УДК 519.6

Об апостериорной аппроксимации множества решений системы уравнений квадратичной структуры с использованием метода Ньютона*

М.Ю. Кокурин, А.И. Козлов

Марийский государственный университет, пл. им. Ленина, 1, Йошкар-Ола, 424001
E-mails: kokurinm@yandex.ru (Кокурин М.Ю.), matemanaliz@rambler.ru (Козлов А.И.)

Кокурин М.Ю., Козлов А.И. Об апостериорной аппроксимации множества решений системы уравнений квадратичной структуры с использованием метода Ньютона // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние.—Новосибирск, 2014.—Т. 17, № 1.—С. 53–65.

Для квадратичных систем алгебраических уравнений предлагается алгоритм апостериорной аппроксимации выпуклой оболочки множества решений по результатам шага метода Ньютона. Приведены результаты численных экспериментов.

Ключевые слова: *квадратичный оператор, метод Ньютона, апостериорная оценка, числовой образ, выпуклая оболочка.*

Kokurin M.Yu., Kozlov A.I. On a posteriori approximation of a set of solutions to a system of quadratic equations with the use of the Newton method // Siberian J. Num. Math. / Sib. Branch of Russ. Acad. of Sci.—Novosibirsk, 2014.—Vol. 17, № 1.—P. 53–65.

For quadratic systems of algebraic equations we propose an algorithm generating a posteriori estimates of a convex hull of a set of solutions using the results of a step of the Newton method. Results of numerical tests are given.

Key words: *quadratic operator, the Newton method, a posteriori estimate, numerical range, convex hull.*

1. Введение

Объектом исследования в работе является система нелинейных уравнений:

$$F(x) = 0, \quad x \in \mathbb{R}^n, \tag{1}$$

определяемая оператором $F : \mathbb{R}^n \to \mathbb{R}^n$ квадратичной структуры

$$F(x) = \frac{1}{2}\Lambda[x]^2 + Ax + f. \tag{2}$$

Здесь $A \in \mathbb{R}^{n\times n}$, $f \in \mathbb{R}^n$, $\Lambda[x]^2 = \Lambda[x, x]$,

$$\Lambda[x, y] = \big((Q_1x, y), \ldots, (Q_nx, y)\big)^\top, \quad x, y \in \mathbb{R}^n. \tag{3}$$

Считаем, что $Q_j^\top = Q_j \in \mathbb{R}^{n\times n}$, $j = 1, \ldots, n$, и множество решений X^* задачи (1)–(3) непусто.

*Работа выполнена при поддержке РФФИ (проект № 12-01-00239a).

В общем случае вещественное алгебраическое многообразие X^* может иметь достаточно сложную структуру, в том числе содержать наряду с отдельными точками подмногообразия различных положительных размерностей. Традиционная логика применения численных методов решения систем нелинейных уравнений сводится к отысканию в пределе итерационного процесса некоторого одного элемента из множества решений. В том случае, когда наряду с отысканием отдельного решения интерес представляет и выяснение глобальной структуры множества решений, подобная техника оказывается малоэффективной. В настоящей работе предлагается способ дополнения одной из стандартных вычислительных процедур — метода Ньютона — апостериорным алгоритмом локализации множества X^* в целом. Предлагаемая схема позволяет единообразно конструировать различные внешние выпуклые многогранные аппроксимации выпуклой оболочки $\mathrm{conv}(X^*)$ множества решений. Возможности рассматриваемой техники иллюстрируются численными примерами.

2. Вспомогательные построения

Зафиксируем начальное приближение $x_0 \in \mathbb{R}^n$. Предполагая регулярность оператора $F'(x_0)$, т. е. невырожденность матрицы $F'(x_0) \in \mathbb{R}^{n\times n}$, выполним один шаг метода Ньютона и определим приближение

$$N(x_0) = x_0 - F'(x_0)^{-1}F(x_0). \tag{4}$$

Здесь на основании (2), (3):

$$F'(x)h = Ah + \Lambda[x, h], \quad x, h \in \mathbb{R}^n. \tag{5}$$

Для произвольного элемента $x^* \in X^*$ с использованием (2), (4), (5) получаем

$$N(x_0) - x^* = F'(x_0)^{-1}\bigl[F'(x_0)(x_0 - x^*) - F(x_0)\bigr] = \frac{1}{2}F'(x_0)^{-1}\Lambda[x_0 - x^*]^2. \tag{6}$$

Согласно (6):

$$x^* = N(x_0) - \frac{1}{2}F'(x_0)^{-1}\Lambda[x_0 - x^*]^2. \tag{7}$$

Равенство (7) является основным в последующих построениях. Оно позволяет по результату шага метода Ньютона уточнить имеющуюся априорную информацию об искомом решении или обо всем множестве решений. Именно, если априори известно, что $x^* \in G \subset \mathbb{R}^n$, то в силу (7) дополнительно выполняется

$$x^* \in N(x_0) + L(A, \Lambda, x_0)\Lambda[x_0 - G]^2, \tag{8}$$

где

$$L(A, \Lambda, x_0) = -\frac{1}{2}F'(x_0)^{-1},\ x_0 - G = \{x_0 - x : x \in G\},\ \Lambda[x_0 - G]^2 = \bigl\{\Lambda[x]^2 : x \in x_0 - G\bigr\},$$

$L(A, \Lambda, x_0)\Lambda[x_0 - G]^2$ — образ множества $\Lambda[x_0 - G]^2$ под действием оператора $L(A, \Lambda, x_0)$. Таким образом, после шага метода Ньютона исходная априорная информация $x^* \in G$ уточняется следующим образом:

$$x^* \in G \cap \bigl\{N(x_0) + L(A, \Lambda, x_0)\Lambda[x_0 - G]^2\bigr\}.$$

Поскольку точное описание образа $\Lambda[x_0 - G]^2$ множества $x_0 - G$ под действием квадратичного отображения $\Lambda[\cdot]^2$ возможно лишь в редких случаях, на практике естественно пользоваться подходящими внешними аппроксимациями $\widetilde{\Lambda} \supset \Lambda[x_0 - G]^2$.

