-1}(\tilde{X}^{\otimes}_{Q_B}, \tilde{E}^{\otimes}_{h_{j*}})[\mathcal{A}_{svec}\, \tilde{X}^{\otimes}_{Q_B} + \varepsilon\, \mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}}]\mathrm{svec}\, C.$$

Применим формулу Шермана—Моррисона—Вудберри для обращения матрицы $\tilde{\Gamma}(\tilde{X}^{\otimes}_{Q_B}, \tilde{E}^{\otimes}_{h_{j*}})$. Полагая для упрощения записи $\Gamma = \Gamma(\tilde{X}^{\otimes}_{Q_B})$, имеем

$$\tilde{\Gamma}^{-1} = \Gamma^{-1} - \varepsilon(1 + \varepsilon\psi)^{-1}\Gamma^{-1}\mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}} \mathcal{A}^T_{svec}\, \Gamma^{-1}.$$

Обозначим далее для сокращения записи $F_1 = \mathcal{A}_{svec}\, \tilde{X}^{\otimes}_{Q_B} \mathrm{svec}\, C$, $F_2 = \mathcal{A}_{svec}\, \tilde{E}_{h_{j*}}\, \mathrm{svec}\, C$. Тогда

$$\tilde{u} = \Gamma^{-1}F_1 - \varepsilon c(\varepsilon,\psi)\Gamma^{-1}\mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}} \mathcal{A}^T_{svec}\, \Gamma^{-1}F_1 + \varepsilon\, \Gamma^{-1}F_2 - \varepsilon^2 c(\varepsilon,\psi)\Gamma^{-1}\mathcal{A}_{svec}\, E^{\otimes}_{h_{j*}} \mathcal{A}^T_{svec}\, \Gamma^{-1}F_2$$

или, с учетом выражения для $u = \Gamma^{-1}F_1$, имеем

$$\tilde{u} = \{I_m - \varepsilon c(\varepsilon,\psi)\Gamma^{-1}\mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}} \mathcal{A}^T_{svec}\}u + \varepsilon\, \Gamma^{-1}\{I_m - \varepsilon c(\varepsilon,\psi)\mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}} \mathcal{A}^T_{svec}\, \Gamma^{-1}\}F_2.$$

Если, кроме того, принять во внимание выражение для F_2, то получаем

$$\tilde{u} = \{I_m - \varepsilon c(\varepsilon,\psi)\Gamma^{-1}\mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}} \mathcal{A}^T_{svec}\}u + \varepsilon\, \Gamma^{-1}\{\mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}} - \varepsilon\psi c(\varepsilon,\psi)\mathcal{A}_{svec}\, \tilde{E}^{\otimes}_{h_{j*}}\}\mathrm{svec}\, C$$

или

$$\tilde{u} = \{I_m - \varepsilon c(\varepsilon,\psi)\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}}\,\mathcal{A}^T_{svec}\}u + \varepsilon c(\varepsilon,\psi)\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}}\,\mathrm{svec}\,C. \tag{4.8}$$

Из (4.8) приходим к выводу, что для $\tilde{u}$ справедливо разложение $\tilde{u} = u + \varepsilon\Delta u$, в котором

$$\Delta u = c(\varepsilon,\psi)[\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}} - \Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}}\,\mathcal{A}^T_{svec}\,\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{X}^{\otimes}_{Q_B}]\,\mathrm{svec}\,C,$$

причем Δu с помощью матрицы $\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})$, введенной в (3.6), можно записать в виде (4.4).

Лемма доказана.

Найдем далее выражение для $\Delta^{Q_B,h_{j_*}}(X)$.

Утверждение 4.1. *Пусть выполнены условия леммы* 4.1. *Тогда направление* $\Delta^{Q_B,h_{j_*}}(X)$ *в векторном представлении имеет вид*

$$\Delta^{Q_B,h_{j_*}}_{svec}(X) = [\tilde{X}^{\otimes}_{Q_B}\mathcal{P}(\tilde{X}^{\otimes}_{Q_B}) + \tilde{c}(\varepsilon,\psi)\mathcal{P}^T(\tilde{X}^{\otimes}_{Q_B})\tilde{E}^{\otimes}_{h_{j_*}}\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})]\,\mathrm{svec}C, \tag{4.9}$$

где $\tilde{c}(\varepsilon,\psi) = \varepsilon c(\varepsilon,\psi) > 0$.

