ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Май—июнь 2014. Том 21, № 3. С. 87–102

УДК 519.85

# НАХОЖДЕНИЕ РАССТОЯНИЯ МЕЖДУ ЭЛЛИПСОИДАМИ *)

*Г. Ш. Тамасян, А. А. Чумаков*

**Аннотация.** Рассматривается задача нахождения ближайших точек между эллипсоидами. С использованием теории точных штрафных функций и аппарата негладкого анализа построены новые алгоритмы решения этой задачи. Предложены два итерационных метода (наискорейшего и гиподифференциального) спуска, которые (по сравнению с ранее известными) обладают определёнными преимуществами, в частности, они универсальные и менее трудоёмкие. Разработано программное обеспечение в системах MATLAB и Maple, реализующее эти алгоритмы.

**Ключевые слова:** негладкий анализ, ближайшее расстояние, эллипсоид, точные штрафные функции, субдифференциал, метод гиподифференциального спуска.

## Введение

Рассматривается следующая задача вычислительной геометрии: требуется найти ближайшие точки между эллипсоидами. Для решения данной задачи известны несколько подходов, которые можно условно поделить на «геометрические» [7, 12] и «алгебраические» [8, 10, 11].

В [12] предложен итерационный метод поиска минимального расстояния между двумя эллипсоидами. В нём на каждой итерации строятся шары, вписанные в эллипсоиды, касающиеся граничной точки эллипсоида. В [7] предложена модификация этого алгоритма. Заметим, что метод «шаров» не трудоёмкий с вычислительной точки зрения, однако он не применим для квадрик, отличных от эллипсоидов.

Поставленная проблема является задачей условной оптимизации и для её решения можно применить классический метод множителей Лагранжа. В результате получается система алгебраических уравнений. Полученную систему в [8, 10, 11] сводят к алгебраическому уравнению от

*) Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проекты 12–01–00752 и 14–01–31521 мол_а) и гранта СПбГУ 9.38.205.2014.

одной переменной, среди корней которого находится квадрат искомого расстояния. В [11] полином получается в результате применения методов теории исключения переменных [6] и имеет степень не менее $n(n+1)$.

В настоящей работе данная проблема условной оптимизации сводится к безусловной задаче с использованием теории точных штрафных функций [1, 2, 5]. Показано, что построенная штрафная функция является точной штрафной функцией, и её исследование производится с помощью аппарата негладкого анализа [3, 4]. Разработаны два новых итерационных метода наискорейшего и гиподифференциального спуска. Отметим, что наиболее трудоёмкие этапы в подобных методах [9], а именно, поиск направления спуска и величины шага спуска в этом направлении, вычисляются аналитически.

## 1. Формулировка задачи

Пусть $x \in \mathbb{R}^n$. Требуется найти ближайшие точки между эллипсоидами $\Omega_1$ и $\Omega_2$, где

$$\Omega_k = \left\{ x \in \mathbb{R}^n \mid h_k(x) = \frac{1}{2} x^T A_k x + x^T B_k + C_k = 0 \right\}, \quad k = 1, 2. \tag{1}$$

Здесь $A_k$ — постоянные симметричные и положительно определённые $(n \times n)$-матрицы, $B_k$ — постоянные $(n \times 1)$-вектор-столбцы, $C_k \in \mathbb{R}$ — константы.

Пусть $x_1, x_2 \in \mathbb{R}^n$. Поставленную проблему можно решить как задачу условной оптимизации

$$\|x_1 - x_2\| = \left[ \sum_{j=1}^{n} \left( x_1^j - x_2^j \right)^2 \right]^{\frac{1}{2}} \longrightarrow \min_{\substack{x_1 \in \Omega_1, \\ x_2 \in \Omega_2}} .$$

## 2. Точная штрафная функция

Пусть $z = (x_1, x_2)$. Введём функцию $f(z) = \dfrac{1}{2}\|x_1 - x_2\|^2$ и множество

$$Z = \{ z = (x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n \mid h_1(x_1) = 0, \ h_2(x_2) = 0 \} .$$

Множество $Z$ можно представить в эквивалентном виде

$$Z = \left\{ z \in \mathbb{R}^{2n} \mid \varphi(z) = 0 \right\} ,$$

где $\varphi(z) = |h_1(x_1)| + |h_2(x_2)|$.

Пусть $\lambda \geqslant 0$ фиксировано. Введём функцию $\Phi_\lambda(z) = f(z) + \lambda\varphi(z)$. Функция $\Phi_\lambda(z)$ называется *штрафной функцией*, а число $\lambda$ — *штрафом*. В [1, 2] представлен ряд теорем, при выполнении которых $\Phi_\lambda(z)$ является функцией точного штрафа.

Таким образом, задача минимизации функции $f(z)$ на множестве $Z$ сведена к задаче минимизации функции $\Phi_\lambda(z)$ на всём пространстве $\mathbb{R}^{2n}$ при $\lambda > \lambda^*$, где $\lambda^*$ — константа точного штрафа [1, 2], для которой можно найти оценки.

Пусть $z = (x_1, x_2)$, $y = (y_1, y_2) \in \mathbb{R}^{2n}$. Введём в $\mathbb{R}^{2n}$ евклидову норму

$$\|z\| = \sqrt{\|x_1\|^2 + \|x_2\|^2} = \sqrt{\sum_{k=1}^{n}\left(x_1^k\right)^2 + \sum_{k=1}^{n}\left(x_2^k\right)^2}$$

и метрику

$$\rho(z, y) = \|z - y\| = \sqrt{\sum_{k=1}^{n}\left(x_1^k - y_1^k\right)^2 + \sum_{k=1}^{n}\left(x_2^k - y_2^k\right)^2}.$$

Далее потребуется

**Теорема 1** [2]. *Если $z_* \in Z$ — точка локального минимума функции $f$ на множестве $Z$ в метрике $\rho$, то найдётся $\lambda^* < \infty$ такое, что при $\lambda > \lambda^*$ точка $z_*$ является точкой локального минимума функционала $\Phi_\lambda(z)$ на всём пространстве $\mathbb{R}^{2n}$ в той же метрике $\rho$.*

**Замечание 1.** Для доказательства теоремы 1 надо показать, что существуют $a > 0$ и $\delta > 0$ такие, что

$$\varphi^{\downarrow}(z) = \liminf_{\substack{y \in \mathbb{R}^{2n},\\ y \to z}} \frac{\varphi(y) - \varphi(z)}{\rho(y, z)} \leqslant -a < 0, \quad z \in Z_\delta \backslash Z, \tag{2}$$

где $Z_\delta = \left\{z \in \mathbb{R}^{2n} \;\middle|\; \varphi(z) < \delta\right\}$.

