ЛИТЕРАТУРА

1. Rutishauser H. On Jacobi rotation patterns.— Proceedings A. M. S. Symposia in Applied Mathematics, 1963, v. 15, p. 219—239.
2. Golub G., Kahan W. Calculating the singular values and pseudoinverse of a matrix.— J. SIAM Numer. Anal. Ser. B, 1965, v. 2, N 2, p. 205—224.
3. Уилкинсон Дж. Х. Алгебраическая проблема собственных значений.— М.: Наука, 1970.— 564 с.
4. Годунов С. К., Костин В. И., Митченко А. Д. Вычисление собственного вектора симметрической трехдиагональной матрицы.— Сиб. мат. журн., 1985, т. XXVI. № 5.
5. Годунов С. К. Решение систем линейных уравнений.— Новосибирск: Наука. Сиб. отд-ние, 1980.— 177 с.

ОБОБЩЕННЫЕ РЕШЕНИЯ И РЕГУЛЯРИЗАЦИЯ СИСТЕМ НЕРАВЕНСТВ

В. А. БУЛАВСКИЙ

Рассматривается подход к решению несовместных систем неравенств, включающий построение их обобщенных решений и регуляризацию. Этот подход представляет естественное обобщение развитого А. Н. Тихоновым [1, 2] способа регуляризации систем линейных алгебраических уравнений и обладает следующими свойствами. Во-первых, он не выводит за рамки линейных задач; во-вторых, позволяет построить класс релаксационных алгоритмов, сходящихся независимо от совместности решаемой задачи, который включает в себя, например, методы последовательного проектирования [3—6] и градиентный метод. Наконец, предлагаемый подход применим для коррекции несовместных (включая двойственную несовместность) задач линейного программирования, которым в последнее время уделяется внимание [7]. Материал статьи является развитием ранее опубликованных результатов автора [5, 8, 9] и основан на использовании аппарата задач с условиями дополнительности. Укажем также, что вводимые понятия и значительная часть результатов могут быть перенесены на нелинейный случай.

В заключение введения оговорим некоторые обозначения и терминологию. Все изложение ведется для конечномерных пространств, и векторы трактуются как векторы-столбцы, так что при умножении на матрицу они стоят справа. Транспонирование обозначается штрихом в качестве верхнего индекса, например, если x — вектор-столбец, то x' — вектор-строка. Для положительной и отрицательной частей числа α приняты обозначения α_+ и α_-. Эти же обозначения используются для векторов: символ x_+ обозначает вектор, составленный из положительных частей компонент вектора x. Покомпонентно применяется и знак неравенства между векторами. Наконец, термины «положительно определенная матрица Q» или «положительно полуопределенная матрица Q» не предполагают симметрию матрицы Q, а лишь означают, что выполняются неравенства $z'Qz > 0$ при $z \neq 0$ или $z'Qz \geqslant 0$ при всех z.

§ 1. ЗАДАЧИ С УСЛОВИЯМИ ДОПОЛНИТЕЛЬНОСТИ

Линейной задачей с условиями дополнительности называют задачу

$$Mz + q \geqslant 0, \quad z \geqslant 0, \quad z'(Mz + q) = 0, \tag{1.1}$$

где $z \in R^s$ — искомый вектор, а матрица $M \in R^{s \times s}$ и свободный член $q \in R^s$ заданы. В последние годы задачам такого типа посвящено много исследований. Обзор результатов и библиографию можно найти, например, в [10]. Сформулируем лишь две теоремы существования для задачи (1.1) и получим из них некоторые следствия, используемые в дальнейшем.

Теорема 1.1. *Если у матрицы M все главные миноры строго положительные, то задача (1.1) имеет ровно одно решение при всяком q.*

Теорема 1.2. *Если матрица M положительно полуопределена, то задача (1.1) имеет решение при совместной системе*

$$Mz+q\geqslant 0,\quad z\geqslant 0. \tag{1.2}$$

Одно из доказательств теоремы 1.2 приведено в [5].

Следующая теорема устанавливает признак разрешимости задачи (1.1), вполне аналогичный известному признаку для систем линейных уравнений.

Теорема 1.3. *Если матрица M положительно полуопределена, то для разрешимости задачи (1.1) необходимо и достаточно, чтобы при всяком ζ, удовлетворяющем однородным условиям*

$$M'\zeta\leqslant 0,\quad \zeta\geqslant 0, \tag{1.3}$$

выполнялось неравенство

$$q'\zeta\geqslant 0. \tag{1.4}$$

При этом два решения z_1 и z_2, если они существуют, удовлетворяют равенству $(z_1-z_2)'M(z_1-z_2)=0$.

Доказательство. Первая часть утверждения теоремы фактически является известным признаком совместности системы неравенств (1.2). Согласно теореме Минковского — Фаркаша, неравенство (1.4) — следствие системы (1.3) в том и только том случае, когда существует решение системы (1.2). Для завершения доказательства этой части теоремы остается сослаться на теорему 1.2. Если z_1 и z_2 — два решения задачи (1.1), то

$$z_1'(Mz_1+q)=z_2'(Mz_2+q)=0,$$

$$z_1'(Mz_2+q)\geqslant 0,\quad z_2'(Mz_1+q)\geqslant 0.$$

Комбинируя приведенные соотношения, получим, что $(z_1-z_2)'M(z_1-z_2)\leqslant\leqslant 0$, а так как знак $<$ здесь невозможен, то теорема доказана.

Следствием полученного признака является тот факт, что задача (1.1) всегда разрешима, если матрица M положительно определена. Впрочем, это следует и из теоремы 1.1, так как положительно определенная матрица имеет строго положительные главные миноры. Отметим также, что ввиду полуопределенности матрицы M неравенства (1.3) эквивалентны тому, что ζ решает следующую однородную задачу с условиями дополнительности:

$$M\zeta\geqslant 0,\quad \zeta\geqslant 0,\quad \zeta'M\zeta=0. \tag{1.5}$$

Действительно, из (1.3) вытекает, что $\zeta'M\zeta\leqslant 0$ и, следовательно, $\zeta'M\zeta==0$, т. е. $(M+M')\zeta=0$, тогда $M\zeta=-M'\zeta\geqslant 0$. Наоборот, из (1.5) точно так же следует, что $M'\zeta\leqslant 0$.