Предположим, что какая-либо априорная информация о расположении X^* отсутствует, тогда естественно положить $G = \mathbb{R}^n$. Согласно (8), в этом случае имеет место включение

$$X^* \subset N(x_0) + L(A, \Lambda, x_0)\Lambda[\mathbb{R}^n]^2. \tag{9}$$

Правая часть включения (9) определяет внешнюю коническую аппроксимацию множества X^*. В случае $n \geq 3$ численное определение множества $\Lambda[\mathbb{R}^n]^2$, стоящего в правой части (9), как правило, затруднительно. Одна из причин заключается в том, что при указанных значениях n это множество в общем случае невыпукло [1]. Численно реализуемые варианты намеченной схемы внешней аппроксимации множества решений (1) могут быть связаны с построением обозримой внешней выпуклой аппроксимации $\widetilde{\Lambda}$ для конуса $\Lambda[\mathbb{R}^n]^2$. Поскольку в этом случае множество $L(A, \Lambda, x_0)\widetilde{\Lambda}$ выпукло, включение (9) допускает следующее уточнение.

Теорема 1. *Пусть множество $\widetilde{\Lambda}$ выпукло и $\widetilde{\Lambda} \supset \Lambda[\mathbb{R}^n]^2$. Тогда имеет место включение*

$$\operatorname{conv}(X^*) \subset \Xi(x_0) \equiv N(x_0) + L(A, \Lambda, x_0)\widetilde{\Lambda}. \tag{10}$$

Следующее утверждение позволяет в отдельных случаях получить с использованием $\widetilde{\Lambda}$ качественную информацию о структуре множества X^*.

Теорема 2. *Предположим, что $\widetilde{\Lambda} \neq \mathbb{R}^n$ — выпуклый конус, $\Lambda[\mathbb{R}^n]^2 \subset \widetilde{\Lambda}$ и точка $x^* \in X^*$ такова, что $|F'(x^*)| \neq 0$. Тогда $x^* \notin \operatorname{intconv}(X^*)$, т. е. x^* — граничная точка $\operatorname{conv}(X^*)$.*

Доказательство. Предположим, что $x^* \in \operatorname{intconv}(X^*)$. В силу условия $|F'(x^*)| \neq 0$, метод Ньютона локально сходится к точке x^* [2, с. 333]. Поэтому, выбрав x_0 достаточно близко к x^*, получим также $N(x_0) \in \operatorname{intconv}(X^*)$. Поскольку $\widetilde{\Lambda} \neq \mathbb{R}^n$ и $L(A, \Lambda, x_0) : \mathbb{R}^n \to \mathbb{R}^n$ — линейное отображение, выполняется $L(A, \Lambda, x_0)\widetilde{\Lambda} \neq \mathbb{R}^n$. Следовательно, выпуклый конус $N(x_0) + L(A, \Lambda, x_0)\widetilde{\Lambda}$ с вершиной в точке $N(x_0)$ не содержит никакую открытую окрестность x^*, что противоречит включению (10). □

В ходе доказательства установлено следующее утверждение, показывающее, что из любой начальной точки x_0 метод Ньютона не может сделать шаг внутрь $\operatorname{conv}(X^*)$.

Теорема 3. *Пусть $\widetilde{\Lambda}$ удовлетворяет условию теоремы 2 и точка $x_0 \in \mathbb{R}^n$. Тогда $N(x_0) \notin \operatorname{intconv}(X^*)$, т. е. $N(x_0)$ — граничная либо внешняя точка для $\operatorname{conv}(X^*)$.*

3. Построение конических аппроксимаций

Опишем предлагаемый способ построения выпуклой конической аппроксимации $\widetilde{\Lambda}$ для невыпуклого в общем случае конуса $\Lambda[\mathbb{R}^n]^2$.

1) Рассмотрим вначале случай четной размерности $n = 2m$. Обозначим

$$K_j = \left\{ \big((Q_{2j-1}x, x), (Q_{2j}x, x)\big)^\top : \ x \in \mathbb{R}^n \right\}, \quad j = 1, \ldots, m.$$

Очевидно, K_j есть конус с центром в начале координат плоскости $\mathbb{R}^2$, при этом граничные лучи не обязательно принадлежат K_j. Для построения K_j введем матрицу

$D_j = Q_{2j-1} + iQ_{2j} \in \mathbb{C}^{n\times n}$ и рассмотрим ее числовой образ

$$W(D_j) = \big\{(D_j z, z)_{\mathbb{C}^n} : \ \|z\|_{\mathbb{C}^n} = 1\big\}, \quad j = 1, \ldots, m.$$

Известно [3, гл. 11], что $W(D_j)$ — выпуклое замкнутое множество на плоскости $\mathbb{C} \sim \mathbb{R}^2$, содержащее спектр матрицы D_j. Кроме того, K_j есть коническая оболочка $W(D_j)$ [4]:

$$K_j = \operatorname{cone}(W(D_j)), \quad j = 1, \ldots, m.$$

Имеет место включение

$$\Lambda[\mathbb{R}^n]^2 = \big\{\big((Q_1 x, x), \ldots, (Q_n x, x)\big) : \ x \in \mathbb{R}^n\big\} \subset \overline{K}_1 \times \ldots \times \overline{K}_m, \tag{11}$$

где черта сверху означает замыкание. На основании (11) положим

$$\widetilde{\Lambda} = \overline{K}_1 \times \ldots \times \overline{K}_m.$$

Для построения числового образа $W(D_j)$ $(j = 1, \ldots, m)$ будем пользоваться следующим утверждением.

Теорема 4 [5]. *Пусть* $D_j = Q_{2j-1} + iQ_{2j} \in \mathbb{C}^{n\times n}$, $Q_{2j-1}^\top = Q_{2j-1}$, $Q_{2j}^\top = Q_{2j} \in \mathbb{R}^{n\times n}$ *и*

$$Q_{2j-1,t} = Q_{2j-1}\cos t + Q_{2j}\sin t, \quad Q_{2j,t} = Q_{2j}\cos t - Q_{2j-1}\sin t, \quad t \in [0, 2\pi),$$

матрица P_t *осуществляет ортогональное проектирование* $\mathbb{R}^n$ *на собственное подпространство матрицы* $Q_{2j-1,t}$, *соответствующее ее наибольшему собственному значению. Пусть из всех собственных векторов* v_t, $\|v_t\| = 1$, *соответствующих наибольшему собственному значению матрицы* $Q_{2j-1,t}$, *выбраны* v_t^+, v_t^-, *являющиеся также собственными векторами матрицы* $P_t Q_{2j,t} P_t$ *с наибольшим и наименьшим собственными значениями соответственно. Тогда для каждого* $t \in [0, 2\pi]$ *точки* $(D_j v_t^+, v_t^+)$ *и* $(D_j v_t^-, v_t^-)$ *принадлежат границе числового образа* $W(D_j)$, $j = 1, \ldots, m$.

Теорема 4 является основой следующего алгоритма для построения замкнутой конической оболочки $\overline{K}_j$, $j = 1, \ldots, m$.

Алгоритм 1.

