О ЧУВСТВИТЕЛЬНОСТИ ЕВКЛИДОВОЙ ПРОЕКЦИИ

ЖВМ и МФ, 2014, том 54, № 3, с. 392–403

А. Ф. Измаилов, А. С. Куренной

Докладчик: Т. Бардадым

Киев, 06.06.17

OUTLINE

- Постановка задачи
- Основные результаты

Изучается структура и поведение евклидовых проекций точки на множество, задаваемое параметрическими ограничениями.

Параметрическое семейство множеств:

$$D(\sigma) = \{x \in R^n \mid h(\sigma, x) = 0, \quad g(\sigma, x) \le 0\} \qquad (1.1)$$

$$h : R^s \times R^n \to R^l, \quad g : R^s \times R^n \to R^m$$

$\sigma \in R^s$ – параметр

Точечно-множественное отображение проектирования
$\Pi\ :\ R^s \times R^n \to 2^{R^n}$:

$$\Pi(\sigma, y) = \{\pi \in D(\sigma) \mid ||\pi - y|| = \mathsf{dist}(y, D(\sigma))\} \quad (1.2)$$

$$\mathsf{dist}(y, D) = \inf_{d \in D} ||y - d||$$

При заданном базовом значении параметра $\overline{\sigma} \in R^s$ рассмотрим точку $\overline{x} \in D(\overline{\sigma})$. Разумеется, $\Pi(\overline{\sigma}, \overline{x}) = \{\overline{x}\}$.
В работе изучается поведение Π вблизи $(\overline{\sigma}, \overline{x})$.

В многочисленных приложениях возникают вопросы об условиях, обеспечивающих однозначность этого отображения, его непрерывность, липшицевость, гладкость. Задача точного или приближенного поиска евклидовой проекции используется в качестве подзадачи в различных оптимизационных алгоритмах, например в методах (неточного) восстановления допустимости, в методах решения задач допустимости и т.д. Кроме того, задача поиска решения системы уравнений и неравенств, наименее уклоняющегося от заданной точки, имеет очевидное самостоятельное значение.

Множество $\Pi(\sigma, y)$ совпадает с множеством глобальных решений задачи математического программирования

$$\frac{1}{2}||x - y||^2 \to \min, \; h(\sigma, x) = 0, \; g(\sigma, x) \leq 0 \qquad (1.3)$$

относительно неизвестной $x \in R^n$, где $\sigma \in R^s$ и $y \in R^n$ являются параметрами.

Без дополнительных предположений регулярности имеют место следующие очевидные свойства.

Если отображения $h(\overline{\sigma}, \cdot)$ и $g(\overline{\sigma}, \cdot)$ непрерывны вблизи $\overline{x}$, то значения отображения $\Pi(\overline{\sigma}, \cdot)$ на некоторой окрестности $\overline{x}$ непусты и компактны, причем

$$\max_{x \in \Pi(\overline{\sigma}, y)} ||x - \overline{x}|| \to 0 \ \text{ при } \ y \to \overline{x}.$$

Вместе с тем, в любых предположениях непрерывности и даже гладкости, без соответствующих условий регулярности ограничений множество $D(\sigma)$ (а значит и $\Pi(\overline{\sigma}, y)$ для любого $y \in R^n$) может быть пусто при $\sigma \in R^s$, сколь угодно близких к $\overline{\sigma}$, причем даже если это не так, величина $\text{dist}(\overline{x}, D(\sigma))$ может быть отделена от нуля при $\sigma \in R^s$, сколь угодно близких к $\overline{\sigma}$.

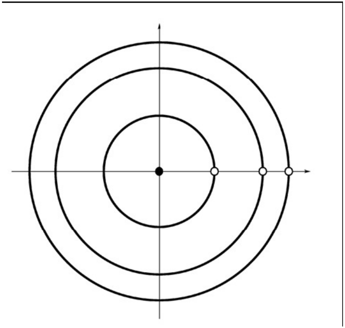
Однако если отображения $h(\overline{\sigma}, \cdot)$ и $g(\overline{\sigma}, \cdot)$ непрерывны на некоторой окрестности $\overline{x}$ для всех $\sigma \in R^s$, достаточно близких к $\overline{\sigma}$, и

$$\mathrm{dist}(\overline{x}, D(\sigma)) \to 0 \ \text{ при } \ \sigma \to \overline{\sigma}. \qquad (1.4)$$

$$\max_{x \in \Pi(\overline{\sigma}, y)} ||x - \overline{x}|| \to 0 \ \text{ при } \ \sigma \to \overline{\sigma}, \quad y \to \overline{x} \qquad (1.5)$$

Если окрестность непрерывности отображений $h(\overline{\sigma}, \cdot)$ и $g(\overline{\sigma}, \cdot)$ зависит от σ, то ни замкнутость, ни непустоту значений Π вблизи $(\overline{\sigma}, \overline{x})$ гарантировать нельзя даже при выполнении условия (1.4).