Д о к а з а т е л ь с т в о. Используя (4.3), (3.5) и учитывая утверждения лемм 3.2 и 4.1, имеем

$$\begin{aligned}\Delta^{Q_B}_{svec}(X) &= \tilde{\mathcal{L}}_n(Q_B\otimes Q_B)\,\mathrm{vec}(D(\eta_B)\circ V^{Q_B}(\tilde{u})) = \tilde{\mathcal{L}}_n(Q_B\otimes Q_B)D^{\otimes}(\eta_B)\,\mathrm{vec}\,V^{Q_B}(\tilde{u})\\ &= \tilde{\mathcal{L}}_n(Q_B\otimes Q_B)D^{\otimes}(\eta_B)(Q_B^T\otimes Q_B^T)\tilde{\mathcal{D}}_n\,\mathrm{svec}\,V(\tilde{u})\\ &= \tilde{\mathcal{L}}_n(Q_B\otimes Q_B)\tilde{\mathcal{D}}_r\tilde{\mathcal{L}}_r D^{\otimes}(\eta_B)\mathcal{D}_r\tilde{\mathcal{L}}_r(Q_B^T\otimes Q_B^T)\tilde{\mathcal{D}}_n\,\mathrm{svec}\,V(\tilde{u})\\ &= \tilde{\mathcal{L}}_n(Q_B\otimes Q_B)\tilde{\mathcal{D}}_r D(\tilde{\eta}^{\otimes}_B)\tilde{\mathcal{L}}_r(Q_B^T\otimes Q_B^T)\tilde{\mathcal{D}}_n\,\mathrm{svec}\,V(\tilde{u}) = \tilde{X}^{\otimes}_{Q_B}(\mathrm{svec}\,V(u) - \varepsilon\mathcal{A}^T_{svec}\Delta u)\\ &= \tilde{X}^{\otimes}_{Q_B}[I_{n_\triangle} - \varepsilon c(\varepsilon,\psi)\mathcal{A}^T_{svec}\,\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}}]\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\,\mathrm{svec}\,C.\end{aligned} \tag{4.10}$$

Соответственно для $\Delta^{h_{j_*}}(X)$ получаем

$$\begin{aligned}\Delta^{h_{j_*}}_{svec}(X) &= \mathrm{svec}\,(h_{j_*}V^{h_{j_*}}(\tilde{u})\,h^T_{j_*}) = \tilde{\mathcal{L}}_n(h_{j_*}\otimes h_{j_*})\,\mathrm{vec}(h^T_{j_*}V(\tilde{u})h_{j_*})\\ &= \tilde{\mathcal{L}}_n(h_{j_*}\otimes h_{j_*})(h^T_{j_*}\otimes h^T_{j_*})\tilde{\mathcal{D}}_n\,\mathrm{svec}\,V(\tilde{u}) = \tilde{\mathcal{L}}_n(h_{j_*}\otimes h_{j_*})(h^T_{j_*}\otimes h^T_{j_*})\tilde{\mathcal{D}}_n[\mathrm{svec}\,V(u) - \varepsilon\mathcal{A}^T_{svec}\,\Delta u]\\ &= \tilde{\mathcal{L}}_n(h_{j_*}\otimes h_{j_*})(h^T_{j_*}\otimes h^T_{j_*})\tilde{\mathcal{D}}_n[I_{n_\triangle} - \varepsilon c(\varepsilon,\psi)\mathcal{A}^T_{svec}\,\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}}]\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\,\mathrm{svec}\,C\\ &= \tilde{\mathcal{D}}^T_n(h_{j_*}\otimes h_{j_*})(h^T_{j_*}\otimes h^T_{j_*})\tilde{\mathcal{D}}_n[I_{n_\triangle} - \varepsilon c(\varepsilon,\psi)\mathcal{A}^T_{svec}\,\Gamma^{-1}\mathcal{A}_{svec}\tilde{E}^{\otimes}_{h_{j_*}}]\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\,\mathrm{svec}\,C\\ &= \tilde{E}^{\otimes}_{h_{j_*}}[I_{n_\triangle} - \varepsilon c(\varepsilon,\psi)\mathcal{A}^T_{svec}\,\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}}]\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\,\mathrm{svec}\,C.\end{aligned}$$

