АЛГЕБРАИЧЕСКИЕ МЕТОДЫ РЕШЕНИЯ ЗАДАЧ НЕЛИНЕЙНОЙ ОПТИМИЗАЦИИ*

А. Ю. Утешев

alexeiuteshev@gmail.com

19 марта 2015 г.

1°. Постановка задачи. Найти

$$\max_{X\in\mathbb{R}^n} F(X)$$

и

$$\max_{X\in\mathbb{R}^n} F(X) \quad \text{s.t.} \quad \left\{ \begin{array}{c} G(X) \geqslant 0 \\ \text{или} \\ G_1(X) \geqslant 0, \dots, G_K(X) \geqslant 0. \end{array} \right.$$

Здесь F, G, $G_1, \dots G_K$ — полиномы.

Методы решения:

- исключение переменных в системах алгебраических уравнений;
- локализация решений таких систем по методу Эрмита.

Замечание 1. Нет предположений о выпуклости.

2°. Примеры конкретных задач.

А. Расстояние между алгебраическими многообразиями.

Расстояние до квадрики (эллипсоида) в $\mathbb{R}^n$

$$X^T \mathbf{A} X + 2\, B^T X - 1 = 0$$

- от точки X_0: $\max |X_0 X|^2$;
- от линейного многообразия $\mathbf{C}Y = H$: $\max |XY|^2$;
- от другой квадрики.

*Семинар по конструктивному негладкому анализу и недифференцируемой оптимизации «CNSA & NDO»: `http://www.apmath.spbu.ru/cnsa/`

B. Стационарные точки функции

$$F(X) = \sum_{j=1}^{N} m_j |XP_j|^L \quad \text{при} \quad \begin{array}{c} \{P_j\}_{j=1}^N \subset \mathbb{R}^n, \\ \{L, \{m_j\}_{j=1}^N\} \subset \mathbb{R}. \end{array}$$

Известные частные случаи задачи:

- $L = 2$: центр масс (барицентр);
- $L = 1$: обобщенная задача Ферма — Торричелли;
- $L = -1$: кулоновский потенциал системы стационарных зарядов или ньютоновский потенциал системы неподвижных материальных точек.

В докладе будут рассмотрены только задачи группы **A.**

3°. Вид решения. На примерах покажем что будет считаться решением задачи.

ПРИМЕР 1. Найти

$$\max_{x \in \mathbb{R}} (-x^6 + 12\, x^2 + 12\, x + 2) \ .$$

Решение. Построим полином, корнями которого являются стационарные значения целевой функции:

$$\mathcal{F}(z) = -z^5 + 10\, z^4 + 472\, z^3 + 16208\, z^2 - 16272\, z - 32800 \ .$$

Утверждается, что если

$$\begin{array}{c} \lambda_1, \lambda_2, \dots \ \text{— корни } F'(x) = 0, \\ \text{то} \\ F(\lambda_1), F(\lambda_2), \dots \ \text{— корни } \mathcal{F}(z) = 0. \end{array}$$

Таким образом:

$$\begin{array}{c} \max = \text{максимальный положительный корень } \mathcal{F}(z), \text{ т.е.} \\ \max \approx 35.6321. \end{array}$$

ПРИМЕР 2. Найти
$$\max\left(\begin{array}{c} -x^4+x^3y+x^2y^2-3\,y^4+ \\ +5\,x^3-2\,x^2y+5\,xy^2- \\ -2\,x^2-3\,xy+4\,y^2-x+2\,y \end{array}\right).$$

Решение. В отличие от предыдущего примера, в этом примере приходится производить дополнительную проверку конечности искомого max. В общем случае, эта задача сводится к выводу условий знакоопределенности однородного полинома (формы) и может быть решена в рамках излагаемой в настоящем докладе идеологии. Не будем на ней здесь останавливаться.

$\max =$ максимальный положительный корень полинома

$$\mathcal{F}(z)=190437227429888\,z^9-237685826728591360\,z^8+$$
$$+\ldots-6432682928247872599337\,z-87899390126339067066 3,$$

$$\max\approx 1098.6931.$$

4°. Как это работает? Теперь объясним как получились ответы в примерах предыдущего параграфа.