Пример 1



$$s = 1, \quad n = 2, \quad l = 1, \quad m = 0,$$

$$h(\sigma, x) = \begin{cases} ||x||^2 - \sigma^2, & x \neq (\sigma, 0). \\ \sigma, & x = (\sigma, 0). \end{cases}$$

Пример 1

Тогда для $\overline{\sigma} = 0$ и $\overline{x} = 0$ выполняется включение $\overline{x} \in D(\overline{\sigma})$, причем для всякого $\sigma \neq 0$ отображение $h(\sigma, \cdot)$ непрерывно на шаре $\{x \in R^2 \mid ||x|| < \sigma\}$, а множество $D(\sigma)$ есть окружность радиуса σ с центром в нуле, с выколотой точкой $(\sigma, 0)$ и эта окружность стягивается к $\overline{x}$ при $\sigma \to 0$. При этом множество $\Pi(\sigma, \overline{x})$ совпадает с незамкнутым множеством $D(\sigma)$, а множество $\Pi(\sigma, (\sigma/2, 0))$ пусто.

Если отображения h и g предполагать непрерывными по совокупности переменных на всем $R^s \times R^n$, то значения Π всегда компактны. Однако если предполагать непрерывность этих отображений только в окрестности $(\overline{\sigma}, \overline{x})$, то в случае нарушения (1.4) опять же нельзя гарантировать замкнутость значений Π вблизи $(\overline{\sigma}, \overline{x})$.

Ограничения-равенства

В случае отсутствия ограничений–неравенств, т.е. для

$$D(\sigma) = \{x \in R^n \mid h(\sigma, x) = 0\} \qquad (2.1)$$

полный ответ на вопросы о поведении проекций может быть получен из стандартных теорем о неявной функции.

Для $\sigma \in R^s$ и $y \in R^n$ будем рассматривать задачу оптимизации

$$\frac{1}{2}||x-y||^2 \to \min, \quad h(\sigma, x) = 0 \qquad (2.2)$$

$$L(\sigma, y, x, \lambda) = \frac{1}{2}||x-y||^2 + \langle \lambda, h(\sigma, x) \rangle$$

$$x - y + \left(\frac{\partial h}{\partial x}(\sigma, x) \right)^T \lambda = 0, \quad h(\sigma, x) = 0 \qquad (2.3)$$

Очевидно, что единственное решение $\overline{x}$ задачи (2.2) при $(\sigma, y) = (\overline{\sigma}, \overline{x})$ является стационарной точкой этой задачи, причем отвечающие этой точке множители Лагранжа характеризуются равенством

$$\left(\frac{\partial h}{\partial x}(\overline{\sigma}, \overline{x})\right)^T \lambda = 0.$$

В частности, если выполнено условие регулярности ограничений

$$\text{rank}\frac{\partial h}{\partial x}(\overline{\sigma}, \overline{x}) = l, \qquad (2.4)$$

то стационарной точке отвечает единственный множитель Лагранжа $\overline{\lambda} = 0$. При этом

$$\frac{\partial^2 L}{\partial x^2}(\sigma, y, \overline{x}, 0) = I$$

откуда следует, что в точке $\overline{x}$ автоматически выполняется достаточное условие второго порядка оптимальности.

Теорема 1

Theorem

Пусть отображение $h : R^s \times R^n \to R^l$ *дважды непрерывно дифференцируемо в некоторой окрестности точки* $(\overline{\sigma}, \overline{x}) \in R^s \times R^n$, *причем* $h(\overline{\sigma}, \overline{x})) = 0$. *Пусть выполнено условие регулярности ограничений (2.4):*

$$\textit{rank}\frac{\partial h}{\partial x}(\overline{\sigma}, \overline{x}) = l, \qquad (2.4)$$

Тогда для любого $\sigma \in R^s$, *достаточно близкого к* $\overline{\sigma}$, *и любого* $y \in R^n$, *достаточно близкого к* $\overline{x}$, *имеет место следующее:*
а) система Лагранжа (2.3) имеет вблизи $(\overline{x}, 0)$ *единственное решение* $(x, \lambda) = (\pi(\sigma, y), \lambda(\sigma, y))$, *причем отображение* $(\pi(\sigma, \cdot), \lambda(\sigma, \cdot))$ *непрерывно дифференцируемо вблизи* $(\overline{\sigma}, \overline{x})$;
б) множество $\Pi(\sigma, y)$, *вводимое согласно (1.2), (2.1), состоит из единственного элемента* $\pi(\sigma, y)$.