Так как $\tilde{E}^{\otimes}_{h_{j_*}}\,\mathcal{A}^T_{svec}\,\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}} = \tilde{G}^T_{h_{j_*}}\,\tilde{G}_{h_{j_*}}\,\mathcal{A}^T_{svec}\,\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{G}^T_{h_{j_*}}\,\tilde{G}_{h_{j_*}} = \psi\tilde{E}^{\otimes}_{h_{j_*}}$, поэтому $\tilde{E}^{\otimes}_{h_{j_*}}[I_{n_\triangle} - \varepsilon c(\varepsilon,\psi)\mathcal{A}^T_{svec}\,\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}}] = (1 - \varepsilon\psi c(\varepsilon,\psi))\tilde{E}^{\otimes}_{h_{j_*}} = c(\varepsilon,\psi)\tilde{E}^{\otimes}_{h_{j_*}}$ и, следовательно,

$$\Delta^{h_{j_*}}_{svec}(X) = c(\varepsilon,\psi)\tilde{E}^{\otimes}_{h_{j_*}}\,\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\,\mathrm{svec}\,C. \tag{4.11}$$

Из (4.10) и (4.11) приходим к (4.9).

Утверждение доказано.

Отметим, что для соответствующего направления $\Delta(X) = \Delta^{Q_B,h_{j_*}}(X)$ в пространстве $\mathbb{S}^n$ сохраняются равенства и неравенство из утверждения 2.2.

5. Итерационный процесс

Рассмотрим обобщение итерационного процесса (2.10), в котором помимо $\Delta_{svec}(X)$ используются направления $\Delta_{svec}^{Q_B}(X)$ или $\Delta_{svec}^{Q_B,h_{j_*}}(X)$.

Пусть теперь направление (2.9) берется в качестве $\Delta_{svec}(X)$ только в том случае, когда $X \succ 0$, т. е. когда точка X принадлежит относительной внутренности множества $\mathcal{F}_P$. В противном случае в качестве $\Delta_{svec}(X)$ берем направление $\Delta_{svec}^{Q_B}(X)$, причем если существует такой собственный вектор h_{j_*} матрицы $V(X)$, которому соответствует отрицательное собственное значение, и при этом h_{j_*} не принадлежит минимальной грани $\mathcal{G}_{\min}(X;\mathbb{S}^n_+)$, то полагаем $\Delta_{svec}(X) = \Delta_{svec}^{Q_B,h_{j_*}}(X)$.

Обозначим через $\mathrm{Matr}(\mathrm{svec}\, M)$ симметричную матрицу M, которой соответствует вектор $\mathrm{svec}\, M$, и рассмотрим теперь вопрос о том, будет ли матрица $\mathrm{Matr}(\mathrm{svec}\ \mathrm{X} - \alpha\Delta_{svec}^{Q_B,h_{j_*}}(X))$ положительно полуопределенной по крайней мере для α достаточно малых.

Предполагая, что точка X является регулярной, разобьем вектор $\Delta_{svec}^{Q_B,h_{j_*}}(X)$ на два слагаемых: $\Delta_{svec}^{Q_B,h_{j_*}}(X) = \Delta_{svec}^{(1)}(X) + \Delta_{svec}^{(2)}(X)$, где

$$\Delta_{svec}^{(1)}(X) = \tilde{X}^{\otimes}_{Q_B}[I_{n_\triangle} - \tilde{c}(\varepsilon,\psi)\mathcal{A}^T_{svec}\,\Gamma^{-1}\mathcal{A}_{svec}\,\tilde{E}^{\otimes}_{h_{j_*}}]\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\,\mathrm{svec}\, C,$$

$$\Delta_{svec}^{(2)}(X) = \tilde{c}(\varepsilon,\psi)\tilde{E}^{\otimes}_{h_{j_*}}\,\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\,\mathrm{svec}\, C.$$

Им соответствуют направления $\Delta^{(1)}(X)$ и $\Delta^{(2)}(X)$ в матричном пространстве $\mathbb{S}^n$. Первое направление $\Delta^{(1)}(X)$ принадлежит грани $\mathcal{G}_{\min}(X;\mathcal{S}^n_+)$, второе направление $\Delta^{(2)}(X)$ выводит за пределы этой грани.