Для этого используются методы **теории исключения** (Elimination Theory): система нелинейных алгебраических уравнений от нескольких переменных преобразуется к эквивалентному виду, в котором в одном из уравнений остается только одна переменная (все остальные *исключаются*).

Ключевые понятия теории: **результант** и **дискриминант** полиномов одной и нескольких переменных.

В случае $\mathbb{R}^1$ формально результант полиномов $F(x)$ и $G(x)$ определяется как[1]:
$$\mathcal{R}(F(x),G(x))=\prod_{j=1}^{\deg G}F(\lambda_j),$$
где $\lambda_1,\lambda_2,\ldots$ — корни полинома $G(x)$.

Фактически же:
$$\mathcal{R}(F(x),G(x))\text{ — полиномиальная функция коэффициентов.}$$

[1] На самом деле, с точностью до сомножителя, но здесь тонкости проигнорируем.

Детерминантное представление (одно из возможных) [1]:

$$\mathcal{R}_x(A_0x^N + A_1x^{N-1} + \cdots + A_N, B_0x^M + B_1x^{M-1} + \cdots + B_M) =$$

$$= \begin{vmatrix} A_0 & A_1 & A_2 & \dots & \dots & A_N & 0 & \dots & 0 & 0 \\ 0 & A_0 & A_1 & \dots & \dots & A_{N-1} & A_N & \dots & 0 & 0 \\ \vdots & & \ddots & & & & & & \ddots & \\ 0 & 0 & \dots & A_0 & \dots & \dots & & \dots & A_{N-1} & A_N \\ 0 & 0 & \dots & & B_0 & B_1 & \dots & \dots & B_{M-1} & B_M \\ 0 & 0 & \dots & B_0 & B_1 & \dots & & \dots & B_M & 0 \\ \vdots & & & \dots & & & & & & \vdots \\ 0 & B_0 & \dots & \dots & & B_M & \dots & & \dots & 0 \\ B_0 & \dots & \dots & & B_M & 0 & \dots & & & 0 \end{vmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}A\\A\\A\\A\end{matrix}}\right\} M \\ \left.\vphantom{\begin{matrix}A\\A\\A\\A\\A\end{matrix}}\right\} N \end{array}$$

Формально в случае $\mathbb{R}^2$ результант определяется, так:

$$\mathcal{R}_{x,y}(F(x,y), G_1(x,y), G_2(x,y)) = \prod_{j=1}^{\deg G_1 \times \deg G_2} F(\alpha_j, \beta_j).$$

Здесь (α_j, β_j) — корни системы уравнений

$$G_1(x,y) = 0, \qquad G_2(x,y) = 0.$$

Фактически: $\mathcal{R}(F(x,y), G_1(x,y), G_2(x,y))$ — полиномиальная функция коэффициентов.

Практически вычисляется:

- построением базиса Грёбнера;
- с помощью детерминантных подходов: в виде определителя блочно-ганкелевой матрицы (многомерные вычеты)

  или по **методу Безу** (посредством введения операции “деления” полинома многих переменных на систему полиномов многих переменных).

Дискриминант полинома определяется через результант[2] [1]:

$$\mathcal{D}(F(x)) = \mathcal{R}(F, F'),$$

$$\mathcal{D}(F(x,y)) = \mathcal{R}_{x,y}(F,\ \partial F/\partial x, \partial F/\partial y).$$

[2]На самом деле, с точностью до сомножителя, но здесь тонкости проигнорируем.

ТЕОРЕМА 1. *Если F имеет единственный кратный корень Λ_* второй кратности, то*

$$\Lambda_* = \textit{рациональная функция коэффициентов } F.$$

ПРИМЕР 3. $\mathcal{D}(A_0x^3 + A_1x^2 + A_2x + A_3) \equiv D(A_0, A_1, A_2, A_3) :=$

$$:= A_1^2A_2^2 - 4\,A_1^3A_3 - 4\,A_0A_2^3 + 18\,A_0A_1A_2A_3 - 27\,A_0^2A_3^2\ . \tag{1}$$

Если существует кратный корень Λ_* кубического полинома, то для него имеет место соотношение:

$$1 : \Lambda_* : \Lambda_*^2 : \Lambda_*^3 = \frac{\partial D}{\partial A_3} : \frac{\partial D}{\partial A_2} : \frac{\partial D}{\partial A_1} : \frac{\partial D}{\partial A_0}. \tag{2}$$

Аналогичный результат справедлив и для полиномов произвольной степени.