В дальнейшем нам потребуется уточнять вид матрицы M. В частности, возникнут задачи, в которых матрица M имеет клеточное строение:

$$M=\begin{bmatrix} P & -A' \\ A & Q\end{bmatrix}, \tag{1.6}$$

где P и Q — квадратные положительно полуопределенные матрицы порядка n и m соответственно, а $A\in R^{m\times n}$. В сделанных предположениях матрица (1.6) является положительно полуопределенной, и к ней применимы теоремы 1.2 и 1.3. Однако для этого случая признак разрешимости задачи (1.1) удобно сформулировать в иной форме. В соответствии с клеточным разбиением (1.6) векторы z и q также представимы в виде пар (x, y) и (c, b), где $x, c \in R^n$, а $y, b \in R^m$. Задача (1.1) тогда перепишется в виде

$$Px - A'y + c \geqslant 0,\quad Ax + Qy + b \geqslant 0,\quad x,\ y \geqslant 0, \tag{1.7}$$

$$x'(Px - A'y + c) = y'(Ax + Qy + b) = 0. \tag{1.8}$$

При нулевых матрицах P и Q задача (1.7) — (1.8) эквивалентна паре двойственных задач линейного программирования. Если одна из матриц, скажем Q, равна нулю, а вторая симметричная, то x-компонента решает задачу квадратичного программирования:

$$\min\left\{\frac{1}{2}x'Px + c'x : Ax + b \geqslant 0,\ x \geqslant 0\right\}.$$

Наконец, если обе матрицы симметричные, то задача (1.7) — (1.8) эквивалентна задаче о седловой точке выпукло-вогнутой функции:

$$V(x, y) = \frac{1}{2}(x'Px - y'Qy) - y'Ax + c'x + b'y$$

на произведении положительных ортантов $R^n_+ \times R^m_+$.

Для (1.7) — (1.8) сузим класс допустимых матриц P и Q, определив его следующим образом.

Определение. *Будем говорить, что матрица Q принадлежит классу Σ, если она положительно полуопределена и равенство $\xi'Q\xi = 0$ выполняется только в случае, когда $Q\xi = Q'\xi = 0$.*

Напомним, что симметрия матриц класса Σ не предполагается. Этому классу принадлежат, очевидно, любые положительно определенные и симметричные положительно полуопределенные матрицы. Общий же вид матриц класса Σ можно описать формулой

$$Q = T'\begin{bmatrix} 0 & 0 \\ 0 & R \end{bmatrix} T, \tag{1.9}$$

где T — ортогональная матрица, а R — положительно определенная. Действительно, то что любая матрица вида (1.9) обладает нужными свойствами, проверяется непосредственно. Пусть теперь $Q \in \Sigma$. Положим $S = \frac{1}{2}(Q + Q')$ и $K = \frac{1}{2}(Q - Q')$. Существует ортогональная матрица T, при которой

$$S = T'\begin{bmatrix} 0 & 0 \\ 0 & D \end{bmatrix} T, \tag{1.10}$$

где D — неособенная диагональная матрица. С разбиением (1.10) согласуем разбиение

$$TKT' = \begin{bmatrix} K_1 & K_2 \\ -K_2' & K_3 \end{bmatrix}. \tag{1.11}$$

Здесь K_1 и K_3 — кососимметрические матрицы. В соответствии с данным разбиением выберем произвольный вектор u с условием, что $u' = (v', 0)$, и положим $\xi = T'u$. Поскольку $Q = S + K$, то $\xi'Q\xi = v'K_1v = 0$ ввиду кососимметричности матрицы K_1. Поэтому

$$0 = \xi'Q = u'TQ = u'TS + u'TK = [v'K_1,\ v'K_2]T.$$

Так как матрица T неособенная, то $v'K_1 = 0$ и $v'K_2 = 0$ при любом векторе v, следовательно, K_1 и K_2 — нулевые матрицы. Остается положить $R = D + K_3$.

Отметим, что класс Σ замкнут относительно сложения матриц.

Теорема 1.4. *Если в задаче (1.7) — (1.8) матрицы P и Q принадлежат классу Σ, то для существования решения необходимо и достаточно, чтобы нашлись векторы $x_0 \geqslant 0$ и $y_0 \geqslant 0$, для которых неравенства*

$$\xi'(c - A'y) \geqslant 0,\quad \eta'(Ax_0 + b) \geqslant 0 \tag{1.12}$$

выполнялись бы при любых ξ и η, удовлетворяющих условиям $P\xi=0$, $Q\eta=0$, $\xi\geqslant 0$, $\eta\geqslant 0$. При этом два решения (x_1, y_1) и (x_2, y_2), если они существуют, удовлетворяют равенствам

$$P(x_1-x_2)=0,\ Q(y_1-y_2)=0.$$

Доказательство. Если задача (1.7) — (1.8) разрешима, то в качестве (x_0, y_0) можно взять ее решение, так что необходимость признака доказательства не требует. Пусть теперь выполнены условия признака. Рассмотрим пару (ξ, η), удовлетворяющую требованиям теоремы 1.3 для матрицы (1.6):

$$P'\xi+A'\eta\leqslant 0,\quad -A\xi+Q'\eta\leqslant 0,\quad \xi\geqslant 0,\ \eta\geqslant 0. \tag{1.13}$$

Умножив первые два неравенства на ξ' и η' соответственно и сложив их, получим $\xi'P\xi+\eta'Q\eta\leqslant 0$. Ввиду полуопределенности матриц P и Q отсюда следует, что $\xi'P\xi=\eta'Q\eta=0$, следовательно, $P'\xi=P\xi=0$ и $Q'\eta=Q\eta=0$. Из (1.13) тогда находим, что $A'\eta\leqslant 0$ и $A\xi\geqslant 0$ и выполнены неравенства (1.12). В силу неотрицательности векторов x_0 и y_0 получаем $\xi'c\geqslant 0$ и $\eta'b\geqslant 0$, т. е. выполнены условия теоремы 1.3 и задача (1.7) — (1.8) имеет решение. Если имеются два решения $z_1=(x_1, y_1)$ и $z_2=(x_2, y_2)$, то $(M+M')(z_1-z_2)=0$, т. е. $(P+P')(x_1-x_2)=0$ и $(Q+Q')(y_1-y_2)=0$. Умножив эти равенства на $(x_1-x_2)'$ и $(y_1-y_2)'$, найдем, что $(x_1-x_2)'P(x_1-x_2)=(y_1-y_2)'Q(y_1-y_2)=0$. Так как матрицы P и Q из класса Σ, то это завершает доказательство теоремы.

До сих пор рассматривались задачи, содержащие только неравенства. Однако встречаются и задачи, у которых часть неравенств первой группы в (1.1) заменена равенствами, а соответствующие компоненты неизвестного вектора свободны по знаку, в частности, задачи вида

$$Px+A'y+c=0,\ Ax+Qy+b\geqslant 0, \tag{1.14}$$

$$y\geqslant 0,\ y'(Ax+Qy+b)=0. \tag{1.15}$$

Представив x как разность его положительной и отрицательной частей и заменив уравнения парами противоположных неравенств, задачу (1.14) — (1.15) можно привести к виду (1.7) — (1.8) и теорему 1.4 трансформировать следующим образом.

Теорема 1.5. *Если матрицы P и Q принадлежат классу Σ, то для существования решения задачи (1.14) — (1.15) необходимо и достаточно, чтобы нашлись такие векторы $y_0\geqslant 0$ и x_0, для которых условия*

$$\xi'(c-A'y_0)=0,\ \eta'(Ax_0+b)\geqslant 0$$

выполнялись бы при всех таких ξ и η, что $P\xi=0$, $Q\eta=0$, $\eta\geqslant 0$. Любые два решения (x_1, y_1) и (x_2, y_2) удовлетворяют равенствам $P(x_1-x_2)=0$ и $Q(y_1-y_2)=0$.