## 3. Классическая вариация функции $\varphi$ и её свойства

Изучим подробнее свойства функции

$$\varphi(z) = |h_1(x_1)| + |h_2(x_2)|,$$

где $h_k(x_k) = \frac{1}{2}x_k^T A_k x_k + x_k^T B_k + C_k$, $k = 1, 2$.

Пусть $z \in \mathbb{R}^{2n}$ фиксировано, $\varepsilon > 0$. Возьмём произвольное $v \in \mathbb{R}^{2n}$ и положим

$$z_\varepsilon = z + \varepsilon v = \begin{pmatrix} x_1 + \varepsilon v_1 \\ x_2 + \varepsilon v_2 \end{pmatrix}. \tag{3}$$

Имеем

$$h_k(x_k + \varepsilon v_k) = h_k(x_k) + \varepsilon v_k^T \left(A_k x_k + B_k\right) + \varepsilon^2 \frac{1}{2} v_k^T A_k v_k, \quad k = 1, 2. \tag{4}$$

Если $z \in Z$, то $h_1(x_1) = h_2(x_2) = 0$ и из (4) получаем явный вид производной функции $\varphi(z)$ в точке $z$ по направлению $v$:

$$\varphi'(z, v) := \lim_{\varepsilon \downarrow 0} \frac{\varphi(z + \varepsilon v) - \varphi(z)}{\varepsilon} = \sum_{k=1}^{2} \left| v_k^T \left(A_k x_k + B_k\right) \right|.$$

Несложно проверить, что это выражение можно переписать в виде

$$\varphi'(z, v) = \max_{\substack{w_k \in [-1,1], \\ k=1,2}} \left[ w_1 v_1^T \left(A_1 x_1 + B_1\right) + w_2 v_2^T \left(A_2 x_2 + B_2\right) \right]. \tag{5}$$

Из (5) заключаем, что функция $\varphi(z)$ субдифференцируема в точке $z$, причём её субдифференциал $\partial\varphi(z)$ равен

$$\partial\varphi(z) = \mathrm{co} \left\{ \begin{pmatrix} w_1 \left(A_1 x_1 + B_1\right) \\ w_2 \left(A_2 x_2 + B_2\right) \end{pmatrix} \mid w_k \in [-1, 1],\ k = 1, 2 \right\},$$

т. е. $\varphi'(z, v) = \max\limits_{W \in \partial\varphi(z)} v^T W$.

Пусть $z \notin Z$, т. е. $\varphi(z) > 0$. Рассмотрим производную функции $\varphi(z)$ в точке $z$ по направлению $v$ и покажем справедливость соотношения (2). Возможны следующие три случая.

Случай 1. Если $h_1(x_1) \neq 0$, $h_2(x_2) \neq 0$, то

$$\varphi'(z, v) = w_1 v_1^T \left(A_1 x_1 + B_1\right) + w_2 v_2^T \left(A_2 x_2 + B_2\right), \tag{6}$$

где $w_1 = \operatorname{sign} h_1(x_1)$, $w_2 = \operatorname{sign} h_2(x_2)$. Из (6) заключаем, что функция $\varphi$ дифференцируема в точке $z$, причём её градиент $\frac{\partial\varphi(z)}{\partial z}$ равен

$$\frac{\partial\varphi(z)}{\partial z} = \begin{pmatrix} \left(A_1 x_1 + B_1\right) \operatorname{sign} h_1(x_1) \\ \left(A_2 x_2 + B_2\right) \operatorname{sign} h_2(x_2) \end{pmatrix}.$$

Случай 2. Если $h_1(x_1) = 0$, $h_2(x_2) \neq 0$, то

$$\varphi'(z, v) = \max_{w_1 \in [-1,1]} \left[ w_1 v_1^T \left(A_1 x_1 + B_1\right) + w_2 v_2^T \left(A_2 x_2 + B_2\right) \right], \tag{7}$$

где либо $w_2 = 1$, либо $w_2 = -1$ в зависимости от знака $h_2(x_2)$. Из (7) заключаем, что функция $\varphi$ субдифференцируема в точке $z$, причём её субдифференциал $\partial\varphi(z)$ равен

$$\partial\varphi(z) = \mathrm{co}\left\{ \begin{pmatrix} -(A_1x_1 + B_1) \\ (A_2x_2 + B_2)\,\mathrm{sign}\,h_2(x_2) \end{pmatrix}, \begin{pmatrix} (A_1x_1 + B_1) \\ (A_2x_2 + B_2)\,\mathrm{sign}\,h_2(x_2) \end{pmatrix} \right\}.$$

Случай 3. Если $h_1(x_1) \neq 0$, $h_2(x_2) = 0$, то

$$\varphi'(z, v) = \max_{w_2 \in [-1,1]} \left[ w_1 v_1^T (A_1x_1 + B_1) + w_2 v_2^T (A_2x_2 + B_2) \right], \tag{8}$$

где либо $w_1 = 1$, либо $w_1 = -1$ в зависимости от знака $h_1(x_1)$. Из (8) заключаем, что функция $\varphi$ субдифференцируема в точке $z$, причём её субдифференциал $\partial\varphi(z)$ равен

$$\partial\varphi(z) = \mathrm{co}\left\{ \begin{pmatrix} (A_1x_1 + B_1)\,\mathrm{sign}\,h_1(x_1) \\ -(A_2x_2 + B_2) \end{pmatrix}, \begin{pmatrix} (A_1x_1 + B_1)\,\mathrm{sign}\,h_1(x_1) \\ (A_2x_2 + B_2) \end{pmatrix} \right\}.$$

Покажем справедливость соотношения (2) в каждом случае.