1. Определяется граница числового образа матрицы $D_j = Q_{2j-1} + iQ_{2j}$, $j = 1, \ldots, m$.
   1.1. Для значений $t = t_k = 2\pi k/N$, $k = 0, 1, \ldots, N-1$, определяются матрицы $Q_{2j-1,t}$, $Q_{2j,t}$.
   1.2. Для матрицы $Q_{2j-1,t}$ находится собственный вектор v_t, $\|v_t\| = 1$, соответствующий наибольшему собственному значению λ_t.
   1.3. Если значению λ_t соответствует одномерное собственное подпространство E_t, то находится одна граничная точка $W(D_j)$ с координатами:
   $$\big((Q_{2j-1,t} v_t, v_t), (Q_{2j,t} v_t, v_t)\big).$$
   1.4. Если $\dim E_t \geq 2$, то определяются векторы $v_t^+, v_t^- \in E_t$, соответствующие наибольшему и наименьшему собственным значениям матрицы $P_t Q_{2j,t} P_t$, после чего находятся две граничные точки $W(D_j)$ с координатами:
   $$\big((Q_{2j-1,t} v_t^+, v_t^+), (Q_{2j,t} v_t^+, v_t^+)\big), \quad \big((Q_{2j-1,t} v_t^-, v_t^-), (Q_{2j,t} v_t^-, v_t^-)\big).$$

2. Находится замкнутая коническая оболочка $\overline{K}_j = \overline{\mathrm{cone}\big(W(D_j)\big)}$.

   2.1. Если граничные точки $W(D_j)$ лежат во всех четырех координатных четвертях, то $\overline{K}_j = \mathbb{R}^2$.

   2.2. В противном случае среди граничных точек выбираются те, радиус-векторы которых образуют наименьший и наибольший углы $\varphi \in (-\pi, \pi]$ с фиксированной осью. Эти радиус-векторы и определяют коническую оболочку $\overline{K}_j$.

Поскольку $\overline{K}_j$ $(j = 1, \ldots, m)$ есть выпуклый конус, могут представиться следующие случаи:

(i) $\overline{K}_j$ — коническая оболочка векторов единичной длины $b_{2j-1}, b_{2j} \in \mathbb{R}^2$, угол между которыми меньше π:

$$\overline{K}_j = \big\{y \in \mathbb{R}^2 : \ y = \alpha_{2j-1} b_{2j-1} + \alpha_{2j} b_{2j}; \quad \alpha_{2j-1}, \alpha_{2j} \geq 0\big\}.$$

(ii) $\overline{K}_j$ — полуплоскость

$$\overline{K}_j = \big\{y \in \mathbb{R}^2 : \ y = \alpha_{2j-1} b_{2j-1} + \alpha_{2j} b_{2j}; \quad \alpha_{2j-1} \in \mathbb{R}, \ \alpha_{2j} \geq 0\big\},$$

определяемая ортогональными единичными векторами $b_{2j-1}, b_{2j} \in \mathbb{R}^2$.

(iii) $\overline{K}_j$ — плоскость, которая определяется ортами $b_{2j-1}, b_{2j} \in \mathbb{R}^2$:

$$\overline{K}_j = \big\{y \in \mathbb{R}^2 : \ y = \alpha_{2j-1} b_{2j-1} + \alpha_{2j} b_{2j}; \quad \alpha_{2j-1}, \alpha_{2j} \in \mathbb{R}\big\}.$$

Поскольку второй и третий случаи отличаются от первого лишь условиями на коэффициенты α_{2j} и α_{2j-1}, дальнейшие построения опишем подробно лишь для случая (i).

В случае (i) для $\widetilde{\Lambda}$ имеем представление

$$\widetilde{\Lambda} = \Big\{x \in \mathbb{R}^n : \ x = \sum_{k=1}^{n} \alpha_k a_k; \quad \alpha_k \geq 0, \ k = 1, \ldots, n\Big\}. \tag{12}$$

В (12) векторы $a_{2j-1}, a_{2j} \in \mathbb{R}^n$ имеют $(2j-1)$-е и $2j$-е координаты, совпадающие с координатами векторов $b_{2j-1}, b_{2j} \in \mathbb{R}^2$, а остальные координаты у этих векторов нулевые: $a_{2j-1} = (0, \ldots, 0, b_{2j-1}^{\top}, 0, \ldots, 0)^{\top}$, $a_{2j} = (0, \ldots, 0, b_{2j}^{\top}, 0, \ldots, 0)^{\top}$, $j = 1, \ldots, m$. Далее, образом выпуклого конуса $\widetilde{\Lambda}$ при действии оператором $L(A, \Lambda, x_0)$ является выпуклый конус

$$\begin{gathered} L(A, \Lambda, x_0)\widetilde{\Lambda} = \Big\{x \in \mathbb{R}^n : x = \sum_{k=1}^{n} \alpha_k c_k(x_0); \quad \alpha_k \geq 0, \ k = 1, \ldots, n\Big\}, \\ c_k(x_0) = L(A, \Lambda, x_0) a_k, \quad k = 1, \ldots, n. \end{gathered} \tag{13}$$

Из (13) следует, что выпуклый конус $\Xi(x_0)$, доставляющий согласно (10) внешнюю аппроксимацию $\mathrm{conv}(X^*)$, имеет вид

$$\Xi(x_0) = \Big\{x \in \mathbb{R}^n : \ x = N(x_0) + \sum_{k=1}^{n} \alpha_k c_k(x_0); \quad \alpha_k \geq 0, \ k = 1, \ldots, n\Big\}. \tag{14}$$

2) Пусть теперь $n = 2m+1$. Будем считать, что $(2m+1)$-е уравнение в (1) определяется неотрицательной матрицей $Q_{2m+1} \geq O$, тогда вместо (11) справедливо включение

$$\Lambda[\mathbb{R}^n]^2 \subset \overline{K}_1 \times \ldots \times \overline{K}_m \times \mathbb{R}_+, \quad \mathbb{R}_+ = [0, +\infty).$$

Поэтому вместо (12) имеем представление

$$\widetilde{\Lambda} = \Big\{x \in \mathbb{R}^n : \ x = \sum_{k=1}^{n-1} \alpha_k a_k + \alpha_n e \,; \quad \alpha_k \geq 0, \ k = 1, \ldots, n\Big\}, \tag{15}$$

где векторы $a_1, \ldots, a_{n-1} \in \mathbb{R}^n$ строятся как и ранее и имеют последнюю компоненту нулевую, а $e = (0, \ldots, 0, 1) \in \mathbb{R}^n$. При этом

$$\begin{gathered} \Xi(x_0) = \Big\{x \in \mathbb{R}^n : \ x = N(x_0) + \sum_{k=1}^{n-1} \alpha_k c_k(x_0) + \alpha_n g(x_0); \quad \alpha_k \geq 0, \ k = 1, \ldots, n\Big\}, \\ g(x_0) = L(A, \Lambda, x_0) e \,. \end{gathered} \tag{16}$$

Случай $Q_{2m+1} \leq O$ сводится к предыдущему умножением $(2m+1)$-го уравнения на -1. Если же матрица Q_{2m+1} знакопеременна, то в (15) и (16) ограничение неотрицательности на коэффициент α_n снимается.