Общий случай

Для функции Лагранжа $L : R^s \times R^n \times R^n \times R^l \times R^m \to R$ задачи (1.3), задаваемой как

$$L(\sigma, y, x, \lambda, \mu) = \frac{1}{2}||x - y||^2 + \langle \lambda, h(\sigma, x) \rangle + \langle \mu, g(\sigma, x) \rangle$$

стационарные точки и отвечающие им множители Лагранжа задачи (1.3) характеризуются системой Каруша–Куна–Таккера

$$x - y + \left(\frac{\partial h}{\partial x}(\sigma, x) \right)^T \lambda + \left(\frac{\partial g}{\partial x}(\sigma, x) \right)^T \mu = 0 \qquad (3.1)$$

$$\mu \geq 0, \quad g(\sigma, x) \leq 0, \quad \langle \mu, g(\sigma, x) = 0 \rangle \qquad (3.2)$$

Очевидно, что единственное решение $\overline{x}$ задачи (2.2) при $(\sigma, y) = (\overline{\sigma}, \overline{x})$ является стационарной точкой этой задачи, причем отвечающие этой точке множители Лагранжа характеризуются соотношениями

$$\left(\frac{\partial h}{\partial x}(\overline{\sigma}, \overline{x})\right)^T \lambda + \left(\frac{\partial g_A}{\partial x}(\overline{\sigma}, \overline{x})\right)^T \mu_A = 0, \;\; \mu_A \geq 0, \;\; \mu_N = 0. \quad (3.3)$$

где

$$A = A(\overline{\sigma}, \overline{x}) = \{i = 1, \ldots, m \mid g_i(\overline{\sigma}, \overline{x}) = 0\},$$

$$N = N(\overline{\sigma}, \overline{x}) = \{i = 1, \ldots, m\} \backslash A$$

В частности, если выполнено условие Мангасаряна–Фромовица, эквивалентная двойственная форма которого состоит в том, что система (3.3) имеет только тривиальное решение, то стационарной точке отвечает единственный множитель Лагранжа $(\overline{\lambda}, \overline{\mu}) = (0, 0)$. При этом

$$\frac{\partial^2 L}{\partial x^2}(\sigma, y, \overline{x}, 0, 0) = I, \qquad (3.4)$$

откуда следует, что в точке $\overline{x}$ автоматически выполняется любое разумное достаточное условие второго порядка оптимальности, например так называемое сильное достаточное условие второго порядка...

Теорема 2

Theorem

Пусть отображения $h\ :\ R^s \times R^n \to R^l$ *и* $g\ :\ R^s \times R^n \to R^m$ *дважды непрерывно дифференцируемы в некоторой окрестности точки* $(\overline{\sigma}, \overline{x})$*, причем* $h(\overline{\sigma}, \overline{x}) = 0$*,* $g(\overline{\sigma}, \overline{x}) \leq 0$*.*
Пусть выполнено условие Мангасаряна–Фромовица, т.е. система (3.3) имеет только тривиальное решение.
Тогда для любого $\sigma \in R^s$*, достаточно близкого к* $\overline{\sigma}$*, и любого* $y \in R^n$*, достаточно близкого к* $\overline{x}$*, имеет место следующее:*
а) задача (2.2) имеет вблизи $\overline{x}$ *единственную стационарную точку* $x = \pi(\sigma, y)$*, причем отображение* $\pi(\cdot)$ *непрерывно вблизи* $(\overline{\sigma}, \overline{x})$*;*
б) множество $\Pi(\sigma, y)$*, вводимое согласно (1.1), (1.2), состоит из единственного элемента* $\pi(\sigma, y)$*.*

При наличии ограничений–неравенств рассчитывать на гладкость проекции (как это было в теореме 1) не приходится. Тем не менее, при более сильных условиях регулярности можно усилить утверждение теоремы 2. В частности, липшицевость проекции можно получить при выполнении так называемого условия линейной независимости

$$\operatorname{rank}\begin{pmatrix} \frac{\partial h}{\partial x}(\overline{\sigma}, \overline{x}) \\ \frac{\partial g_A}{\partial x}(\overline{\sigma}, \overline{x}) \end{pmatrix} = l + |A| \qquad (3.5)$$

где через $|J|$ обозначается количество элементов в конечном множестве J. Это условие сильнее условия Мангасаряна–Фромовица, и, в частности, при выполнении (3.5) стационарной точке $\overline{x}$ по-прежнему отвечает единственный множитель Лагранжа $(\overline{\lambda}, \overline{\mu}) = (0, 0)$, причем по-прежнему выполняется сильное достаточное условие второго порядка.