В случае 1 для любого $\alpha > 0$ положим $z_\alpha = z + \alpha v^*$, где

$$v^* = -\begin{pmatrix} (A_1x_1 + B_1)\,\mathrm{sign}\,h_1(x_1) \\ (A_2x_2 + B_2)\,\mathrm{sign}\,h_2(x_2) \end{pmatrix}.$$

Тогда, используя (4) и (6), имеем

$$\begin{aligned} \varphi(z_\alpha) = \varphi(z) + \alpha v^{*T} \begin{pmatrix} (A_1x_1 + B_1)\,\mathrm{sign}\,h_1(x_1) \\ (A_2x_2 + B_2)\,\mathrm{sign}\,h_2(x_2) \end{pmatrix} + o_1(\alpha) \\ = \varphi(z) - \alpha H_1^*(z) + o_1(\alpha), \end{aligned} \tag{9}$$

где $o_1(\alpha)/\alpha \xrightarrow[\alpha\downarrow 0]{} 0$, $H_1^*(z) = \left\| \begin{matrix} A_1x_1 + B_1 \\ A_2x_2 + B_2 \end{matrix} \right\|^2$. Заметим, что найдётся $\delta_1 > 0$ такое, что $H_1^*(z) > 0$ для всех $z \in Z_{\delta_1} \backslash Z$.

Далее,

$$\rho(z_\alpha, z) = \alpha \left\| \begin{pmatrix} A_1x_1 + B_1 \\ A_2x_2 + B_2 \end{pmatrix} \right\|. \tag{10}$$

Из (2) имеем

$$\varphi^{\downarrow}(z) \leqslant \liminf_{\alpha\downarrow 0} \frac{\varphi(z_\alpha) - \varphi(z)}{\rho(z_\alpha, z)}. \tag{11}$$

Подставляя (9) и (10) в (11), получаем

$$\varphi^{\downarrow}(z) \leqslant -\left\| \begin{pmatrix} A_1x_1 + B_1 \\ A_2x_2 + B_2 \end{pmatrix} \right\| < 0. \tag{12}$$

В случае 2 для любого $\alpha > 0$ положим $z_\alpha = z + \alpha v^*$, где

$$v^* = -\begin{pmatrix} \mathbb{O}_n \\ (A_2x_2 + B_2)\operatorname{sign} h_2(x_2) \end{pmatrix}.$$

Тогда, используя (4) и (7), имеем

$$\varphi(z_\alpha) = \varphi(z) - \alpha\|A_2x_2 + B_2\|^2 + o_2(\alpha) = \varphi(z) - \alpha H_2^*(z) + o_2(\alpha), \tag{13}$$

где $o_2(\alpha)/\alpha \xrightarrow[\alpha\downarrow 0]{} 0$. Заметим, что найдётся $\delta_2 > 0$ такое, что $H_2^*(z) > 0$ для всех $z \in Z_{\delta_2} \backslash Z$.

Далее,

$$\rho(z_\alpha, z) = \alpha \left\| A_2x_2 + B_2 \right\|. \tag{14}$$

Подставляя (13), (14) в (11), получаем

$$\varphi^{\downarrow}(z) \leqslant -\left\| A_2x_2 + B_2 \right\| < 0. \tag{15}$$

Наконец, рассмотрим случай 3. Как и в случае 2, для любого $\alpha > 0$ положим $z_\alpha = z + \alpha v^*$, где

$$v^* = -\begin{pmatrix} (A_1x_1 + B_1)\operatorname{sign} h_1(x_1) \\ \mathbb{O}_n \end{pmatrix}.$$

Тогда, используя (4) и (8), имеем

$$\varphi(z_\alpha) = \varphi(z) - \alpha\|A_1x_1 + B_1\|^2 + o_3(\alpha) = \varphi(z) - \alpha H_3^*(z) + o_3(\alpha), \tag{16}$$

где $o_3(\alpha)/\alpha \xrightarrow[\alpha\downarrow 0]{} 0$. Заметим, что найдётся $\delta_3 > 0$ такое, что $H_3^*(z) > 0$ для всех $z \in Z_{\delta_3} \backslash Z$.

Далее,

$$\rho(z_\alpha, z) = \alpha\|A_1x_1 + B_1\|. \tag{17}$$

Подставляя (16) и (17) в (11), получаем

$$\varphi^{\downarrow}(z) \leqslant -\left\| A_1x_1 + B_1 \right\| < 0. \tag{18}$$

Выберем $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Тогда (2) и справедливость теоремы 1 вытекают из (12), (15) и (18).

Таким образом, задача минимизации функционала $f$ на множестве $Z$ сведена к задаче минимизации функционала $\Phi_\lambda(z)$ на всём пространстве $\mathbb{R}^{2n}$ при $\lambda > \lambda^*$.

### 4. Классическая вариация функции $f(z)$

Для классической вариации (3) имеем

$$f(z_\varepsilon) = f(z) + \varepsilon v^T \begin{pmatrix} x_1 - x_2 \\ x_2 - x_1 \end{pmatrix} + \frac{1}{2}\varepsilon^2 \left\| v_1 - v_2 \right\|^2. \tag{19}$$

Отсюда следует, что функция $f$ дифференцируема в точке $z$, а

$$Q(z) = \begin{pmatrix} x_1 - x_2 \\ x_2 - x_1 \end{pmatrix} \tag{20}$$

является градиентом $f$ в $z$.