Выбирая в вышеописанной схеме различные начальные точки $x_0 = x_0^{(l)}$, $l = 1, \ldots, s$, в качестве внешней аппроксимации $\operatorname{conv}(X^*)$ получаем выпуклый многогранник

$$\widehat{\Xi} = \bigcap_{l=1}^{s} \Xi\big(x_0^{(l)}\big). \tag{17}$$

На практике наряду с параметрически заданными конусами (14) и (16) удобно иметь аппроксимации $\widetilde{\Xi} \supset \operatorname{conv}(X^*)$, определяемые системами линейных неравенств и допускающие поэтому более простую геометрическую интерпретацию. Простейший способ построения $\widetilde{\Xi}$ сводится к тому, что выбирается пробная точка ξ и вычисляется ее проекция $p(\xi)$ на конус $\Xi(x_0)$. Предположим, что $\xi \notin \Xi(x_0)$. Обозначая через

$$\Pi\big(\Xi(x_0), \xi\big) = \big\{x \in \mathbb{R}^n : \ \big(\xi - p(\xi), x - p(\xi)\big) \leq 0\big\} \tag{18}$$

опорное полупространство к $\Xi(x_0)$ с внешней нормалью $\xi - p(\xi)$, проходящее через точку $p(\xi)$, получаем

$$\operatorname{conv}(X^*) \subset \Xi(x_0) \subset \Pi\big(\Xi(x_0), \xi\big).$$

Нахождение проекции $p(\xi)$ сводится к задаче квадратичного программирования с простыми ограничениями. Например, в случае $n = 2m$ имеем $p(\xi) = N(x_0) + \sum_{k=1}^{n} \alpha_k^* c_k(x_0)$, вектор $(\alpha_1^*, \ldots, \alpha_n^*)$ является решением задачи

$$\min\Big\{\Big\|\xi - N(x_0) - \sum_{k=1}^{n} \alpha_k c_k(x_0)\Big\|^2 : \ \alpha_k \geq 0, \quad k = 1, \ldots, n\Big\}.$$

Выбрав различные точки ξ_l, $l = 1, \ldots, r$, положим

$$\widetilde{\Xi} = \bigcap_{l=1}^{r} \Pi\big(\Xi(x_0), \xi_l\big). \tag{19}$$

Очевидно, что $\widetilde{\Xi}$ представляет собой выпуклый многогранный конус, содержащий $\Xi(x_0)$.

Тот же прием применим и к многограннику $\widehat{\Xi}$, определенному в (17) и доставляющему внешнюю выпуклую аппроксимацию $\operatorname{conv}(X^*)$. При этом полагаем

$$\widetilde{\widehat{\Xi}} = \bigcap_{l=1}^{r} \Pi\big(\widehat{\Xi}, \xi_l\big). \tag{20}$$

Пусть, например, $n = 2m$ и $s = 2$. Тогда проекция $p(\xi)$, определяющая опорное полупространство $\Pi(\widehat{\Xi}, \xi)$, имеет вид $p(\xi) = N(x_0^{(1)}) + \sum_{k=1}^n \alpha_k^{(1)*} c_k(x_0^{(1)})$, где для отыскания вектора $(\alpha_1^{(1)*}, \dots, \alpha_n^{(1)*})$ решается задача квадратичного программирования с простыми ограничениями-неравенствами:

$$\min\Big\{\Big\|\xi - N(x_0^{(1)}) - \sum_{k=1}^n \alpha_k^{(1)} c_k(x_0^{(1)})\Big\|^2 :\ N(x_0^{(1)}) + \sum_{k=1}^n \alpha_k^{(1)} c_k(x_0^{(1)}) = N(x_0^{(2)}) + \sum_{k=1}^n \alpha_k^{(2)} c_k(x_0^{(2)});\ \alpha_k^{(1)},\ \alpha_k^{(2)} \ge 0,\ k = 1, \dots, n\Big\}. \tag{21}$$

4. Основной алгоритм и численные эксперименты

Вышеприведенные построения приводят к следующему алгоритму аппроксимации множества решений X^* системы (1)–(3).

Алгоритм 2.

1. Для каждой из матриц $D_j = Q_{2j-1} + iQ_{2j}$ с использованием алгоритма 1 определяются конусы $\overline{K}_j$, где $K_j = \mathrm{cone}(W(D_j))$, $j = 1, \dots, m$; $m = [n/2]$.
2. В зависимости от четности n определяются векторы $(a_1, \dots, a_n)$ или $(a_1, \dots, a_{n-1}, e)$, описывающие согласно (12) или (15) внешнюю выпуклую коническую аппроксимацию $\widetilde{\Lambda}$ для $\Lambda[\mathbb{R}^n]^2$.
3. Для выбранного начального приближения $x_0 \in \mathbb{R}^n$ согласно (4) определяется приближение метода Ньютона $N(x_0)$.
4. Согласно (14) или (16), определяется внешняя выпуклая коническая аппроксимация $\Xi(x_0)$ для $\mathrm{conv}(X^*)$.
5. При необходимости для выбранного набора точек $x_0^{(l)}$, $l = 1, \dots, s$, согласно (17) строится внешняя выпуклая многогранная аппроксимация $\widehat{\Xi}$ для $\mathrm{conv}(X^*)$. Для $\widehat{\Xi}$ в соответствии с (18), (20) может быть построена внешняя аппроксимация $\widetilde{\Xi}$, определяемая системой линейных неравенств.

Первые два шага алгоритма реализуются только один раз, в отличие от них шаги 3–5 могут повторяться неоднократно с различными начальными точками.

Приведем примеры применения алгоритма. Ограничимся системами (1) при $n = 3, 4$. Все вычисления проводятся с точностью 10^{-4}.