5°. Метрические задачи для квадрик. Рассматриваются многообразия в $\mathbb{R}^n$.

ТЕОРЕМА 2. *Квадрат расстояния от точки X_0 до многообразия*

$$X^\top \mathbf{A} X + 2\,B^\top X - 1 = 0\ , \quad (\mathbf{A} = \mathbf{A}^\top),$$

равен минимальному положительному корню ***уравнения расстояний***

$$\mathcal{F}(z) = 0 \quad \textit{при} \quad \mathcal{F}(z) := \mathcal{D}_\mu\left(\Phi(\mu, z)\right)\ .$$

Здесь

$$\Phi(\mu, z) := \det\left(\left[\begin{array}{cc} \mathbf{A} & B \\ B^\top & -1 \end{array}\right] + \mu\left[\begin{array}{cc} -E & X_0 \\ X_0^\top & z - X_0^\top X_0 \end{array}\right]\right)\ , \tag{3}$$

а E — единичная матрица порядка n. Дополнительно предполагается, что указанный корень не является кратным.

ПРИМЕР 4. Для вычисления расстояние от точки (x_0, y_0) до эллипса

$$G(x, y) := \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$$

строим сначала полином (3):

$$\Phi(\mu, z) = \mu^3 + A_1\mu^2 + A_2\mu + A_3\ .$$

Здесь

$$
\begin{array}{rcl}
A_1 & = & x_0^2+y_0^2-a^2-b^2-z \\
A_2 & = & a^2b^2\left\{\left(\dfrac{1}{a^2}+\dfrac{1}{b^2}\right)z-G(x_0,y_0)\right\}, \\
A_3 & = & -a^2b^2\,z\ .
\end{array}
$$

Вычисляем дискриминант этого полинома (как полинома от переменной μ) по формуле (1). Уравнение расстояний:

$$
\begin{gathered}
\mathcal{F}(z):=(a^2-b^2)^2z^4- \\
-2\,(a^2-b^2)\Big\{(a^2-b^2)(a^2+b^2+x_0^2+y_0^2)+a^2y_0^2-b^2x_0^2\Big\}z^3+ \\
+\dots+ \\
+a^4b^4G(x_0,y_0)^2\left\{(x_0^2+y_0^2-a^2-b^2)^2+4\,a^2b^2G(x_0,y_0)\right\}=0\,.
\end{gathered}
$$

ТЕОРЕМА 3. *Координаты ближайшей к X_0 точки на квадрике:*

$$
X_*=\textit{рациональные функции}\ (z_*)\ \ ,
$$

где z_* — минимальный положительный корень полинома $\mathcal{F}(z)$.

ПРИМЕР 5. Координаты ближайшей к точке (x_0,y_0) точки эллипса

$$
\frac{x^2}{a^2}+\frac{y^2}{b^2}-1=0
$$

определяются формулами

$$
x_*=\frac{a^2x_0}{a^2-\mu_*},\ \ y_*=\frac{b^2y_0}{b^2-\mu_*}
$$

при μ_*, вычисляемом по формуле (2):

$$
\mu_*=\frac{9\,A_3-A_1A_2}{2(A_1^2-3\,A_2)}
$$

при значениях A_1,A_2,A_3 взятых из предыдущего примера.

Другие метрические задачи с эллипсоидом в $\mathbb{R}^n$:

- расстояние от линейного многообразия $\mathbf{C}X=H$;
- расстояние от другой квадрики;
- расстояние от пересечения этих многообразий (например, от эллипса в $\mathbb{R}^3$).