### 5. Необходимые условия минимума

Пусть $z_* \in Z$ — точка локального минимума функционала $f$ на множестве $Z$. Ранее установлено, что найдётся $\lambda^* < \infty$ такое, что при $\lambda > \lambda^*$ точка $z_*$ будет точкой локального минимума функционала $\Phi_\lambda(z) = f(z) + \lambda\varphi(z)$ на всём пространстве $\mathbb{R}^{2n}$. Зафиксируем произвольное $\lambda > \lambda^*$. Пусть $\varepsilon > 0$, выберем $v \in \mathbb{R}^{2n}$ и положим

$$z_\varepsilon = z_* + \varepsilon v. \tag{21}$$

Поскольку $\varphi(z_*) = 0$, для вариации (21) из (5) и (19) имеем

$$\begin{aligned} \Phi_\lambda(z_\varepsilon) = \Phi_\lambda(z_*) + \varepsilon \big[ & v^T Q(z_*) \\ & + \lambda \big\{ \left| v_1^T (A_1 x_1 + B_1) \right| + \left| v_2^T (A_2 x_2 + B_2) \right| \big\} \big] + o(\varepsilon), \end{aligned} \tag{22}$$

где $Q(z_*)$ — выражение (20) при $z = z_*$, $o(\varepsilon)/\varepsilon \underset{\varepsilon \downarrow 0}{\longrightarrow} 0$. Соотношение (22) можно переписать в виде

$$\Phi_\lambda(z_\varepsilon) = \Phi_\lambda(z_*) + \varepsilon \max_{\substack{w_k \in [-1,1], \\ k=1,2}} v^T \left[ Q(z_*) + \lambda \begin{pmatrix} w_1 (A_1 x_1 + B_1) \\ w_2 (A_2 x_2 + B_2) \end{pmatrix} \right] + o(\varepsilon). \tag{23}$$

**Теорема 2** [2]. *Для того чтобы $z_* \in Z$ была точкой глобального или локального минимума функции $f$ на множестве $Z$ в метрике $\rho$, необходимо, чтобы $\Phi_\lambda^\downarrow(z_*) \geqslant 0$.*

Учитывая произвольность $v \in \mathbb{R}^{2n}$, из (23) и теоремы 2 имеем необходимое условие минимума

$$\max_{\substack{w_k \in [-1,1], \\ k=1,2}} v^T \left[ Q(z_*) + \lambda \begin{pmatrix} w_1 (A_1 x_1 + B_1) \\ w_2 (A_2 x_2 + B_2) \end{pmatrix} \right] \geqslant 0, \quad v \in \mathbb{R}^{2n}. \tag{24}$$

**Теорема 3.** *Соотношение* (24) *эквивалентно следующему условию: существуют* $w_1^*$, $w_2^* \in [-1,1]$ *такие, что*

$$\begin{pmatrix} x_1^* - x_2^* \\ x_2^* - x_1^* \end{pmatrix} + \lambda \begin{pmatrix} w_1^*\left(A_1x_1 + B_1\right) \\ w_2^*\left(A_2x_2 + B_2\right) \end{pmatrix} = \mathbb{O}_{2n}.$$

Доказательство аналогично доказательству леммы 4.6.1 в [2] или теоремы 3 в [9]

### 6. Направление наискорейшего спуска функции $\Phi_\lambda$

Предположим, что точка $z = (x_1, x_2) \in \mathbb{R}^{2n}$ такая, что $x_1 \neq x_2$ и $A_kx_k + B_k \neq 0$ при $k = 1, 2$. Ранее установлено, что функция $\varphi(z)$ дифференцируема по направлениям. Поскольку функция $f(z)$ дифференцируема, штрафная функция $\Phi_\lambda(z) = f(z) + \lambda\varphi(z)$ тоже дифференцируема по направлениям в точке $z$.

Возможны следующие 4 случая: $h_1(x_1) \neq 0$, $h_2(x_2) \neq 0$; $h_1(x_1) = 0$, $h_2(x_2) \neq 0$; $h_1(x_1) \neq 0$, $h_2(x_2) = 0$; $h_1(x_1) = 0$, $h_2(x_2) = 0$.

Случай 1: $h_1(x_1) \neq 0$, $h_2(x_2) \neq 0$. Из (6) и (19) заключаем, что $\Phi_\lambda$ дифференцируема в точке $z$, причём её градиент $G_\lambda(z) := \frac{\partial\Phi_\lambda(z)}{\partial z}$ равен

$$G_\lambda(z) = \begin{pmatrix} x_1 - x_2 \\ x_2 - x_1 \end{pmatrix} + \lambda \begin{pmatrix} (A_1x_1 + B_1)\operatorname{sign} h_1(x_1) \\ (A_2x_2 + B_2)\operatorname{sign} h_2(x_2) \end{pmatrix}. \qquad (25)$$

Направление $g_\lambda(z) = -G_\lambda(z)/\|G_\lambda(z)\|$ является направлением наискорейшего спуска функционала $\Phi_\lambda$ в точке $z$.

Случай 2: $h_1(x_1) = 0$, $h_2(x_2) \neq 0$. Из (7) и (19) заключаем, что функция $\Phi_\lambda$ субдифференцируема в точке $z$, причём её субдифференциал $\partial\Phi_\lambda(z)$ равен

$$\begin{aligned}\partial\Phi_\lambda(z) = &\begin{pmatrix} x_1 - x_2 \\ x_2 - x_1 \end{pmatrix} \\ &+ \lambda \operatorname{co}\left\{\begin{pmatrix} -\left(A_1x_1 + B_1\right) \\ (A_2x_2 + B_2)\operatorname{sign} h_2(x_2) \end{pmatrix}, \begin{pmatrix} \left(A_1x_1 + B_1\right) \\ (A_2x_2 + B_2)\operatorname{sign} h_2(x_2) \end{pmatrix}\right\}. \qquad (26)\end{aligned}$$

Каждый элемент $W \in \partial\Phi_\lambda(z)$ можно описать следующим образом:

$$W(\mu) = \begin{pmatrix} x_1 - x_2 + \mu\lambda\left(A_1x_1 + B_1\right) \\ x_2 - x_1 + \lambda\left(A_2x_2 + B_2\right)\operatorname{sign} h_2(x_2) \end{pmatrix},$$