§ 2. ОБОБЩЕННЫЕ РЕШЕНИЯ

Задачи, которые будем рассматривать в этом параграфе, формулируются следующим образом. Заданы множество $X\subset R^n$, матрица $H\in R^{s\times n}$ и правая часть $h\in R^s$. Требуется найти вектор $x\in R^n$ из условий

$$Hx+h\geqslant 0,\ x\in X. \tag{2.1}$$

В целом система не предполагается совместной, но множество X будем считать непустым. Более того, хотя определения и результаты этого и следующего параграфа можно распространить на достаточно широкий класс выпуклых множеств X, мы останемся в рамках линейных задач, предположив, что X — выпуклое многогранное множество, т. е. описывается некоторой конечной системой линейных неравенств.

Стандартным способом определения обобщенного решения является замена задачи (2.1) экстремальной задачей

$$\min\{\|(Hx+h)_-\|:x\in X\},$$

где выбор нормы может диктоваться различными соображениями, однако чаще всего используют евклидову норму. Подойдем к определению обобщенного решения несколько с иной стороны и, оставаясь в рамках линейных задач, получим более общее понятие. Сопоставим неравенствам системы (2.1) неизвестные заранее неотрицательные весовые множители $u\in R^s_+$, интуитивный смысл которых — характеризация степени жесткости каждого из неравенств или, иначе, его «вклад» в несовместность всей системы. Если система совместна, то столбец u автоматически окажется нулевым. В дальнейшем для краткости эти весовые множители будем называть оценками.

Выберем матрицу $R\in R^{s\times s}$ так, чтобы она принадлежала классу Σ и условия $w\geqslant 0$, $w'Rw=0$ выполнялись бы только для $w=0$. Согласно теореме 1.3, при каждом $x\in X$ можно определить вектор u из условий

$$u\geqslant 0,\quad Hx+Ru+h\geqslant 0,\quad u'(Hx+Ru+h)=0. \tag{2.2}$$

Определение. *Пара* $(x,u)\in X\times R^s_+$ *называется компромиссной парой для задачи* (2.1), *если выполнены соотношения* (2.2) *и при данном* u *вектор* x *является решением задачи линейного программирования*

$$\max\{(u'H)z:z\in X\}. \tag{2.3}$$

Вектор x *будем называть обобщенным решением задачи* (2.1), *а вектор* u *— оценками.*

Теорема 2.1. *При сделанных предположениях относительно матрицы* R *система* (2.1) *с непустым выпуклым многогранным множеством* X *имеет компромиссную пару. Если при этом матрица* R *положительно определена, то оценки* u *определяются однозначно, а совокупность обобщенных решений образует выпуклое многогранное множество.*

Доказательство. Пусть множество X задается системой неравенств

$$Gx+g\geqslant 0. \tag{2.4}$$

Согласно признаку оптимальности для задачи линейного программирования, вектор $x\in X$ решает задачу (2.3) в том и только том случае, когда существует вектор $v\geqslant 0$, при котором $H'u+G'v=0$ и $v'(Gx+g)=0$. Таким образом, нахождение компромиссной пары (x,u) эквивалентно решению задачи (1.14)—(1.15), если принять $P=0$, $c=0$ и положить

$$y=\begin{bmatrix}v\\u\end{bmatrix},\quad Q=\begin{bmatrix}0&0\\0&R\end{bmatrix},\quad A=\begin{bmatrix}G\\H\end{bmatrix},\quad b=\begin{bmatrix}g\\h\end{bmatrix}. \tag{2.5}$$

Для выполнения условий теоремы 1.5 в нашем случае можно положить $y_0=0$, а в качестве x_0 взять любую точку множества X, или любое решение системы (2.4). При положительно определенной матрице R из теоремы 1.5 следует единственность оценок, так как равенство $Q(y_2-y_1)=0$ означает, что $R(u_1-u_2)=0$, и $u_1-u_2=0$ ввиду неособенности R. Что касается x-компоненты компромиссной пары, то для ее определения при фиксированных оценках u получаем совокупность условий (2.2)—(2.4), каждое из которых определяет выпуклое многогранное множество. Таким же является их (непустое) пересечение. Теорема доказана.

Отметим несколько частных случаев.

а. Если $X=R^n$, т. е. система (2.4) фактически отсутствует, то вектор x может решать задачу линейного программирования (2.3) лишь при $H'u=0$. Для определения компромиссной пары в этом случае запишем задачу

$$Hx+Ru+h\geqslant 0,\ u\geqslant 0,\ u'(Ru+h)=0,\ H'u=0.$$

б. Если $X = R^n_+$, т. е. система (2.4) сводится к условию $x \geqslant 0$, то вектор x решает задачу (2.3) при $H'u \leqslant 0$ и $u'Hx = 0$. Для определения компромиссной пары получаем задачу с условиями дополнительности в основной постановке:

$$Hx + Ru + h \geqslant 0,\ H'u \geqslant 0,\ u \geqslant 0,\ x \geqslant 0,$$
$$u'(Hx + Ru + h) = 0,\ x'H'u = 0.$$

в. Пару двойственных задач линейного программирования

$$\min\{c'y : Ay + b \geqslant 0,\ y \geqslant 0\},$$
$$\min\{b'z : -A'z + c \geqslant 0,\ z \geqslant 0\}$$

можно заменить эквивалентной системой

$$Ay + b \geqslant 0,\ -A'z + c \geqslant 0, \tag{2.6}$$
$$y \geqslant 0,\ z \geqslant 0,\ c'y + b'z = 0. \tag{2.7}$$

Рассмотренная пара задач может и не иметь решения из-за несовместности одной или обеих систем ограничений. Для корректировки задачи можно взять в качестве X множество решений системы (2.7), а для неравенств (2.6) выбрать квадратные матрицы R_1 и R_2 подходящих размеров. В качестве обобщенного решения пары задач линейного программирования получим пару (y, z) из условий

$$Ay + R_1u_1 + b \geqslant 0,\ -A'z + R_2u_2 + c \geqslant 0,\ u_1 \geqslant 0,\ u_2 \geqslant 0,$$
$$u_1'(Ay + R_1u_1 + b) = u_2'(-A'z + R_2u_2 + c) = 0.$$

Форма $(u_1'A)y - (u_2'A')z$ должна достигать максимума на множестве решений системы (2.7). В случае, если пара задач разрешима, то $u_1 = 0$ и $u_2 = 0$ и получается ее обычное решение (y, z). Отметим, что часть ограничений (2.6) также можно перевести в описание множества X, если есть уверенность, что оно получится непустым.

г. Рассмотрим случай, когда матрица R единичная. Условия (2.2) теперь означают, что $u = (Hx + h)_-$, т. е. оценки совпадают с невязкой в системе неравенств. Задача (2.3) при этом является линеаризацией в точке x задачи

$$\min\{\|(Hx + h)_-\|^2 : x \in X\}, \tag{2.8}$$

где норма евклидова. Требование, чтобы вектор x являлся решением задачи (2.3), эквивалентно тому, что x — решение задачи (2.8). Таким образом, при $R = I$ обобщенное решение совпадает с классическим определением, а оценки совпадают с невязкой.