Пример 1. Пусть $n = 3$, матрицы Q_j, $j = 1, 2, 3$, имеют вид:

$$Q_1 = \begin{pmatrix} 95.8 & 7.6 & 61.2 \\ 7.6 & 4 & 3.6 \\ 61.2 & 3.6 & 39.6 \end{pmatrix},\ Q_2 = \begin{pmatrix} 101.8 & 17.6 & 57.2 \\ 17.6 & 16 & 1.6 \\ 57.2 & 1.6 & 37.6 \end{pmatrix},\ Q_3 = \begin{pmatrix} 9.36 & 2.16 & 5.28 \\ 2.16 & 1.44 & 0.72 \\ 5.28 & 0.72 & 3.28 \end{pmatrix}.$$

Матрицу A и элемент f определим следующим образом:

$$A = \begin{pmatrix} 7.2 & 3.6 & 3.6 \\ 3.2 & 1.6 & 1.6 \\ 1.44 & 0.72 & 0.72 \end{pmatrix}, \quad f = \begin{pmatrix} 1.72 \\ 0.48 \\ 0.3312 \end{pmatrix}.$$

Множество решений системы (1) с указанными данными есть

$$X^* = \big\{(1.2, -1.4, -1.8)^\top, (2.0, -2.2, -3.0)^\top, (-2.0, -1.8, -3.0)^\top, (2.8, -2.6, -4.2)^\top\big\}.$$

Для заданных Q_1, Q_2 конус $\overline{K}_1$ определяется системой неравенств:

$$\begin{cases} 0.0843\,x_1 - 135.5716\,x_2 \le 0, \\ 137.4066\,x_1 - 0.0251\,x_2 \ge 0. \end{cases}$$

Учитывая, что $Q_3 \ge O$, аппроксимируем $\Lambda[\mathbb{R}^3]^2$ выпуклым конусом

$$\widetilde{\Lambda} = \left\{ \alpha_1 \begin{pmatrix} 135.5716 \\ 0.0843 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0.0251 \\ 137.4066 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} : \alpha_1, \alpha_2, \alpha_3 \ge 0 \right\}.$$

Выберем начальное приближение $x_0^{(1)} = (1, 1, 1)^\top$ и, выполнив шаг метода Ньютона (4), получим

$$N\big(x_0^{(1)}\big) = \big(1.5120, -0.5020, -1.0180\big)^\top.$$

Реализуя п. 4 алгоритма 2, получаем для $\operatorname{conv}(X^*)$ следующую внешнюю выпуклую коническую аппроксимацию:

$$\begin{aligned} &\Xi\big(x_0^{(1)}\big) = N\big(x_0^{(1)}\big) + L\big(A, \Lambda, x_0^{(1)}\big)\widetilde{\Lambda} \\ &= \left\{ \begin{pmatrix} 1.5120 \\ -0.5020 \\ -1.0180 \end{pmatrix} + \alpha_1 \begin{pmatrix} 2.4031 \\ 0.2034 \\ -4.4858 \end{pmatrix} + \alpha_2 \begin{pmatrix} -6.9040 \\ 4.8815 \\ 10.1326 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0.3007 \\ -0.3641 \\ -0.4149 \end{pmatrix} : \alpha_1, \alpha_2, \alpha_3 \ge 0 \right\}. \end{aligned}$$

Для нового начального приближения $x_0^{(2)} = (0, -4, -4)^\top$ аналогично получаем $N\big(x_0^{(2)}\big) = \big(0.9957, -3.0007, -3.4936\big)^\top$ и имеем другую коническую аппроксимацию

$$\begin{aligned} &\Xi\big(x_0^{(2)}\big) = N\big(x_0^{(2)}\big) + L\big(A, \Lambda, x_0^{(2)}\big)\widetilde{\Lambda} \\ &= \left\{ \begin{pmatrix} 0.9957 \\ -3.0007 \\ -3.49360 \end{pmatrix} + \alpha_1 \begin{pmatrix} -0.7568 \\ -0.5464 \\ 1.6859 \end{pmatrix} + \alpha_2 \begin{pmatrix} 4.0248 \\ -3.0135 \\ -5.8975 \end{pmatrix} + \alpha_3 \begin{pmatrix} -0.2355 \\ 0.2672 \\ 0.3307 \end{pmatrix} : \alpha_1, \alpha_2, \alpha_3 \ge 0 \right\}. \end{aligned}$$

Таким образом, $\operatorname{conv}(X^*) \subset \widehat{\Xi} = \Xi\big(x_0^{(1)}\big) \cap \Xi\big(x_0^{(2)}\big)$.

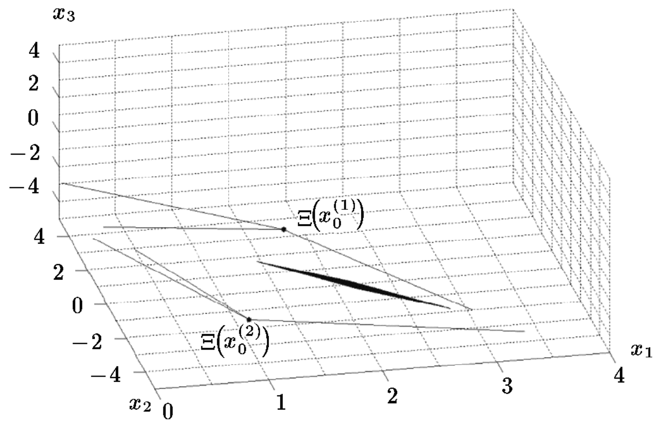


Рис. 1. Пример 1. Выпуклая оболочка решений и конусы

На рис. 1 конусы $\Xi\big(x_0^{(1)}\big)$ и $\Xi\big(x_0^{(2)}\big)$ показаны в виде трехгранных углов, содержащих внутри множество $\operatorname{conv}(X^*)$, изображенное в виде черного четырехугольника.

Пример 2. Положим $n = 4$,

$$Q_1 = \begin{pmatrix} 625 & -480 & -540 & 115 \\ -480 & 576 & 0 & -192 \\ -540 & 0 & 1296 & 108 \\ 115 & -192 & 108 & 73 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 625 & -270 & -960 & 10 \\ -270 & 324 & 0 & -108 \\ -960 & 0 & 2304 & 192 \\ 10 & -108 & 192 & 52 \end{pmatrix},$$

$$Q_3 = \begin{pmatrix} 144 & -72 & 72 & -144 \\ -72 & 520 & -476 & -280 \\ 72 & -476 & 436 & 248 \\ -144 & -280 & 248 & 400 \end{pmatrix}, \quad Q_4 = \begin{pmatrix} 16 & -8 & 8 & -16 \\ -8 & 4360 & -3964 & -3160 \\ 8 & -3964 & 3604 & 2872 \\ -16 & -3160 & 2872 & 2320 \end{pmatrix},$$

$$A = \begin{pmatrix} 680 & -384 & -864 & 56 \\ 820 & -216 & -1536 & -56 \\ -504 & -144 & 108 & 792 \\ -56 & -3536 & 3212 & 2648 \end{pmatrix}, \quad f = \begin{pmatrix} -366 \\ -534 \\ -864 \\ -1376 \end{pmatrix}.$$