где $\mu \in [-1,1]$. Найдём минимальный по норме субградиент $W \in \partial\Phi_\lambda(z)$, т. е. решим задачу

$$\min_{W\in\partial\Phi_\lambda(z)} \|W\|^2 = \min_{\mu\in[-1,1]} [\|x_1 - x_2 + \mu\lambda\,(A_1x_1 + B_1)\,\|^2 \\ + \|x_2 - x_1 + \lambda\,(A_2x_2 + B_2)\,\mathrm{sign}\,h_2(x_2)\|^2] = \|W^*\|^2.$$

Но прежде вычислим

$$\min_{\mu\in\mathbb{R}} [\|x_1 - x_2 + \mu\lambda\,(A_1x_1 + B_1)\,\|^2 \\ + \|x_2 - x_1 + \lambda\,(A_2x_2 + B_2)\,\mathrm{sign}\,h_2(x_2)\|^2] \\ = \min_{\mu\in\mathbb{R}} [\mu^2\lambda^2\|A_1x_1 + B_1\|^2 + 2\mu\lambda(x_1 - x_2)^T\,(A_1x_1 + B_1) \\ + \|x_1 - x_2\|^2 + \|x_2 - x_1 + \lambda\,(A_2x_2 + B_2)\,\mathrm{sign}\,h_2(x_2)\|^2].$$

При сделанных предположениях коэффициент при $\mu^2$ положителен, поэтому минимум достигается в единственной точке

$$\mu^* = -\frac{(x_1 - x_2)^T\,(A_1x_1 + B_1)}{\lambda\|A_1x_1 + B_1\|^2}.$$

Напомним, что мы можем выбрать любое $\lambda > \lambda^*$ (см. теорему 1), поэтому $\mu^* \in [-1,1]$ при достаточно больших $\lambda$.

Вектор $W^* = \begin{pmatrix} x_1 - x_2 + \mu^*\lambda\,(A_1x_1 + B_1) \\ x_2 - x_1 + \lambda\,(A_2x_2 + B_2)\,\mathrm{sign}\,h_2(x_2) \end{pmatrix}$ является наименьшим субградиентом функции $\Phi_\lambda$ в точке $z$. Направление наискорейшего спуска функционала $\Phi_\lambda$ в точке $z$ равно $g_\lambda(z) = -W^*/\|W^*\|$.

Случай 3: $h_1(x_1) \neq 0$, $h_2(x_2) = 0$. Из (8) и (19) заключаем, что функция $\Phi_\lambda$ субдифференцируема в точке $z$, причём её субдифференциал $\partial\Phi_\lambda(z)$ равен

$$\partial\Phi_\lambda(z) = \begin{pmatrix} x_1 - x_2 \\ x_2 - x_1 \end{pmatrix} \\ + \lambda\,\mathrm{co}\left\{ \begin{pmatrix} (A_1x_1 + B_1)\,\mathrm{sign}\,h_1(x_1) \\ -\,(A_2x_2 + B_2) \end{pmatrix}, \begin{pmatrix} (A_1x_1 + B_1)\,\mathrm{sign}\,h_1(x_1) \\ (A_2x_2 + B_2) \end{pmatrix} \right\}. \quad (27)$$

Каждый элемент $W \in \partial\Phi_\lambda(z)$ можно описать следующим образом:

$$W(\mu) = \begin{pmatrix} x_1 - x_2 + \lambda\,(A_1x_1 + B_1)\,\mathrm{sign}\,h_1(x_1) \\ x_2 - x_1 + \mu\lambda\,(A_2x_2 + B_2) \end{pmatrix},$$

где $\mu \in [-1,1]$. Найдём минимальный по норме субградиент $W \in \partial\Phi_\lambda(z)$, т. е. решим задачу

$$\min_{W\in\partial\Phi_\lambda(z)} \|W\|^2 = \min_{\mu\in[-1,1]}[\|x_1 - x_2 + \lambda\,(A_1x_1 + B_1)\operatorname{sign} h_1(x_1)\|^2 \\ + \|x_2 - x_1 + \mu\lambda\,(A_2x_2 + B_2)\,\|^2] = \|W^*\|^2.$$

Вначале найдём

$$\min_{\mu\in\mathbb{R}}[\|x_1 - x_2 + \lambda\,(A_1x_1 + B_1)\operatorname{sign} h_1(x_1)\|^2 \\ + \|x_2 - x_1 + \mu\lambda\,(A_2x_2 + B_2)\,\|^2] \\ = \min_{\mu\in\mathbb{R}}[\mu^2\lambda^2\|A_2x_2 + B_2\|^2 + 2\mu\lambda(x_2 - x_1)^T\,(A_2x_2 + B_2) \\ + \|x_2 - x_1\|^2 + \|x_1 - x_2 + \lambda\,(A_1x_1 + B_1)\operatorname{sign} h_1(x_1)\|^2].$$

При сделанных предположениях коэффициент при $\mu^2$ положителен, поэтому минимум достигается в единственной точке

$$\mu^* = -\frac{(x_2 - x_1)^T\,(A_2x_2 + B_2)}{\lambda\|A_2x_2 + B_2\|^2}.$$

При достаточно больших $\lambda$ будет $\mu^* \in [-1,1]$.

Вектор $W^* = \begin{pmatrix} x_1 - x_2 + \lambda\,(A_1x_1 + B_1)\operatorname{sign} h_1(x_1) \\ x_2 - x_1 + \mu^*\lambda\,(A_2x_2 + B_2) \end{pmatrix}$ является наименьшим субградиентом функции $\Phi_\lambda$ в точке $z$. Направление наискорейшего спуска функционала $\Phi_\lambda$ в точке $z$ равно $g_\lambda(z) = -W^*/\|W^*\|$.