д. Пусть теперь $R = ee'$, где $e' = (1, 1, \ldots, 1)$, т. е. R — квадратная матрица, все элементы которой равны единице. Положив $\varepsilon = e'u$, из (2.2) получим

$$\varepsilon e + Hx + h \geqslant 0. \tag{2.9}$$

Если система (2.1) совместна, то $Ru = 0$ и $\varepsilon = 0$. При несовместной системе (2.1) $u \neq 0$ и $\varepsilon > 0$, так что из условий дополнительности в (2.2) по некоторым компонентам в (2.9) достигается равенство, и поэтому

$$\varepsilon = \max_i\{(Hx + h)^i_-\} = T(x),$$

где верхний индекс означает номер компоненты. Возьмем любой другой $z \in X$ и положим $\bar\varepsilon = T(z)$. Тогда $Hz + h + \bar\varepsilon e \geqslant 0$, и поскольку вектор x решает задачу (2.3), получим

$$\varepsilon^2 = u'Ru = -u'(Hx + h) \leqslant -u'(Hz + h) \leqslant (u'e) \cdot \bar\varepsilon = \varepsilon \cdot \bar\varepsilon.$$

Так как $\varepsilon > 0$, то $\varepsilon \leqslant \bar\varepsilon$. Таким образом, при $R = ee'$ обобщенное решение минимизирует чебышевскую норму невязки в системе (2.1).

е. Предположим, что $X=R^n$, а система неравенств в (2.1) разбита на r подсистем

$$H_i x + h_i \geqslant 0, \quad i=1,\ 2,\ \ldots,\ r, \tag{2.10}$$

каждая из которых предполагается совместной. Не исключается случай, когда каждая подсистема состоит из одного неравенства. Метод последовательного проектирования заключается в том, что итерируемый вектор по очереди проектируется на множества решений подсистем (2.10). Каждое такое проектирование может быть описано следующим образом. Если вектор z надо спроектировать на множество решений i-й подсистемы, то проекция ищется в виде $z+H'_i u_i$, где вектор u_i должен удовлетворять условиям:

$$H_i(z+H'_i u_i)+h_i \geqslant 0, \quad u_i \geqslant 0,$$
$$u'_i\left[H_i(z+H'_i u_i)+h_i\right]=0.$$

Если к началу цикла проектирований имелся вектор x, то к моменту проектирования на i-е множество получим вектор

$$z = x + \sum_{k=1}^{i-1} H'_k u_k,$$

который позволяет для определения очередного u_i сформулировать следующую задачу с условиями дополнительности:

$$H_i x + \sum_{k=1}^{i} (H_i H'_k) u_k + h_i \geqslant 0, \quad u_i \geqslant 0,$$
$$u'_i \left[H_i x + \sum_{k=1}^{i} (H_i H'_k) u_k + h_i \right] = 0.$$

Если векторы u_i собрать в один вектор u и составить нижнюю блочно-треугольную матрицу

$$R = \begin{bmatrix} (H_1H'_1) & & & \\ (H_2H'_1) & (H_2H'_2) & & \\ \cdots & \cdots & \cdots & \cdots \\ (H_rH'_1) & (H_rH'_2) & \cdots & (H_rH'_r) \end{bmatrix},$$

то весь цикл проектирований можно записать в виде задачи (2.2). При этом вместо вектора x к началу очередного цикла проектирований получим вектор $x+H'u$. Условие $H'u=0$, которое возникает вместо задачи (2.3) при $X=R^n$, означает, что вектор x является неподвижной точкой циклического процесса проектирования. В случае $u=0$ вектор x не меняется на протяжении всего цикла и, следовательно, решает систему (2.10). Если же $u \neq 0$, точнее $Ru \neq 0$, то система (2.10) несовместна и неподвижной точкой циклического процесса проектирования служит обобщенное решение системы (2.10), определяемое построенной матрицей R. Сам процесс циклического проектирования сходится к этому обобщенному решению и укладывается в рамки общей схемы метода релаксации, рассматриваемого в § 4.

§ 3. РЕГУЛЯРИЗАЦИЯ ЗАДАЧИ

Если система (2.1) совместна, но имеет не единственное решение, то в качестве регуляризованного решения выбирают такое, которое по некоторой норме является ближайшим к заданному вектору a. Часто в качестве нормы берут евклидову, а в качестве a нулевой вектор, т. е. решают задачу о минимизации величины $\|x\|^2$ на множестве решений системы (2.1). Рассмотрим более общий случай, когда регуляризация задается некоторой положительно определенной матрицей $N \in R^{n\times n}$ и вектором a, причем требуется найти такое решение x, при котором неравенство

$$(Nx-a)'(z-x)\geqslant 0 \tag{3.1}$$

выполняется для любого решения z системы (2.1). Случай минимизации квадратичной нормы получается, если матрица N единичная.

Предположим теперь, что совместность системы (2.1) не предполагается и вместо обычного решения рассматривается обобщенное. Будем в дальнейшем считать, что матрица R положительно определена. В этом случае оценки u находятся однозначно, и если выбрать некоторую компромиссную пару (x_0, u), то совокупность обобщенных решений z описывается условиями

$$Hz+Ru+h\geqslant 0,\quad Gz+g\geqslant 0, \tag{3.2}$$

$$u'(Hz+Ru+h)=0,\quad u'H(z-x_0)=0. \tag{3.3}$$

Множество всех обобщенных решений обозначим через X^*. Как отмечалось, это многогранное множество не пусто.

Определение. *Регуляризованным обобщенным решением задачи* (2.1) *при выбранных матрицах R и N, а также векторе a называется такой вектор $x\in X^*$, при котором неравенство* (3.1) *выполняется для всех $z\in X^*$.*

Теорема 3.1. *Если матрицы R и N положительно определенные, то регуляризованное обобщенное решение существует и единственно.*

Доказательство. Для краткости записи заменим систему (3.2)—(3.3), описывающую множество X^*, некоторой эквивалентной ей системой $Fz+f\geqslant 0$. Для регуляризованного решения x следствием такой системы должно быть неравенство (3.1). По теореме Минковского — Фаркаша это эквивалентно существованию вектора $w\geqslant 0$, при котором

$$Nx-a=F'w,\quad w'f+(Nx-a)'x\leqslant 0.$$

Но поскольку $x\in X^*$, т. е. $Fx+f\geqslant 0$, то

$$0\leqslant w'(Fx+f)=w'f+(Nx-a)'x\leqslant 0$$

и $w'(Fx+f)=0$. Таким образом, задача нахождения регуляризованного обобщенного решения эквивалентна задаче

$$Nx-F'w-a=0,\ Fx+f\geqslant 0,$$
$$w'\geqslant 0,\ w'(Fx+f)=0.$$

Ввиду положительной определенности матрицы N и совместности системы $Fx+f\geqslant 0$, данная задача по теореме 1.5 имеет решение (x, w), причем x определено однозначно. Теорема доказана.