ТЕОРЕМА 4. *Уравнение расстояний для любой из указанных задач строится посредством вычисления дискриминанта.*

Например, уравнение расстояний от эллипсоида до линейного многообразия

$$\mathbf{C}_{k\times n}X = H$$

имеет вид

$$\mathcal{F}(z) := \mathcal{D}_{\mu}\left(\mu^k \left|\begin{array}{ccc} \mathbf{A} & B & \mathbf{C} \\ B^T & -1+\mu z & -H^T \\ \mathbf{C}^T & -H & \dfrac{1}{\mu}\mathbf{C}^T\mathbf{C} \end{array}\right|\right) = 0\,.$$

Уравнение расстояний для двух центрированных (вложенных) эллипсоидов

$$X^T\mathbf{A}_1X = 1, \quad X^T\mathbf{A}_2X = 1$$

имеет вид:

$$\mathcal{F}(z) := \mathcal{D}_{\lambda}(\det\{\lambda\mathbf{A}_1 + (z-\lambda)\mathbf{A}_2 - \lambda(z-\lambda)\mathbf{A}_2\mathbf{A}_1\}) = 0\,.$$

Уравнение расстояний для двух произвольных эллипсоидов

$$X^T\mathbf{A}_1X + 2B_1^TX = 1, \quad X^T\mathbf{A}_2X + 2B_2^TX = 1$$

имеет вид:

$$\mathcal{F}(z) := \mathcal{D}_{\mu_1,\mu_2}\left(\det\left\{\begin{array}{l} \mu_1\left[\begin{array}{cc} \mathbf{A}_1 & B_1 \\ B_1^T & -1 \end{array}\right] + \mu_2\left[\begin{array}{cc} \mathbf{A}_2 & B_2 \\ B_2^T & -1 \end{array}\right] - \\ -\left[\begin{array}{cc} \mathbf{A}_2\mathbf{A}_1 & \mathbf{A}_2B_1 \\ B_2^T\mathbf{A}_1 & B_2^TB_1 - \mu_1\mu_2 z \end{array}\right] \end{array}\right\}\right) = 0\,.$$

6°. Преимущества и недостатки аналитического подхода.

“Плюсы”:

- достоверность (нет проблем с точностью вычислений и со сходимостью; нет проблемы различения глобальный - локальный экстремум);
- возможность исследования свойств решения в зависимости от параметров (возможность моделирования);
- изящество.

“Минусы”:

- существенна полиномиальность задачи (методы теории исключения работают только для алгебраических уравнений);
- дороговизна;
- неявность решения — как корня алгебраического уравнения.

ПРИМЕР 6. Уравнение расстояний для эллипсов

$$-\frac{1}{2}\,x_1^2+\frac{1}{2}\,x_1x_2-\frac{3}{2}\,x_2^2+\frac{5}{2}\,x_1+4\,x_2=1 \quad \text{и} \quad -\frac{1}{84}\,x_1^2-\frac{4}{189}\,x_2^2-\frac{1}{3}\,x_1=1$$

имеет вид

$$\mathcal{F}(z) := 12866891832025\,z^{12} - 2445505463588880\,z^{11} - 1086711163754965271 6\,z^{10} -$$

$$-3123865087697933253136\,z^9 + \ldots + 60730952901233749068462660878127980941312 = 0\,.$$

И такой рост размеров коэффициентов уравнения расстояний сравнительно с коэффициентами исходных многообразий — обычное дело.

ЛИТЕРАТУРА

1. Калинина Е.А., Утешев А.Ю. Теория исключения. Учеб. пособие. СПб.: НИИ Химии СПбГУ, 2002. 72 с.
2. Uteshev A.Yu., Yashina M.V. *Metric problems for quadrics in multidimensional space.* J. Symbolic Computation. 2015.
3. Uteshev A.Yu., Yashina M.V. *Distance computation from an ellipsoid to a linear or a quadric surface in* $\mathbb{R}^n$. Lecture Notes in Computer Science. 2007.
4. Утешев А.Ю., Яшина М.В. *Нахождение расстояния от эллипсоида до плоскости и квадрики в* $\mathbb{R}^n$. Доклады АН. 2008.
5. Uteshev A.Yu., Cherkasov T.M. *The search for the maximum of a polynomial.* J.Symbolic Computation. 1998.
6. Утешев А.Ю. Вычисление расстояний между геометрическими объектами. *Записная книжка профессора Утешева.* http://pmpu.ru/vf4/algebra2/optimiz/distance