Случай 4: $h_1(x_1) = 0$, $h_2(x_2) = 0$. Из (5) и (19) заключаем, что функция $\Phi_\lambda$ субдифференцируема в точке $z$, причём её субдифференциал $\partial\Phi_\lambda(z)$ равен

$$\partial\Phi_\lambda(z) = \begin{pmatrix} x_1 - x_2 \\ x_2 - x_1 \end{pmatrix} \\ + \lambda\operatorname{co}\left\{\begin{pmatrix} w_1\,(A_1x_1 + B_1) \\ w_2\,(A_2x_2 + B_2) \end{pmatrix} \;\middle|\; w_k \in [-1,1],\ k = 1,2\right\}. \tag{28}$$

Любой элемент $W \in \partial\Phi_\lambda(z)$ можно представить в следующем виде:

$$W(\mu_1,\mu_2) = \begin{pmatrix} x_1 - x_2 + \mu_1\,(A_1x_1 + B_1) \\ x_2 - x_1 + \mu_2\,(A_2x_2 + B_2) \end{pmatrix},$$

где $\mu_k \in [-\lambda, \lambda]$, $k = 1, 2$. Найдём минимальный по норме субградиент $W \in \partial\Phi_\lambda(z)$, т. е. решим задачу

$$\min_{W\in\partial\Phi_\lambda(z)} \|W\|^2 = \min_{\substack{\mu_k\in[-\lambda,\lambda],\\ k=1,2}} [\|x_1 - x_2 + \mu_1 (A_1x_1 + B_1)\|^2 + \|x_2 - x_1 + \mu_2 (A_2x_2 + B_2)\|^2] = \|W^*\|^2.$$

Вначале найдём

$$\begin{aligned} &\min_{\substack{\mu_k\in\mathbb{R},\\ k=1,2}} [\|x_1 - x_2 + \mu_1 (A_1x_1 + B_1)\|^2 + \|x_2 - x_1 + \mu_2 (A_2x_2 + B_2)\|^2] \\ &= \min_{\substack{\mu_k\in\mathbb{R},\\ k=1,2}} \big[\mu_1^2\|A_1x_1 + B_1\|^2 + \mu_2^2\|A_2x_2 + B_2\|^2 + 2\mu_1(x_1 - x_2)^T (A_1x_1 + B_1) \\ &\qquad + 2\mu_2(x_2 - x_1)^T (A_2x_2 + B_2) + 2\|x_1 - x_2\|^2\big]. \end{aligned}$$

Приравнивая нулю производные по $\mu_1$ и $\mu_2$, получим систему

$$\begin{pmatrix} \|A_1x_1 + B_1\|^2 & 0 \\ 0 & \|A_2x_2 + B_2\|^2 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} (x_2 - x_1)^T (A_1x_1 + B_1) \\ (x_1 - x_2)^T (A_2x_2 + B_2) \end{pmatrix}.$$

При сделанных предположениях определитель этой системы положителен, поэтому она имеет единственное решение

$$\begin{pmatrix} \mu_1^* \\ \mu_2^* \end{pmatrix} = \begin{pmatrix} \dfrac{(x_2 - x_1)^T (A_1x_1 + B_1)}{\|A_1x_1 + B_1\|^2} \\ \dfrac{(x_1 - x_2)^T (A_2x_2 + B_2)}{\|A_2x_2 + B_2\|^2} \end{pmatrix}.$$

Вектор $W^* = \begin{pmatrix} x_1 - x_2 + \mu_1^* (A_1x_1 + B_1) \\ x_2 - x_1 + \mu_2^* (A_2x_2 + B_2) \end{pmatrix}$ является наименьшим субградиентом функции $\Phi_\lambda$ в точке $z$ и не зависит от $\lambda$.

Если $\|W^*\| > 0$, то направление $g(z) = -W^*/\|W^*\|$ является направлением наискорейшего спуска функционала $\Phi_\lambda$ в точке $z$.

Таким образом, используя классическую вариацию (3), имеем

$$\Phi_\lambda(z_\varepsilon) = \Phi_\lambda(z) + \varepsilon H_\lambda(z, v) + o(\varepsilon), \tag{29}$$

где $H_\lambda(z, v) = \max\limits_{W\in\partial\Phi_\lambda(z)} v^T W$.

## 7. Метод наискорейшего спуска

Пусть $z_* \in Z$ — точка минимума функционала $\Phi_\lambda(z)$ на $\mathbb{R}^{2n}$. Тогда $\varphi(z_*) = 0$. Как следует из теоремы 3, необходимое условие (24) эквивалентно существованию $w_k^* \in [-1, 1]$, $k = 1, 2$, для которых

$$\begin{pmatrix} x_1^* - x_2^* \\ x_2^* - x_1^* \end{pmatrix} + \lambda \begin{pmatrix} w_1^* \, (A_1 x_1 + B_1) \\ w_2^* \, (A_2 x_2 + B_2) \end{pmatrix} = \mathbb{O}_{2n}.$$

Это условие эквивалентно включению

$$\mathbb{O}_{2n} \in \partial \Phi_\lambda(z_*). \tag{30}$$

Точка $z_*$, в которой выполнено условие (30), называется *стационарной.*

Итак, если точка $z \in \mathbb{R}^{2n}$ не является стационарной точкой функционала $\Phi_\lambda$, то можно найти направление наискорейшего спуска функционала $\Phi_\lambda$ в точке $z$. Опишем метод наискорейшего спуска для нахождения стационарных точек.

Выберем произвольное $z^0 \in \mathbb{R}^{2n}$. Пусть уже найдено $z^k \in \mathbb{R}^{2n}$. Если $\varphi(z^k) = 0$ и выполнено условие (30), то точка $z^k$ является стационарной, и процесс прекращается. Если $\varphi(z^k) \neq 0$ или $\varphi(z^k) = 0$, но условие (30) не выполнено, то возьмём функцию $G_{k\lambda} = G_\lambda(z^k)$ — наименьший по норме субградиент функционала $\Phi_\lambda$ в точке $z^k$.

Далее решается задача одномерной минимизации

$$\min_{\beta \geqslant 0} \Phi_\lambda(z^k - \beta G_{k\lambda}) = \Phi_\lambda(z^k - \beta_k G_{k\lambda}). \tag{31}$$

Положим $z^{k+1} = z^k - \beta_k G_{k\lambda}$. Имеем $\Phi_\lambda(z^{k+1}) < \Phi_\lambda(z^k)$.