Имеем $\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\mathrm{svec}\, C = \mathrm{svec}\, V(u)$. Следовательно, $\tilde{G}_{h_{j_*}}\tilde{\mathcal{D}}_n\,\mathrm{svec}\, V(u) = \mathrm{vec}\, V^{h_{j_*}}(u) = \theta_{j_*} < 0$. Поэтому $\tilde{E}^{\otimes}_{h_{j_*}}\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})\,\mathrm{svec}\, C = \tilde{E}^{\otimes}_{h_{j_*}}\mathrm{svec}\, V(u) = \theta_{j_*}\tilde{\mathcal{D}}^T_n\,\mathrm{svec}\, H_{j_*} = \theta_{j_*}\tilde{\mathcal{L}}_n\,\mathrm{svec}\, H_{j_*}$. Двигаясь из точки X с шагом $\alpha > 0$ вдоль направления, обратному к направлению $\Delta(X) = \Delta^{(1)}(X) + \Delta^{(2)}(X)$, получаем

$$X(\alpha) = X - \alpha\Delta(X) = Q_B\,(Z - \alpha\Delta(Z))\,Q_B^T - \alpha\tilde{c}(\varepsilon,\psi)\theta_{j_*}H_{j_*}, \tag{5.1}$$

где $Z = D(\eta_B)$, а $\Delta(Z)$ — симметричная матрица порядка r, которой соответствует вектор

$$\Delta_{vec}(Z) = D(\eta_B^{\otimes})(Q_B^T\otimes Q_B^T)\tilde{\mathcal{D}}_n[I_{n_\triangle} - \tilde{c}(\varepsilon,\psi)\mathcal{A}^T_{vec}\,\Gamma^{-1}\mathcal{A}_{vec}\,\tilde{E}^{\otimes}_{h_{j_*}}]\,\mathrm{svec}\, V(u).$$

Понятно, что матрица $Z - \alpha\Delta(Z)$ при α достаточно малых будет оставаться положительно определенной. Так как H_{j_*} — положительно полуопределенная матрица единичного ранга и $\theta_{j_*} < 0$, то второе слагаемое в (5.1) также будет оставаться положительно полуопределенной матрицей, причем для любого $\alpha > 0$. Таким образом, максимально возможный шаг α_k определяется из условия, когда положительно определенная матрица $Z - \alpha\Delta(Z)$ при увеличении α впервые становится положительно полуопределенной.

Такой шаг α_k на k-й итерации можно найти, если воспользоваться, например, коньюктивным приведением двух симметричных матриц Z_k и $\Delta(Z_k)$, одна из которых является положительно определенной. Тогда, взяв некоторую невырожденную матрицу P порядка r, получаем

$$P^TZ_kP = I_n, \qquad P^T\Delta(Z_k)P = D(\omega_k), \tag{5.2}$$

где $\omega_k = [\omega_k^1,\dots,\omega_k^r]$ — вектор, составленный из собственных чисел матрицы $\hat{Z}_k = Z_k^{-1}\Delta(Z_k)$. Все числа $\omega_k^1,\dots,\omega_k^r$ действительные. Согласно (5.2) имеем $\Delta(Z_k) = (P^{-1})^TD(\omega_k)P^{-1}$ и $Z_k = (P^{-1})^TP^{-1}$. Поэтому

$$Z_{k+1} = (P^{-1})^T[I_k - \alpha_kD(\omega_k)]P^{-1} = (P^{-1})^TD(e - \alpha_k\omega_k)P^{-1},$$

где e — r-мерный вектор, состоящий из единиц.

Пусть $\omega_k^{\max}$ — максимальное положительное число из набора собственных чисел $\omega_k^1, \ldots, \omega_k^n$ и $\hat{\alpha}_k = 1/\omega_k^{\max}$. В качестве α_k берем $\hat{\alpha}_k$. Отсутствие такого конечного положительного числа $\omega_k^{\max}$ означает, что задача (1.1) не имеет решения.

Если все точки траектории оказываются регулярными, то итерационный процесс (2.10) полностью определен. Более того, если сделать дополнительное предположение, что $m = r_\triangle$ для некоторого $r > 0$ $\big($такое условие выполняется, в частности, когда аффинное многообразие $\mathcal{N}_A$ задается через линейное отображение одного матричного пространства в другое матричное пространство (см., например, [9])$\big)$, то в тех случаях, когда текущая регулярная точка X_k оказывается крайней, последующая точка X_{k+1} также будет крайней, т. е. метод начинает себя вести как симплекс-метод.

Покажем, что рассмотренный итерационный процесс позволяет для любой стартовой точки $X_0 \in \mathcal{F}_P$ попасть в некоторую окрестность $\mathcal{S}(X_*)$ решения X_* задачи (1.1), где уже можно применять основной вариант метода (2.10). Обозначим $\mathcal{U}_f = \{X \in \mathcal{F}_P \colon C \bullet X \le f\}$. Если f_* — оптимальное значение целевой функции в задаче (1.1) и множество $\mathcal{U}_{f_*}$ состоит из единственной точки (например, если для решений прямой и двойственной задач выполнено условие строгой дополнительности), то всегда можно указать $\tilde{f} > f_*$ такое, что $\mathcal{U}_{\tilde{f}} \subseteq \mathcal{S}(X_*)$. Считаем, что все точки траектории $\{X_k\}$ являются регулярными.