Решать задачу о нахождении регуляризованного обобщенного решения в два этапа не всегда удобно. Естественно свести оба этапа в один, использовав стандартный способ штрафующего множителя. Выберем параметр $\alpha>0$ и рассмотрим задачу о нахождении пары $(x, u)\in X\times R^s_+$, удовлетворяющей соотношениям (2.2), и такой, что вектор x решает задачу линейного программирования

$$\max\{[H'u-\alpha(Nx-a)]'z : z\in X\}. \tag{3.4}$$

Решение задачи обозначим через (x_α, u_α). При $\alpha=0$ задача (3.4) совпадает с (2.3), так что в качестве ее решения (x_0, u_0) получается любая компромиссная пара. В дальнейшем под x_0 будем понимать регуляризованное обобщенное решение, поэтому пара (x_0, u_0) уже будет определена однозначно.

Теорема 3.2. *При любом $\alpha>0$ пара (x_α, u_α) существует, единственна, и при $\alpha\to 0$*

$$\overline{\lim}\,\frac{\|u_\alpha-u_0\|}{\alpha}<+\infty,\quad \overline{\lim}\,\frac{\|x_\alpha-x_0\|}{\alpha}<+\infty.$$

Доказательство. Пусть, как и раньше, множество X описывается системой (2.4). Введя в рассмотрение для задачи (3.4) множители Лагранжа v_α, найдем, что тройка $(x_\alpha, u_\alpha, v_\alpha)$ должна удовлетворять соотношениям

$$\alpha N x_\alpha - G' v_\alpha - H' u_\alpha - \alpha a = 0, \tag{3.5}$$

$$G x_\alpha + g \geqslant 0, \quad v_\alpha \geqslant 0, \tag{3.6}$$

$$H x_\alpha + R u_\alpha + h \geqslant 0, \quad u_\alpha \geqslant 0, \tag{3.7}$$

$$v'_\alpha (G x_\alpha + g) = 0, \quad u'_\alpha (H x_\alpha + R u_\alpha + h) = 0. \tag{3.8}$$

Этим же соотношениям при некотором v_0 удовлетворяет и тройка (x_0, u_0, v_0), если положить $\alpha = 0$. Если взять $v_\alpha = 0$, $u_\alpha = 0$ и $x_\alpha \in X$, то неравенства (3.6) будут выполнены, и поскольку матрицы αN и R положительно определенные, то можно применить теорему 1.5, приняв снова обозначения (2.5) и положив $P = \alpha N$ и $c = -\alpha a$. Таким образом, при $\alpha > 0$ существует решение $(x_\alpha, u_\alpha, v_\alpha)$ задачи (3.5) — (3.8), причем пара (x_α, u_α) определена однозначно.

Получим теперь некоторые оценки. Из (3.7) находим, что

$$u'_0 (H x_\alpha + R u_\alpha + h) \geqslant 0, \quad u'_\alpha (H x_0 + R u_0 + h) \geqslant 0.$$

Так как при этом

$$u'_0 (H x_0 + R u_0 + h) = u'_\alpha (H x_\alpha + R u_\alpha + h) = 0,$$

то

$$(u_\alpha - u_0)' R (u_\alpha - u_0) + (u_\alpha - u_0)' H (x_\alpha - x_0) \leqslant 0. \tag{3.9}$$

Поскольку матрица R положительно определенная, то существует число $\gamma > 0$, при котором $z' R z \geqslant \gamma \|z\|^2$ для любого z, и (3.9) можно переписать в виде

$$\gamma \|u_\alpha - u_0\|^2 + (u_\alpha - u_0)' H (x_\alpha - x_0) \leqslant 0. \tag{3.10}$$

Подобно неравенству (3.9) из (3.6) и (3.8) выводим $(v_\alpha - v_0)' G (x_\alpha - x_0) \leqslant 0$. С учетом этого неравенства из (3.5) и аналогичного равенства при $\alpha = 0$ находим

$$\alpha \delta \|x_\alpha - x_0\|^2 - (u_\alpha - u_0)' H (x_\alpha - x_0) \leqslant \alpha (x_\alpha - x_0)' (a - N x_0),$$

где константа $\delta > 0$ вычислена по положительно определенной матрице N. Сложив полученное неравенство с (3.10), запишем

$$\gamma \|u_\alpha - u_0\|^2 + \alpha \delta \|x_\alpha - x_0\|^2 \leqslant \alpha (x_\alpha - x_0)' (a - N x_0). \tag{3.11}$$

Из (3.11) находим предварительные оценки

$$\|x_\alpha - x_0\|^2 \leqslant \|a - N x_0\| / \delta,$$

$$\|u_\alpha - u_0\|^2 \leqslant \alpha \cdot \|a - N x_0\|^2 / \gamma \delta.$$

Таким образом, совокупность $\{x_\alpha\}$ ограниченная и $\|u_\alpha - u_0\| \to 0$ при $\alpha \to 0$.

Сравним теперь задачи (3.4) и (2.3). Так как

$$\lim_{\alpha \to 0} (H' u_\alpha + \alpha a - \alpha N x_\alpha) = H' u_0,$$

а многогранное множество X допустимых векторов у этих задач одинаково и не зависит от α, то при достаточно малых α всякое решение задачи (3.4) будет являться и решением задачи (2.3). Поэтому при достаточно малых α оказывается $u'_0 H (x_\alpha - x_0) = 0$. С другой стороны, регуляризованное обобщенное решение x_0 по определению минимизирует форму $(N x_0 - a)' z$ на множестве решений системы (3.2) — (3.3). Поэтому существуют множители Лагранжа $p \geqslant 0$, $q \geqslant 0$, λ, μ, при которых

$$N x_0 - a = H' p + G' q + (\lambda + \mu) H' u_0,$$

$$p' (H x_0 + R u_0 + h) = 0, \quad q' (G x_0 + g) = 0.$$

Из этих равенств получаем, что при достаточно малых α

$$(Nx_0 - a)'(x_\alpha - x_0) = (p'H + q'G)(x_\alpha - x_0) + (\lambda + \mu)u_0'H(x_\alpha - x_0) =$$
$$= (p'H + q'G)(x_\alpha - x_0),$$
$$p'H(x_\alpha - x_0) + p'R(u_\alpha - u_0) \geqslant 0, \quad q'G(x_\alpha - x_0) \geqslant 0.$$

Отсюда

$$(a - Nx_0)'(x_\alpha - x_0) \leqslant p'R(u_\alpha - u_0),$$

и вместо (3.11) напишем

$$\gamma\|u_\alpha - u_0\|^2 + \alpha\delta\|x_\alpha - x_0\|^2 \leqslant \alpha(p'R)(u_\alpha - u_0).$$

Из последнего неравенства следует, что

$$\|u_\alpha - u_0\| \leqslant \alpha \cdot \|p'R\|/\gamma, \quad \|x_\alpha - x_0\| \leqslant \sqrt{\alpha} \cdot \|p'R\|/\sqrt{\gamma\delta}.$$

Таким образом, $\lim x_\alpha = x_0$.