Множество X^* в этом примере состоит из следующих 16 точек:

(1.2187, 1.7070, 1.4183, −0.9257)	(−0.1911, 0.1783, 0.5860, −1.9873)
(−1.9915, −2.4968, 0.4628, −5.5117)	(−3.4013, −4.0255, −0.3694, −6.5732)
(−0.9979, −0.6242, 0.6157, −2.3779)	(−2.4076, −2.1529, −0.2166, −3.4395)
(−4.2081, −4.8280, −0.3397,−6.9639)	(−5.6178, −6.3567, −1.1720, −8.0255)
(4.2845, 4.3567, 2.5053, 1.3588)	(2.8747, 2.8280, 1.6730, 0.2972)
(1.0743, 0.1529, 1.5499, −3.2272)	(−0.3355, −1.3758, 0.7176, −4.2887)
(2.0679, 2.0255, 1.7028, −0.0934)	(0.6582, 0.4968, 0.8705, −1.1550)
(−1.1423, −2.1783, 0.7473,−4.6794)	(−2.5520, −3.7070, −0.0849, −5.7410)

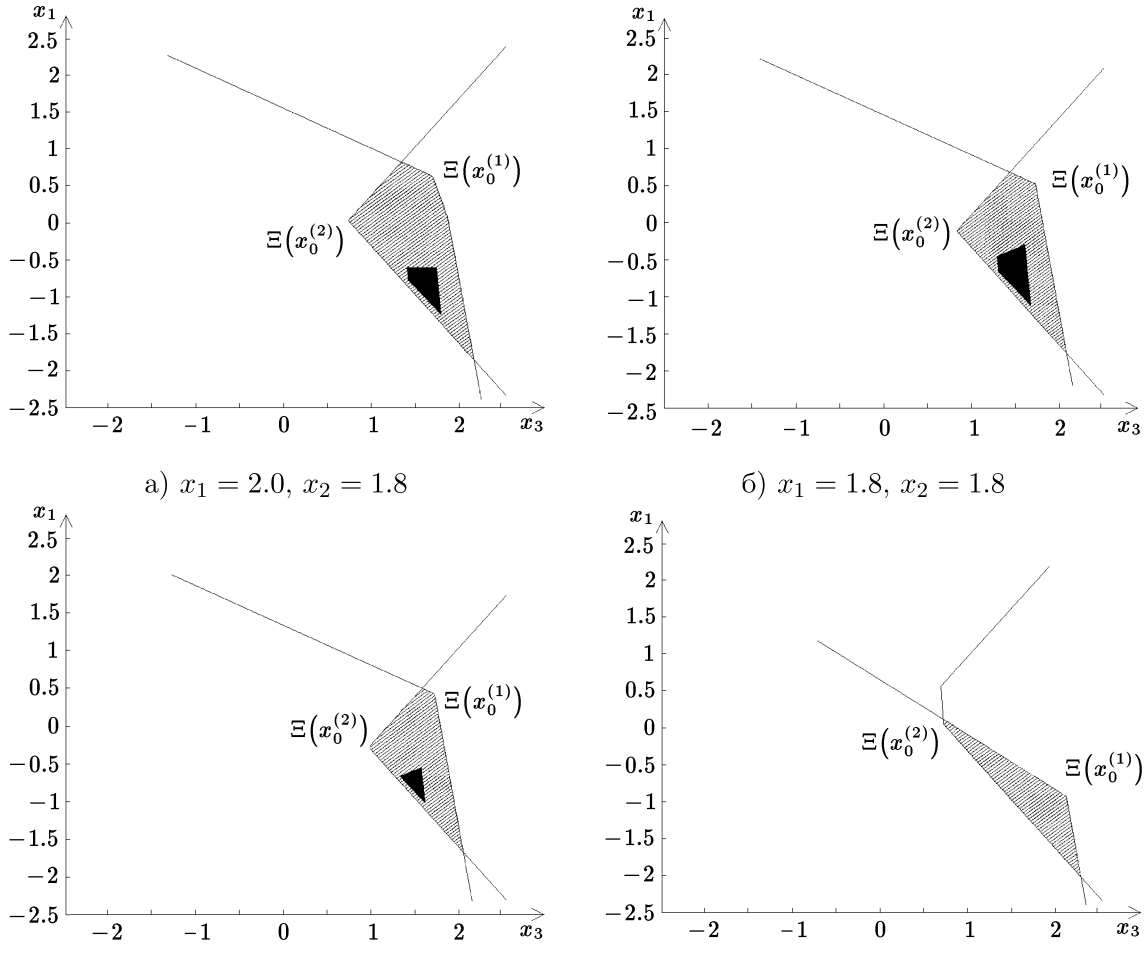
Выберем $x_0^{(1)} = (2, 2, 2, 2)$ и получим выпуклый конус

$$\Xi\bigl(x_0^{(1)}\bigr) = N\bigl(x_0^{(1)}\bigr)+L\bigl(A, \Lambda, x_0^{(1)}\bigr)\widetilde{\Lambda}= \left\{ \begin{pmatrix} 1.2819 \\ 1.1907 \\ 1.5521 \\ 0.0341 \end{pmatrix} + \alpha_1 \begin{pmatrix} -0.4707 \\ -0.3819 \\ -0.1402 \\ -0.3498 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0.1271 \\ 0.0655 \\ -0.0024 \\ 0.0931 \end{pmatrix} + \right.$$
$$\left. \alpha_3 \begin{pmatrix} -0.1100 \\ -0.1434 \\ -0.0329 \\ -0.1552 \end{pmatrix} + \alpha_4 \begin{pmatrix} 0.0345 \\ 0.0394 \\ 0.0117 \\ 0.0319 \end{pmatrix} : \alpha_1, \alpha_2, \alpha_3, \alpha_4 \geq 0 \right\}.$$