Теорема 5.1. *Пусть точка $X_0 \in \mathcal{F}_P$ такова, что множество $\mathcal{U}_{f_0}$ ограничено. Тогда можно указать номер $K \ge 1$ такой, что $X_k \in \mathcal{U}_{\tilde{f}}$ для $k \ge K$.*

Д о к а з а т е л ь с т в о. Предположим, что итерационный процесс генерирует бесконечную последовательность точек, ни одна из которых не принадлежит множеству $\mathcal{U}_{\tilde{f}}$. Так как множество $\mathcal{U}_{f_0}$ ограничено, то у последовательности $\{X_k\}$ существуют предельные точки. Пусть $X_{k_l} \to \bar{X}$.

Считаем для общности, что в качестве направлений перемещения используются возмущенные направления $\Delta_{svec}^{Q_B, h_{j*}}(X)$, вычисленные в соответствующих точках X_{k_l} (в противном случае следует в качестве $\tilde{E}^{\otimes}_{h_{j*}}$ взять нулевую матрицу). Обе матрицы $\tilde{X}^{\otimes}_{Q_B}\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})$ и $\mathcal{P}^T(\tilde{X}^{\otimes}_{Q_B})\tilde{E}^{\otimes}_{h_{j*}}\mathcal{P}(\tilde{X}^{\otimes}_{Q_B})$ в (4.9) — симметричные и положительно полуопределенные. Симметричность первой матрицы следует из разложения $\tilde{X}^{\otimes}_{Q_B}\mathcal{P}(\tilde{X}^{\otimes}_{Q_B}) = (\tilde{X}^{\otimes}_{Q_B})^{1/2}\tilde{\mathcal{P}}(\tilde{X}^{\otimes}_{Q_B})(\tilde{X}^{\otimes}_{Q_B})^{1/2}$. Здесь

$$\tilde{\mathcal{P}}(\tilde{X}^{\otimes}_{Q_B}) = I_{n_\triangle} - (\tilde{X}^{\otimes}_{Q_B})^{1/2}\mathcal{A}^T_{svec}\Gamma^{-1}(\tilde{X}^{\otimes}_{Q_B})\mathcal{A}_{svec}(\tilde{X}^{\otimes}_{Q_B})^{1/2}$$

— матрица ортогонального проектирования.

Так как все матрицы, входящие в (4.9)), ограничены в совокупности по норме в некоторой окрестности точки $\bar{X}$, то длина шагов α_{k_l} не может стремиться к нулю, т. е. они ограничены снизу положительным числом. Кроме того, последовательность значений целевой функции $\{C \bullet X_{k_l}\}$ монотонно убывающая и ограниченная снизу значением $C \bullet \bar{X}$. Поэтому в предельной точке $X = \bar{X}$ должно выполняться

$$\|\tilde{\mathcal{P}}(\tilde{X}^{\otimes}_{Q_B})(\tilde{X}^{\otimes}_{Q_B})^{1/2}\mathrm{svec}\, C\|_F = 0, \tag{5.3}$$

где $\|M\|_F$ — норма Фробениуса матрицы M, причем в (5.3) ортогональная матрица Q из (2.12), а также ее разбиение на подматрицы Q_B и Q_N соответствуют точке $\bar{X}$. Отсюда следует, что вектор $(\tilde{X}^{\otimes}_{Q_B})^{1/2}\mathrm{svec}\, C$ принадлежит пространству столбцов матрицы $(\tilde{X}^{\otimes}_{Q_B})^{1/2}\mathcal{A}^T_{svec}$, т. е. можно указать такое $\bar{u} \in \mathbb{R}^m$, что

$$(\tilde{X}^{\otimes}_{Q_B})^{1/2}\mathrm{svec}\, C = (\tilde{X}^{\otimes}_{Q_B})^{1/2}\mathcal{A}^T_{svec}\bar{u}. \tag{5.4}$$