Нам осталось уточнить оценку для $\|x_\alpha - x_0\|$. При каждом $\alpha > 0$ в системе (3.5) — (3.7) отберем ограничения, обратившиеся в равенство на тройке $(x_\alpha, u_\alpha, v_\alpha)$, и полученную систему для краткости перепишем в виде

$$\alpha N x_\alpha + T y_\alpha - \alpha a = 0, \tag{3.12}$$
$$S x_\alpha + V y_\alpha + b = 0. \tag{3.13}$$

Здесь положено $y_\alpha = (u_\alpha, v_\alpha)$. Так как в каждой паре неравенств (3.6) — (3.7) обязательно выполняется хоть одно равенство, то оно будет включено в систему (3.12) — (3.13) и, следовательно, условия дополнительности (3.8) будут выполняться на любом решении системы (3.12) — (3.13). Кроме того, если (x, y) — решение этой системы, отличное от (x_α, y_α), то при любом ε пара

$$(x_\alpha + \varepsilon(x - x_\alpha),\ y_\alpha + \varepsilon(y - y_\alpha))$$

также будет решением системы (3.12) — (3.13), а при достаточно малом ε будет удовлетворять отброшенным строгим неравенствам в (3.6) — (3.7). Таким образом, данная пара будет удовлетворять всем условиям (3.5) — (3.8). Поскольку такими условиями, как показано выше, x_α определяется однозначно, то оно однозначно определяется и системой (3.12) — (3.13). Значит, для нахождения x_α можно сначала отбросить зависимые уравнения, затем исключить переменные y_α, а к оставшейся системе относительно x_α применить формулы Крамера. Получим, что каждая компонента вектора x_α определяется системой (3.12) — (3.13) в виде отношения многочленов от α. В зависимости от того, какая получилась система, такие представления для x_α различны. Однако их будет конечное число. Так как при $\alpha \to 0$ отношение двух многочленов всегда имеет предел, возможно, бесконечный, то при достаточно малых α могут встречаться лишь представления, имеющие предел при $\alpha \to 0$, равный x_0. Для любого такого представления отклонение от x_0 имеет порядок малости не ниже α, а поскольку их конечное число, то и $\|x_\alpha - x_0\|$ имеет порядок малости не ниже α. Теорема доказана.

§ 4. ПОСЛЕДОВАТЕЛЬНАЯ РЕЛАКСАЦИЯ

Для поиска пары (x_α, u_α) рассмотрим метод релаксации, каждый шаг которого состоит в следующем. Пусть известно приближение x_k (нижний индекс будем использовать и для указания номера итерации: путаницы с обозначением x_α и u_α это вызвать не должно). Определим вектор u_k из условий

$$Ru_k + (Hx_k + h) \geqslant 0, \quad u_k \geqslant 0, \tag{4.1}$$
$$u_k'[Ru_k + (Hx_k + h)] = 0. \tag{4.2}$$

Поскольку матрица R предполагается положительно определенной, то согласно теореме 1.1 вектор u_k находится однозначно. Затем, выбрав шаг $\lambda_k \geqslant 0$, положим

$$\eta_k = H'u_k + \alpha(a - Nx_k),$$
$$x_{k+1} = \pi(x_k + \lambda_k \eta_k),$$

где π — операция проектирования на множество X. Прежде чем обсуждать вопросы выбора шага λ_k и сходимости, докажем следующую лемму.

Лемма. *Пусть π — оператор проектирования на непустое выпуклое замкнутое множество X. Тогда он является оператором нестрогого сжатия, т. е. $\|\pi(y) - \pi(x)\| \leqslant \|y - x\|$ при всех x и y, а если в этом неравенстве достигается равенство, то $\pi(y) - \pi(x) = y - x$.*

Доказательство. Так как $\pi(x) \in X$, то в силу выпуклости множества X выполняются неравенства

$$[y - \pi(y)]'[\pi(y) - \pi(x)] \geqslant 0, \; [x - \pi(x)]'[\pi(x) - \pi(y)] \geqslant 0.$$

Сложив их, найдем, что

$$(y - x)' \cdot [\pi(y) - \pi(x)] \geqslant \|\pi(y) - \pi(x)\|^2.$$

Применив неравенство Коши — Буняковского, получим

$$\|y - x\| \cdot \|\pi(y) - \pi(x)\| \geqslant (y - x)'[\pi(y) - \pi(x)] \geqslant \|\pi(y) - \pi(x)\|^2,$$

откуда следует свойство нестрогого сжатия. Если же $\|y - x\| = \|\pi(y) - \pi(x)\|$, то в левом неравенстве достигается равенство, что возможно лишь при положительной пропорциональности векторов $y - x$ и $\pi(y) - \pi(x)$. Если к тому же равны их нормы, то получаем нужное равенство. Лемма доказана.

Положим $\eta_\alpha = H'u_\alpha + \alpha(a - Nx_\alpha)$. Так как при $\alpha \geqslant 0$ вектор u_α определен однозначно, то однозначно определено и направление η_α при $\alpha \geqslant 0$. Далее, поскольку x_α максимизирует форму $\eta'_\alpha z$ на множестве X, то $\pi(x_\alpha + \lambda_k \eta_\alpha) = x_\alpha$. Поэтому

$$\|x_{k+1} - x_\alpha\|^2 = \|\pi(x_k + \lambda_k \eta_k) - \pi(x_\alpha + \lambda_k \eta_\alpha)\|^2 \leqslant \|x_k - x_\alpha + \lambda_k(\eta_k - \eta_\alpha)\|^2. \quad (4.3)$$

Используя (4.1), (4.2), (3.7) и (3.8), аналогично формуле (3.9), получим

$$(u_k - u_\alpha)'R(u_k - u_\alpha) + (u_k - u_\alpha)'H(x_k - x_\alpha) \leqslant 0.$$

Но $(\eta_k - \eta_\alpha) = H'(u_k - u_\alpha) - \alpha N(x_k - x_\alpha)$. Поэтому

$$(x_k - x_\alpha)'(\eta_k - \eta_\alpha) \leqslant -[(u_k - u_\alpha)'R(u_k - u_\alpha) + \alpha(x_k - x_\alpha)'N(x_k - x_\alpha)].$$

Таким образом, из (4.3) находим

$$\|x_{k+1} - x_\alpha\|^2 \leqslant \|x_k - x_\alpha\|^2 + \lambda_k^2 \|\eta_k - \eta_\alpha\|^2 - 2\lambda_k V_k, \quad (4.4)$$

где для краткости положим

$$V_k = (u_k - u_\alpha)'R(u_k - u_\alpha) + \alpha(x_k - x_\alpha)'N(x_k - x_\alpha).$$

Напомним, что $V_k \geqslant \gamma\|u_k - u_\alpha\|^2 + \alpha\delta\|x_k - x_\alpha\|^2$, где $\gamma > 0$ и $\delta > 0$.