Выбирая новое начальное приближение $x_0^{(2)} = (-2, -2, -2, -2)$, будем иметь другой выпуклый конус, содержащий $\mathrm{conv}(X^*)$:

$$\Xi\bigl(x_0^{(2)}\bigr)=N\bigl(x_0^{(2)}\bigr)+L\bigl(A, \Lambda, x_0^{(2)}\bigr)\widetilde{\Lambda}= \left\{ \begin{pmatrix} 0,0802 \\ -0,0693 \\ -0,1895 \\ -1,4085 \end{pmatrix} + \alpha_1 \begin{pmatrix} -1.7897 \\ -1.5551 \\ -0.6435 \\ -1.3339 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1.0526 \\ 0.9245 \\ 0.3890 \\ 0.7849 \end{pmatrix} + \right.$$
$$\left. \alpha_3 \begin{pmatrix} -0.0959 \\ -0.1286 \\ -0.0278 \\ -0.1460 \end{pmatrix} + \alpha_4 \begin{pmatrix} -0.0920 \\ -0.0936 \\ -0.0341 \\ -0.0510 \end{pmatrix} : \alpha_1, \alpha_2, \alpha_3, \alpha_4 \geq 0 \right\}.$$

Таким образом, $\mathrm{conv}(X^*) \subset \widehat{\Xi} = \Xi\bigl(x_0^{(1)}\bigr) \cap \Xi\bigl(x_0^{(2)}\bigr)$. Наглядное представление о многограннике $\widehat{\Xi} \subset \mathbb{R}^4$ можно получить, построив его сечения двумерными плоскостями.



а) $x_1 = 2.0$, $x_2 = 1.8$ б) $x_1 = 1.8$, $x_2 = 1.8$

в) $x_1 = 1.6$, $x_2 = 1.8$ г) $x_1 = 2.4$, $x_2 = 1.8$

Рис. 2. Пример 2. Сечения $\mathrm{conv}(X^*)$ и $\Xi(x_0^{(1)})$, $\Xi(x_0^{(2)})$

На рис. 2 показаны некоторые сечения конусов $\Xi(x_0^{(1)})$, $\Xi(x_0^{(2)})$, сечения $\mathrm{conv}(X^*)$ изображены черными многоугольниками, сечения $\widehat{\Xi} = \Xi(x_0^{(1)}) \cap \Xi(x_0^{(2)})$ заштрихованы. При $x_1 = 2.4$, $x_2 = 1.8$ сечение $\mathrm{conv}(X^*)$ пусто (рис. 2, г).

Геометрическое представление о многогранном множестве $\widehat{\Xi}$ можно получить также, выбрав различные точки ξ_l и записав согласно (20) пересечение соответствующих опорных полупространств. Для точек $\xi_l = (\pm 8, \pm 8, \pm 8, \pm 8)$, $l = 1, \ldots, 16$, решая задачи (21), получаем следующий набор линейных неравенств, описывающих внешнюю многогранную аппроксимацию $\widehat{\Xi}$:

$$\begin{cases} -0.4618x_1 - 0.4044x_2 + 1.9171x_3 + 0.2945x_4 \le 1.9121; \\ -0.8218x_1 - 2.3433x_2 + 2.0929x_3 + 2.8249x_4 \le -0.4991; \\ -4.6754x_1 + 4.3785x_2 - 4.1930x_3 + 3.1911x_4 \le -4.3785; \\ -0.3655x_1 + 2.4478x_2 - 2.2419x_3 - 1.2818x_4 \le 2.0312; \\ -0.5370x_1 - 0.8098x_2 + 0.7118x_3 + 1.3212x_4 \le -0.5029; \\ -3.2329x_1 + 3.8323x_2 - 5.9744x_3 + 2.7830x_4 \le -3.3126. \end{cases} \tag{22}$$

Точки (8, 8, 8, 8), (8, 8, −8, 8), (8, 8, −8,−8), (8, −8, 8, 8), (8, −8, 8, −8), (8, −8, −8, 8), (8, −8, −8, −8), (−8, 8, −8, −8), (−8, −8, 8, 8), (−8, −8, 8, −8) оказались лежащими внутри $\widehat{\Xi}$. Сечения опорных полупространств и многогранника $\mathrm{conv}(X^*)$ двумерной плоскостью $x_1 = 1.8$, $x_2 = 1.8$ показаны на рис. 3. Номер полуплоскости соответствует порядковому номеру неравенства в (22), описывающего опорное полупространство. Сечение множества решений системы (22) на рис. 3 заштриховано.

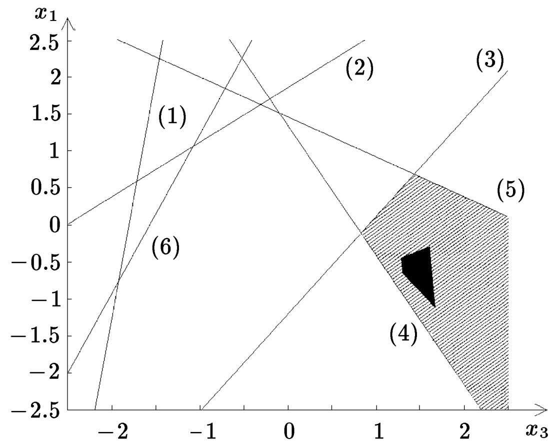


Рис. 3. Пример 2. Сечение $\mathrm{conv}(X^*)$ и полупространств (22)

Пример 3. Рассмотрим систему уравнений:

$$\begin{cases} x_2^2 - x_1x_3 = 0; \\ x^2 - 2x_1x_3 = 0; \\ x_1^2 + x_2^2 + 2x_1 = 0; \\ x_3^2 + x_4^2 - 1 = 0. \end{cases} \tag{23}$$

Множество решений X^* системы (23) состоит из окружности $x_1 = x_2 = 0$, $x_3^2 + x_4^2 = 1$ и двух отдельных точек $(-2, 0, 0, -1)$, $(-2, 0, 0, 1)$. Конус $\overline{K}_1$ совпадет с $\mathbb{R}^2$, $\overline{K}_2$ является положительным квадрантом $\mathbb{R}^2$. Сечение $\mathrm{conv}(X^*)$ гиперплоскостью $x_2 = 0$ показано на рис. 4.