К сожалению, описанный процесс может и не привести к стационарной точке, поскольку субдифференциальное отображение $\partial \Phi_\lambda(z)$ не является непрерывным как функция $z$ в метрике Хаусдорфа [2, 4].

## 8. Метод гиподифференциального спуска

Вместо разложения (29) для вариации (3) можно получить другое представление. Имеем

$$\Phi_\lambda(z_\varepsilon) = \Phi_\lambda(z) + H_\lambda(\varepsilon, z, v) + o(\varepsilon), \tag{32}$$

где

$$\begin{aligned}H_\lambda(\varepsilon, z, v) = \varepsilon v^T \begin{pmatrix} x_1 - x_2 \\ x_2 - x_1 \end{pmatrix} \\ + \lambda\Big[ \max \big\{ h_1(x_1) - |h_1(x_1)| + \varepsilon v_1^T\left(A_1x_1 + B_1\right), \\ - h_1(x_1) - |h_1(x_1)| - \varepsilon v_1^T\left(A_1x_1 + B_1\right)\big\} \\ + \max \big\{ h_2(x_2) - |h_2(x_2)| + \varepsilon v_2^T\left(A_2x_2 + B_2\right), \\ - h_2(x_2) - |h_2(x_2)| - \varepsilon v_2^T\left(A_2x_2 + B_2\right)\big\}\Big].\end{aligned}$$

Из (32) следует, что функционал $\Phi_\lambda(z)$ гиподифференцируем в точке $z$ и его гиподифференциал равен

$$\begin{aligned}d\Phi_\lambda(z) = \left\{ \begin{pmatrix} 0 \\ Q(z) \end{pmatrix} \right\} \\ + \lambda\Bigg( \mathrm{co} \left\{ \begin{pmatrix} h_1(x_1) - |h_1(x_1)| \\ h_1'(x_1) \end{pmatrix}, \begin{pmatrix} -h_1(x_1) - |h_1(x_1)| \\ -\, h_1'(x_1) \end{pmatrix} \right\} \\ + \mathrm{co} \left\{ \begin{pmatrix} h_2(x_2) - |h_2(x_2)| \\ h_2'(x_2) \end{pmatrix}, \begin{pmatrix} -h_2(x_2) - |h_2(x_2)| \\ -\, h_2'(x_2) \end{pmatrix} \right\} \Bigg),\end{aligned}$$

где

$$Q(z) = \begin{pmatrix} x_1 - x_2 \\ -(x_1 - x_2) \end{pmatrix}, \quad h_1'(x_1) = \begin{pmatrix} A_1x_1 + B_1 \\ \mathbb{O}_n \end{pmatrix},$$

$$h_2'(x_2) = \begin{pmatrix} \mathbb{O}_n \\ A_2x_2 + B_2 \end{pmatrix}, \quad \mathbb{O}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Отметим, что отображение $d\Phi_\lambda(z)$ непрерывно в метрике Хаусдорфа [2, 4]. Нетрудно также показать, что необходимое условие минимума функции $\Phi_\lambda(z)$ эквивалентно условию

$$[0, \mathbb{O}_{2n}] \in d\Phi_\lambda(z_*). \tag{33}$$

Перед тем как приступить к поиску минимального по норме гипоградиента $W \in d\Phi_\lambda(z)$, заметим, что гиподифференциал $d\Phi_\lambda(z)$ можно представить в виде $d\Phi_\lambda(z) = \{W(\mu_1, \mu_2) \mid |\mu_1| \leqslant \lambda,\ |\mu_2| \leqslant \lambda\}$, где

$$W(\mu_1, \mu_2) = \begin{pmatrix} \mu_1 h_1(x_1) + \mu_2 h_2(x_2) - \lambda\varphi(z) \\ \mu_1 h_1'(x_1) + \mu_2 h_2'(x_2) + Q(z) \end{pmatrix}.$$

Найдём

$$\min_{W\in d\Phi_\lambda(z)} \|W\|^2 = \min_{\substack{\mu_k\in[-\lambda,\lambda],\\ k=1,2}} \|W(\mu_1,\mu_2)\|^2 = \|W^*\|^2.$$

Однако вначале решим задачу

$$\min_{\substack{\mu_k\in\mathbb{R},\\ k=1,2}} \|W(\mu_1,\mu_2)\|^2.$$

Имеем

$$\begin{aligned} L(\mu_1,\mu_2) &:= \|W(\mu_1,\mu_2)\|^2 \\ &= \left[\mu_1 h_1(x_1) + \mu_2 h_2(x_2) - \lambda\varphi(z)\right]^2 + \|\mu_1 h_1'(x_1) + \mu_2 h_2'(x_2) + Q(z)\|^2. \end{aligned}$$

Приравняем к нулю производные функции $L(\mu_1,\mu_2)$ по $\mu_1$ и $\mu_2$. Получим систему

$$D\mu = \eta, \tag{34}$$

где

$$D = \begin{pmatrix} h_1^2(x_1) + \left(h_1'(x_1)\right)^T h_1'(x_1) & h_1(x_1)h_2(x_2) \\ h_1(x_1)h_2(x_2) & h_2^2(x_2) + \left(h_2'(x_2)\right)^T h_2'(x_2) \end{pmatrix},$$

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \eta = \begin{pmatrix} \lambda\varphi(z)h_1(x_1) - \left(h_1'(x_1)\right)^T Q(z) \\ \lambda\varphi(z)h_2(x_1) - \left(h_2'(x_2)\right)^T Q(z) \end{pmatrix}.$$