Выберем константу β из условия, что

$$\beta[w'Rw + \alpha z'Nz] \geqslant \|H'w - \alpha Nz\|^2 \quad (4.5)$$

при всех w и z. Например, если известны константы β_R и β_N, для которых при всех w и z

$$\beta_R \cdot (w'Rw) \geqslant \|H'w\|^2, \; \beta_N \cdot (z'Nz) \geqslant \|Nz\|^2,$$

то, поскольку при любом $\Delta > 0$

$$\|H'w - \alpha Nz\|^2 \leqslant (1 + \Delta)\|H'w\|^2 + \left(1 + \frac{1}{\Delta}\right)\alpha^2 \|Nz\|^2 \leqslant$$

$$\leqslant \beta_R(1 + \Delta)(w'Rw) + \alpha^2 \beta_N \left(1 + \frac{1}{\Delta}\right)(z'Nz),$$

можно выбрать $\Delta=\alpha\beta_N/\beta_R$ и положить $\beta=\beta_R+\alpha\beta_N$. В частности, годятся $\beta_R=\|HH'\|/\gamma$ и $\beta_N=\|N'N\|/\delta$. Наконец, если R и N единичные матрицы, то $\gamma=\delta=\|NN'\|=1$ и можно принять $\beta=\|HH'\|+\alpha$.

Теорема 4.1. *Если при некотором $\varepsilon\in(0,\ 1]$ шаги λ_k удовлетворяют неравенствам $\varepsilon\leqslant\beta\lambda_k\leqslant 2-\varepsilon$, то последовательность $\{(x_k,\ u_k)\}$ сходится к паре $(x_\alpha,\ u_\alpha)$ при $\alpha>0$ и к некоторой компромиссной паре, если $\alpha=0$.*

Доказательство. Пара $(x_\alpha,\ u_\alpha)$ определена однозначно, если $\alpha>0$. Если же $\alpha=0$, то под x_0 будем понимать любое обобщенное решение. Оценки u_0 и в этом случае однозначно определены. Положив в (4.5) $w=u_k-u_\alpha$ и $z=x_k-x_\alpha$, получим

$$\beta V_k\geqslant\|\eta_k-\eta_\alpha\|^2. \tag{4.6}$$

Поэтому из (4.4) следует

$$\|x_{k+1}-x_\alpha\|^2\leqslant\|x_k-x_\alpha\|^2-\lambda_k(2-\beta\lambda_k)V_k$$

и в силу выбора λ_k

$$\|x_{k+1}-x_\alpha\|^2\leqslant\|x_k-x_\alpha\|^2-\frac{\varepsilon(2-\varepsilon)}{\beta}V_k. \tag{4.7}$$

Просуммировав эти неравенства, находим

$$\sum_{k=1}^{\infty}V_k\leqslant\frac{\beta}{\varepsilon(2-\varepsilon)}\|x_1-x_\alpha\|^2.$$

Поскольку члены ряда неотрицательные, то из его сходимости вытекает, что $V_k\to 0$, т. е. $\lim\|u_k-u_\alpha\|=0$, а если $\alpha>0$, то и $\lim\|x_k-x_\alpha\|=0$.

Рассмотрим отдельно случай $\alpha=0$. Снова используя (4.6), из (4.4) найдем

$$\|x_{k+1}-x_0\|^2\leqslant\|x_k-x_0\|^2-\frac{\lambda_k(2-\beta\lambda_k)}{\beta}\|\eta_k-\eta_0\|^2,$$

и так же, как для V_k, установим, что $\lim\|\eta_k-\eta_0\|=0$. Далее, из (4.7) следует ограниченность последовательности $\{x_k\}$ и монотонное убывание отклонений $\|x_k-x_0\|$. Таким образом, существует $\lim\|x_k-x_0\|=\mu\geqslant 0$. Так как последовательность $\{\lambda_k\}$ тоже ограниченная, то можно выбрать подпоследовательность шагов $\{k_\nu\}$, для которой существуют пределы $\lim x_{k_\nu}=x$ и $\lim\lambda_{k_\nu}=\lambda>0$. Подставим в (4.3) $\alpha=0$, $k=k_\nu$ и перейдем к пределу. В силу непрерывности нормы и операции проектирования на выпуклое замкнутое множество получим

$$\mu^2=\|\pi(x+\lambda\eta_0)-\pi(x_0+\lambda\eta_0)\|\leqslant\|x-x_0\|^2=\mu^2.$$

Следовательно, фактически в средней части этого соотношения выполняется равенство и согласно лемме

$$\pi(x+\lambda\eta_0)-\pi(x_0+\lambda\eta_0)=x-x_0.$$

Однако x_0 — обобщенное решение, которое максимизирует форму $\eta_0'z=(u_0'H)z$ на множестве X. Поэтому $\pi(x_0+\lambda\eta_0)=x_0$, откуда $\pi(x+\lambda\eta_0)=x$. Так как $\lambda>0$, то вектор x также максимизирует форму $\eta_0'z$ на множестве X. Кроме того, если перейти к пределу в соотношениях (4.1)—(4.2), то получим, что x удовлетворяет условиям (2.2) и, следовательно, является обобщенным решением. Но при выводе формулы (4.7) для случая $\alpha=0$ в качестве x_0 можно взять любое обобщенное решение. В частности, можно считать, что $x_0=x$. Таким образом, монотонно убывающая последовательность $\{\|x_k-x\|\}$ содержит сходящуюся к нулю подпоследовательность. Поэтому $\lim\|x_k-x\|=0$. Теорема доказана.

Рассмотрим вопрос о скорости сходимости. Если $\alpha>0$, то ввиду неравенства $V_k\geqslant\alpha\delta\|x_k-x_\alpha\|^2$ получаем оценку

$$\|x_{k+1}-x_\alpha\|^2\leqslant\left(1-\frac{\alpha\delta\varepsilon(2-\varepsilon)}{\beta}\right)\|x_k-x_\alpha\|^2,$$

т. е. скорость сходимости линейная с коэффициентом

$$q=\sqrt{1-\frac{\alpha\delta\varepsilon(2-\varepsilon)}{\beta}}.$$

При очень малых α коэффициент q близок к единице, и оценка гарантирует лишь медленную сходимость. Однако следует заметить, что регуляризация требуется обычно в плохо обусловленных задачах, и медленная сходимость здесь по существу.