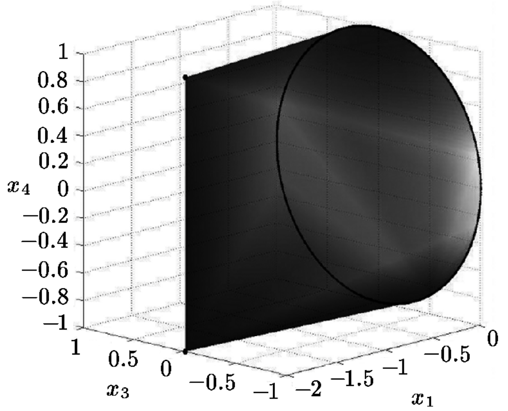


Рис. 4. Пример 3. Сечение $\mathrm{conv}(X^*)$ гиперплоскостью $x_3 = 0$

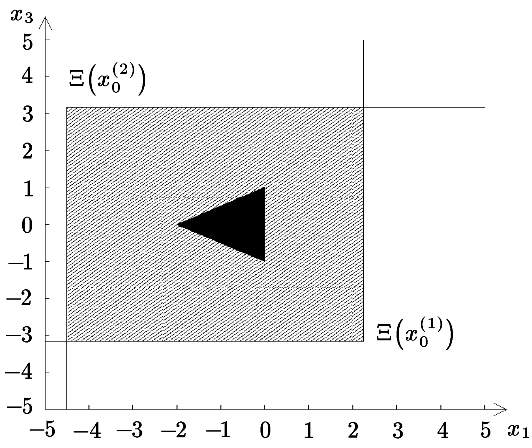


Рис. 5. Пример 3. Сечения $\mathrm{conv}(X^*)$ и $\Xi\big(x_0^{(1)}\big)$, $\Xi\big(x_0^{(2)}\big)$ плоскостью $x_2 = 0$, $x_4 = 0$

Выберем начальные приближения $x_0^{(1)} = (3, 3, -3, 3)$ и $x_0^{(2)} = (-3, 3, 3, 3)$, получим конусы $\Xi\big(x_0^{(1)}\big)$, $\Xi\big(x_0^{(2)}\big)$, содержащие $\mathrm{conv}(X^*)$. Сечением $\mathrm{conv}(X^*)$ плоскостью $x_2 = 0$, $x_4 = 0$ является треугольник с вершинами $(-2, 0)$, $(0, 1)$, $(0, -1)$. На рис. 5 показаны сечения конусов $\Xi\big(x_0^{(1)}\big)$, $\Xi\big(x_0^{(2)}\big)$ этой плоскостью, черным треугольником изображено сечение $\operatorname{conv}(X^*)$ и заштриховано сечение многогранника $\widehat{\Xi} = \Xi\big(x_0^{(1)}\big) \cap \Xi\big(x_0^{(2)}\big)$.

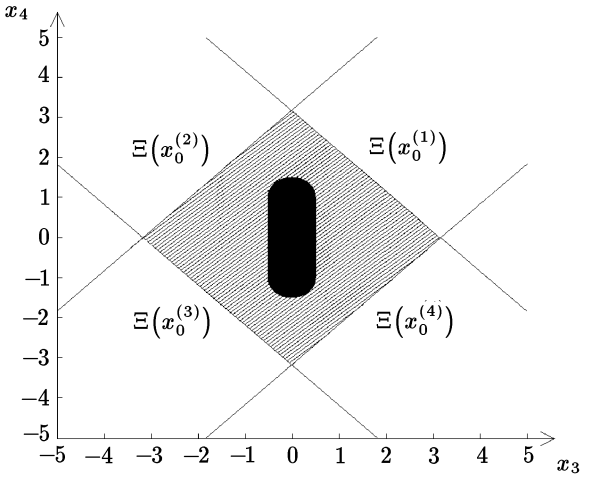


Рис. 6. Пример 3. Сечения $\operatorname{conv}(X^*)$ и $\widehat{\Xi}$ плоскостью $x_1 = -1$, $x_2 = 0$

Для других начальных точек $x_0^{(1)} = (3,3,3,3)$, $x_0^{(2)} = (3,3,-3,3)$, $x_0^{(3)} = (3,3,-3,-3)$, $x_0^{(4)} = (-3,3,3,-3)$ на рис. 6 показаны сечения конусов $\Xi\big(x_0^{(j)}\big)$, $j = 1,\ldots,4$, плоскостью $x_1 = -1$, $x_2 = 0$. Черным выделено сечение $\operatorname{conv}(X^*)$ и заштриховано сечение многогранника $\widehat{\Xi} = \bigcap\limits_{j=1}^{4} \Xi\big(x_0^{(j)}\big)$ этой плоскостью.

5. Обсуждение

Разработка методов отыскания или аппроксимации всего множества решений систем полиномиальных уравнений привлекает в последние годы растущее внимание исследователей. Наибольшее развитие в этом направлении получила техника, основанная на идеях гомотопической редукции к исходной системе некоторой модельной системы, решения которой записываются в явном виде [6]. Алгоритмы, использующие эту технику, предназначены для систем с изолированными решениями и включают трудоемкие процедуры численного интегрирования нелинейных систем дифференциальных уравнений с различными начальными условиями. Известны алгебраические подходы, основанные на последовательном исключении неизвестных с использованием базисов Гребнера [7]. К представленной в работе схеме близки методы внешней аппроксимации множества решений полиномиальных систем, основанные на представлении полиномов на кубе $[0,1]^n$ в форме Бернштейна [8]. Эти методы имеют комбинаторный характер и используют последовательное исключение подобластей, полученных половинным делением исходного куба, которые заведомо не могут содержать точек из множества решений. В отличие от перечисленных трудоемких подходов, ориентированных на точную аппроксимацию невыпуклого множества решений, предлагаемая схема имеет целью эффективное отыскание внешних аппроксимаций выпуклой оболочки этого множества. Отметим, что ограничение уравнениями с полиномами второй степени формально не уменьшает общности построений, поскольку произвольные полиномиальные системы сводятся к квадратичным путем последовательных замен вида $y_{ij} = x_i x_j$.

Литература

1. **Calabi E.** Linear systems of real quadratic forms. II // Proc. Amer. Math. Soc. — 1982. — Vol. 84, № 3. — P. 331–334.
2. **Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.** Численные методы. — М.: БИНОМ, 2007.

3. **Годунов С.К.** Современные аспекты линейной алгебры. — Новосибирск: Научная книга, 1997.

4. **Кокурин М.Ю.** О кривой П.А. Широкова и теоремах Хаусдорфа и Дайнса // Ученые записки Казанского государственного университета. Сер. Физико-математические науки. — 2009. — Т. 151, кн. 4. — С. 51–53.

5. **Cowen C.C., Harel E.** An Effective Algorithm for Computing the Numerical Range. — West Lafayette, Indiana: Purdue University, Department of Mathematics, 1995. — (Technical report).

6. **Morgan A.** Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems. — Englewood Cliffs, N.J.: Prentice Hall, 1987.

7. **Кокс Д., Литтл Дж., О′Ши Д.** Идеалы, многообразия и алгоритмы. Введение в вычислительные аспекты алгебраической геометрии и коммутативной алгебры. — М.: Мир, 2000.

8. **Garloff J., Smith A.P.** Investigation of a subdivision based algorithm for solving systems of polynomial equations // J. of Nonlinear Analysis. — 2001. — Vol. 47, № 1. — P. 167–178.

Поступила в редакцию 27 ноября 2012 г.