Вновь предположим, что точка $z = (x_1, x_2) \in \mathbb{R}^{2n}$ такая, что $x_1 \neq x_2$ и $A_k x_k + B_k \neq 0$ при $k = 1, 2$. Тогда $\det D > 0$, и система (34) имеет единственное решение $\mu^* = D^{-1}\eta$. При достаточно больших $\lambda$ будет $|\mu_k^*| \leqslant \lambda$, $k = 1, 2$. Следовательно, при таких $\lambda$

$$W^* = \begin{pmatrix} \mu_1^* h_1(x_1) + \mu_2^* h_2(x_2) - \lambda\varphi(z) \\ \mu_1^* h_1'(x_1) + \mu_2^* h_2'(x_2) + Q(z) \end{pmatrix}.$$

Вектор

$$G^*(z) = \mu_1^* h_1'(x_1) + \mu_2^* h_2'(x_2) + Q(z) \tag{35}$$

является гипоградиентом функционала $\Phi_\lambda(z)$. Если $\|G^*\| > 0$, т. е. точка $z$ не стационарна, то направление $g(z) = -\frac{G^*(z)}{\|G^*(z)\|}$ является направлением спуска функционала $\Phi_\lambda$ в точке $z$. В отличие от направления наискорейшего спуска, направление $g(z)$ как функция $z$ непрерывно.

**Замечание 2.** Если $z \in Z$, то наименьший субградиент (29) совпадает с гипоградиентом (35) функционала $\Phi_\lambda$ в точке $z$.

Таким образом, если точка $z \in \mathbb{R}^{2n}$ не является стационарной точкой функционала $\Phi_\lambda$, то можно найти направление спуска функционала $\Phi_\lambda$ в точке $z$.

Опишем следующий метод гиподифференциального спуска для нахождения стационарных точек, т.е. точек, удовлетворяющих (30) или (33). Выберем произвольное $z_0 \in \mathbb{R}^{2n}$. Пусть уже найдено $z_k \in \mathbb{R}^{2n}$. Если $\varphi(z_k) = 0$ и выполнено (33), то точка $z_k$ стационарна, и процесс прекращается. Если условие (33) не выполнено, то возьмём вектор $G^*(z_k)$ — гипоградиент функции $\Phi_\lambda$ в точке $z_k$. Теперь найдём

$$\min_{\beta \geqslant 0} \Phi_\lambda(z_k - \beta G^*(z_k)) = \Phi_\lambda(z_k - \beta_k G^*(z_k)) \tag{36}$$

и положим $z_{k+1} = z_k - \beta_k G^*(z_k)$. Имеем $\Phi_\lambda(z_{k+1}) < \Phi_\lambda(z_k)$. Пользуясь непрерывностью в метрике Хаусдорфа гиподифференциального отображения как функции $z$, можно доказать, что описанный метод сходится в следующем смысле: $\|G^*(z_k)\| \longrightarrow 0$.

**Замечание 3.** Отметим, что величина шага спуска $\beta$ в задачах (31) и (36) вычисляется аналитически.

**Замечание 4.** В случае, когда один эллипсоид содержит другой, метод «шаров» неприменим, в отличие от предлагаемых нами методов.

**Замечание 5.** Если один или оба эллипсоида вырождаются в точку, то метод «шаров» неприменим (происходит зацикливание), в то время как метод гиподифференциального спуска работает и выдаёт корректные результаты.

**Замечание 6.** Если один из эллипсоидов пустой, то метод гиподифференциального спуска «информирует» об этом.

## Заключение

В работе описаны два алгоритма решения поставленной проблемы. Разработано программное обеспечение в системах MATLAB и Maple, реализующее данные алгоритмы. Из проведённых экспериментов следует, что метод гиподифференциального спуска более универсален, чем метод «шаров» [12], и менее трудоёмок («эффективнее») нежели методы из [8, 11].

## ЛИТЕРАТУРА


1. **Демьянов В. Ф.** Точные штрафные функции в задачах негладкой оптимизации // Вестн. СПбГУ. Сер. 1. — 1994. — Вып. 4, № 22. — С. 21–27.
2. **Демьянов В. Ф.** Условия экстремума и вариационное исчисление. — М.: Высшая школа, 2005. — 335 с.
3. **Демьянов В. Ф., Васильев Л. В.** Недифференцируемая оптимизация. — М.: Наука, 1981. — 384 с.
4. **Демьянов В. Ф., Рубинов А. М.** Основы негладкого анализа и квазидифференциальное исчисление. — М.: Наука, 1990. — 432 с.
5. **Ерёмин И. И.** Метод «штрафов» в выпуклом программировании // Докл. АН СССР. — 1967. — Т. 143, № 4. — С. 748–751.
6. **Калинина Е. А., Утешев А. Ю.** Теория исключения: учебное пособие. — СПб.: НИИ Химии СПбГУ, 2002. — 72 с.
7. **Косолап А. И.** Квадратичные оптимизационные задачи компьютерной геометрии // Искусственный интеллект. — 2010. — № 1. — С. 70–75.
8. **Лебедев Д. М., Полякова Л. Н.** Задача проектирования нулевой точки на квадрик // Вестн. СПбГУ. Сер. 10. — 2013. — Вып. 1. — С. 11–17.
9. **Тамасян Г. Ш.** О методах наискорейшего и гиподифференциального спуска в одной задаче вариационного исчисления // Вычислительные методы и программирование: новые вычислительные технологии. — 2012. — Т. 13. — С. 197–217.
10. **Утешев А. Ю.** Вычисление расстояний между геометрическими объектами // http://pmpu.ru/vf4/algebra2/optimiz/distance.
11. **Утешев А. Ю., Яшина М. В.** Нахождение расстояния от эллипсоида до плоскости и квадрики в $\mathbb{R}^n$ // Докл. АН. — 2008. — Т. 419, № 4. — С. 471–474.
12. **Lin A., Han S. P.** On the distance between two ellipsoids // SIAM J. Optim. — 2002. — Vol. 13. — P. 298–308.


*Тамасян Григорий Шаликович,*
e-mail: grigoriytamasjan@mail.ru
*Чумаков Андрей Александрович,*
e-mail: katandrew@mail.ru

Статья поступила
2 сентября 2013 г.

Переработанный вариант —
11 ноября 2013 г.