Пусть теперь $\alpha=0$. Для простоты предположим, что $X=R^n$, так что проектирование отсутствует и $\eta_0=H'u_0=0$. Сделаем еще предположение о невырожденности: в каждой паре неравенств (3.7) одно из них выполняется как строгое неравенство. Поскольку $x_k\to x_0$ и $u_k\to u_0$, то начиная с некоторого номера k распределение строгих равенств и неравенств в (4.1) стабилизируется и можно отбросить нулевые компоненты векторов u_k и u_0 и соответствующие им строгие неравенства в (4.1) и (3.7). Чтобы не вводить новых обозначений, будем считать, что это уже сделано, и на рассматриваемых шагах процесса

$$Ru_k+Hx_k+h=0,\quad u_k>0,$$
$$Ru_0+Hx_0+h=0,\quad u_0>0.$$

Положим также $\Delta_k=x_k-x_0$. Тогда

$$\eta_k=H'u_k=-H'R^{-1}Hx_k-H'R^{-1}h,$$
$$0=H'u_0=-H'R^{-1}Hx_0-H'R^{-1}h.$$

Поэтому, если положить $B=H'R^{-1}H$, получим, что $\eta_k=-B\Delta_k$. Формулу итерации перепишем в виде

$$\Delta_{k+1}=\Delta_k-\lambda_kB\Delta_k=(I-\lambda_kB)\Delta_k,$$

следовательно,

$$\|\Delta_{k+1}\|^2=\Delta'_{k+1}\cdot\Delta_{k+1}=\|\Delta_k\|^2-2\lambda_k\Delta'_kB\Delta_k+\lambda_k^2\|B\Delta_k\|^2.$$

Воспользуемся теперь формулой (4.5) для выбора β, положив там $\alpha=0$ и $w=R^{-1}H\Delta_k$. Найдем, что

$$\|B\Delta_k\|^2=\|H'R^{-1}H\Delta_k\|^2\leqslant\beta(w'R'w)=\beta(\Delta'_kB\Delta_k),$$

и значит, аналогично формуле (4.7) получим

$$\|\Delta_{k+1}\|^2\leqslant\|\Delta_k\|^2-\frac{\varepsilon(2-\varepsilon)}{\beta}(\Delta'_kB\Delta_k). \tag{4.8}$$

Рассмотрим подпространство $\mathscr{L}=\{Bz:z\in R^n\}$. Так как $\Delta_k\to0$, а на каждой итерации из Δ_k вычитается вектор $\lambda_kB\Delta_k\in\mathscr{L}$, то для всех рассматриваемых шагов $\Delta_k\in\mathscr{L}$. С другой стороны, матрица B положительно полуопределенная и если $B'w=0$ для некоторого $w=Bz\in\mathscr{L}$, то $z'B'Bz=z'B'w=0$, т. е. $w=Bz=0$. Таким образом, для ненулевых векторов $w\in\mathscr{L}$ оказывается $w'Bw>0$. Поэтому существует такая константа $\varkappa>0$, что $w'Bw\geqslant\varkappa\|w\|^2$ при $w\in\mathscr{L}$. В частности, это неравенство справедливо для $w=\Delta_k$. Из (4.8) окончательно получаем

$$\|\Delta_{k+1}\|/\|\Delta_k\|\leqslant\sqrt{1-\varepsilon(2-\varepsilon)\frac{\varkappa}{\beta}}=q.$$

В заключение отметим следующее. Если R — единичная матрица, то в качестве β можно взять норму матрицы HH' — ее наибольшее собственное число, а в качестве $\varkappa$ — наименьшее собственное число этой же матрицы, отличное от нуля. Следовательно, отношение $\beta/\varkappa$ играет роль числа обусловленности.

ЛИТЕРАТУРА

1. **Тихонов А. Н.** Об устойчивости алгоритмов для решения вырожденных систем линейных алгебраических уравнений.— Журн. вычисл. математики и мат. физики, 1965, т. 5, № 4, с. 181—188.

2. **Тихонов А. Н.** О приближенных системах линейных алгебраических уравнений.— Журн. вычисл. математики и мат. физики, 1980, т. 20, № 6, с. 1373—1383.
3. **Булавский В. А.** Итеративный метод решения задач линейного программирования.— Докл. АН СССР, 1961, т. 137, № 2, с. 258—260.
4. **Еремин И. И.** Обобщение релаксационного метода Моцкина — Агмона.— Успехи мат. наук, 1965, т. 20, вып. 2, с. 183—187.
5. **Булавский В. А.** Методы релаксации для систем неравенств. — Новосибирск: НГУ, 1981.— 83 с.
6. **Motzkin T. S., Schoenberg J. J.** The relaxation method for linear inequalities.— Canad. J. Mathem. 1954, N 6, № 3, p. 394—404.
7. **Еремин И. И., Мазуров В. Д., Астафьев Н. Н.** Несобственные задачи линейного и выпуклого программирования.— М.: Наука, 1983.— 336 с.
8. **Булавский В. А.** Релаксация в задачах с неравенствами.— В кн.: Оптимизация. Новосибирск: ИМ СО АН СССР, 1979, вып. 23 (40), с. 32—40.
9. **Булавский В. А.** Квазилинейное программирование и векторная оптимизация.— Докл. АН СССР, 1981, т. 257, № 4, с. 788—791.
10. **Eaves B. C.** The linear complementarity problem.— Manag. Sci., 1971, v. 17, p. 612—634.

ПРИМЕНЕНИЕ ОБОБЩЕННОЙ ТЕОРЕМЫ ШТУРМА ДЛЯ ВЫЧИСЛЕНИЯ СОБСТВЕННЫХ ЧИСЕЛ ОДНОГО КЛАССА МАТРИЦ

В. А. БУЛАВСКИЙ, М. А. ЯКОВЛЕВА

Статья является расширенным изложением и развитием ранее опубликованной работы авторов [1]. Вместо предложенного там термина «обобщенные трехдиагональные матрицы» здесь вводится название «обобщенные якобиевы матрицы», поскольку слово «трехдиагональные» оказалось чересчур наглядным и настоящих трех ненулевых диагоналей у интересующих нас матриц, строго говоря, нет. Основное содержание статьи можно охарактеризовать следующим образом. Рассматривается класс матриц, содержащий, в частности, трехдиагональные матрицы. Для этого класса формулируется и доказывается подходящим образом обобщенная теорема Штурма о числе корней многочлена. На ее основе описывается метод определения границ собственных чисел обобщенной якобиевой матрицы, полностью аналогичный известному методу для трехдиагональных матриц [2, 3]. Для данного метода приводятся условия и оценки, гарантирующие его безаварийную реализацию на ЭВМ и объявленную точность полученного результата.

§ 1. ОБОБЩЕННЫЕ ЯКОБИЕВЫ МАТРИЦЫ

Класс матриц, который вводится в рассмотрение, естественным образом обобщает понятие трехдиагональной матрицы. Фактически это понятие определяется конфигурацией элементов, значения которых не предполагаются нулевыми. Если дополнительно некоторые из них окажутся нулями, то матрица может распасться на независимые диагональные клетки той же структуры. Поэтому в определении удобно описывать не фактическое расположение ненулевых элементов, а допустимую их конфигурацию.

Рассмотрим отображение

$$i \to k(i), \quad i=1,\ 2,\ \ldots,\ n-1, \qquad (1)$$

где $i+1 \leqslant k(i) \leqslant n$ при всех i.

Определение. *Матрица A порядка n с элементами a_{ij} называется обобщенной якобиевой матрицей со структурой* (1), *если у нее отличны от нуля разве лишь диагональные элементы, а также элементы $a_{i,k(i)}$ и $a_{k(i),i}$, $i=1, 2, \ldots, n-1$.*

В дальнейшем для ненулевых элементов обобщенной якобиевой матрицы примем обозначения:

$$d_i = a_{ii},\ a_i = a_{i,k(i)},\ b_i = a_{k(i),i}.$$