Равенство (5.4) сохранится и после умножения его обеих частей на матрицу $(\tilde{X}^{\otimes}_{Q_B})^{1/2}$. Тогда его можно переписать в виде

$$\tilde{\mathcal{D}}_n(Q_B \otimes Q_B)\tilde{\mathcal{L}}_r^T D(\tilde{\eta}^{\otimes}_B)[\mathrm{svec}\, C^{Q_B} - (\mathcal{A}^{Q_B}_{svec})^T\bar{u}] = 0_{n_\triangle}. \tag{5.5}$$

Поскольку матрица, стоящая в (5.5) перед квадратной скобкой, имеет полный ранг по столбцам, то это означает, что $\operatorname{svec} V^{Q_B}(\bar{u}) = 0_{r_\triangle}$. Из-за того что матрица C не принадлежит линейному подпространству, порожденному матрицами A_i, $i \in J^m$, это возможно только в том случае, когда $r_\triangle \le m$, где r — ранг матрицы $\bar{X}$. Таким образом, $\bar{X}$ — крайняя точка множества $\mathcal{F}_P$.

Пусть $M^Q = Q^T M Q$. Имеем $\bar{X}^Q \succeq 0$. Кроме того, из $\bar{X}^Q \bullet V^Q(\bar{u}) = 0$ следует, что $\bar{X} \bullet \bar{V} = 0$, где $\bar{V} = \bar{V}(\bar{u})$. Но точка $\bar{X}$ не есть решение задачи, поэтому у матрицы $\bar{V}$ существует отрицательное собственное значение θ_{j_*}, причем соответствующий собственный вектор h_{j_*} не принадлежит подпространству $\mathcal{L}(Q_B)$, ибо иначе матрица $\bar{V}^{Q_B}$ была ненулевой. Далее, так как $\operatorname{svec} \bar{V} = \mathcal{P}(\tilde{X}^{\otimes}_{Q_B}) \operatorname{svec} C$, то $\langle \operatorname{svec} C, \mathcal{P}^T(\tilde{X}^{\otimes}_{Q_B}) \tilde{E}^{\otimes}_{h_{j_*}} \mathcal{P}(\tilde{X}^{\otimes}_{Q_B}) \operatorname{svec} C \rangle = \theta_{j_*} < 0$. Это приводит к тому, что при X_{k_l}, достаточно близких к $\bar{X}$, значение функции $C \bullet X$ убывает на величину, ограниченную снизу некоторым положительным числом. В результате на некоторой итерации должно стать $C \bullet X_{k_l} < C \bullet \bar{X}$, что невозможно. Поэтому $X_k \in \mathcal{U}_{\tilde{f}}$ для k достаточно больших.

Теорема доказана.

З а м е ч а н и е. В силу произвольности выбора уровня $\tilde{f}$ можно утверждать, что описанный итерационный процесс сходится глобально на множестве $\mathcal{F}_P$.

СПИСОК ЛИТЕРАТУРЫ

1. **Еремин И. И.** Теория линейной оптимизации. Екатеринбург: Изд-во "Екатеринбург", 1999. 312 с.
2. Handbook of Semidefinite Programming / eds. H. Wolkowicz, R. Saigal, L. Vandenberghe. Dordrecht: Kluwer Acad. Publ., 2000. 656 p.
3. **Бабынин М. С., Жадан В. Г.** Прямой метод внутренней точки для линейной задачи полуопределенного программирования // Журн. вычисл. математики и мат. физики. 2008. Т. 48, № 10. С. 1780–1801.
4. **Жадан В. Г.** Прямой мультипликативно-барьерный метод с наискорейшим спуском для линейной задачи полуопределенного программирования // Оптимизация и приложения. Вып. 2. М.: ВЦ РАН, 2011. С. 107–131.
5. **Жадан В. Г.** Конечные барьерно-проективные методы для линейного программирования // Методы оптимизации и их приложения: тр. XI Байкальской междунар. шк.-семинара. Пленарные доклады. Иркутск: Изд-во СЭИ СО РАН, 1998. С. 140–144.
6. **Арнольд В. И.** О матрицах, зависящих от параметров // Успехи мат. наук. 1971. Т. 26, вып. 2(158). С. 101–114.
7. **Alizadeh F., Haeberly J.-P. F., Overton M. L.** Complementarity and nondegeneracy in semidefinite programming // Math. Programming. Ser. B. 1997. Vol. 77, no. 1. P. 111–128.
8. **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.
9. **Lasserre J.B.** Linear programming with positive semi-definite matreces // Math. Problems in Engineering. 1996. Vol. 2, iss. 6. P. 499–522.

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

Поступила 01.02.2014