Теорема 3

Theorem

Пусть отображения $h : R^s \times R^n \to R^l$ *и* $g : R^s \times R^n \to R^m$ *дважды непрерывно дифференцируемы в некоторой окрестности точки* $(\overline{\sigma}, \overline{x}) \in R^s \times R^n$, *причем* $h(\overline{\sigma}, \overline{x}) = 0$, $g(\overline{\sigma}, \overline{x}) \leq 0$.
Пусть выполнено условие линейной независимости (3.5).
Тогда для любого $\sigma \in R^s$, *достаточно близкого к* $\overline{\sigma}$, *и любого* $y \in R^n$, *достаточно близкого к* $\overline{x}$, *имеет место следующее:*
а) система Каруша–Куна–Таккера (3.1), (3.2) имеет вблизи $(\overline{x}, 0, 0)$ *единственное решение* $(x, \lambda, \mu) = (\pi(\sigma, y), \lambda(\sigma, y), \mu(\sigma, y))$, *причем отображение* $(\pi(\cdot), \lambda(\cdot), \mu(\cdot))$ *удовлетворяет условию Липшица вблизи* $(\overline{\sigma}, \overline{x})$;
б) множество $\Pi(\sigma, y)$, *вводимое согласно (1.2), (2.1), состоит из единственного элемента* $\pi(\sigma, y)$.

ВОЗМУЩЕНИЯ ПО НАПРАВЛЕНИЯМ

СЛУЧАЙ ПОСТОЯННОГО ДОПУСТИМОГО МНОЖЕСТВА

Пусть теперь

$$D(\sigma) = \{x \in R^n \mid h(x) = 0, \quad g(x) \le 0\} \qquad (5.1)$$

где $h \ : \ R^n \to R^l$ и $\ : \ R^n \to R^m$ – заданные отображения. Определим отображение проектирования $\Pi \ : \ R^n \to 2^{R^n}$

$$\Pi(y) = \{\pi \in D \mid ||\pi - y|| = \mathsf{dist}(y, D)\} \qquad (5.2)$$

Напомним, что если отображение h аффинно, а компоненты g выпуклы (что влечет замкнутость и выпуклость D), то, как известно, отображение Π глобально однозначно, удовлетворяет условию Липшица и, более того, является неразжимающим.

Для того чтобы охарактеризовать поведение отображения Π в невыпуклом случае, выпишем для всякого $y \in R^n$ аналог задачи (1.3), определяющий множество $\Pi(y)$:

$$\frac{1}{2}||x||^2 - \langle y, x\rangle \to \min, \quad h(x) = 0, \quad g(x) \le 0. \quad (5.3)$$

Параметр y входит в эту задачу весьма специальным образом: только в целевую функцию и только как линейное возмущение. Такая параметризация известна под названием параметризации наклона (от англ. “tilt parameterization”) и обладает определенными специальными свойствами с точки зрения теории чувствительности, отсутствующими в случае более общих параметризаций.

Теорема 5

Theorem

Пусть отображения $h : R^s \to R^l$ *и* $g : R^n \to R^m$ *дважды непрерывно дифференцируемы в некоторой окрестности точки* $\overline{x} \in R^n$, *причем* $h(\overline{x}) = 0$, $g(\overline{x}) \le 0$. *Пусть выполнено условие Мангасаряна–Фромовица, т.е. система*

$$(h'(\overline{x}))^T \eta + (g'_A(\overline{x}))^T \zeta_A = 0, \quad \zeta_A \ge 0, \quad \zeta_N = 0, \quad (5.4)$$

имеет только тривиальное решение, где $A = A(\overline{x}) = $
$= \{i = 1, \ldots, m \mid g_i(\overline{x}) = 0\}, \quad N = N(\overline{x}) = \{1, \ldots, m\} \backslash A.$
Тогда для любого $y \in R^s$, *достаточно близкого к* $\overline{x}$:
а) задача (5.3) имеет вблизи $\overline{x}$ *единственную стационарную точку* $x = \pi(y)$, *причем отображение* $\pi(\cdot)$ *удовлетворяет условию Липшица вблизи* $\overline{x}$;
б) множество $\Pi(\sigma, y)$, *вводимое согласно (5.1), (5.2), состоит из единственного элемента* $\pi(y)$.

12. Lu S. Implications of the constant rank constraint qualification // Math. Program. 2011. V. 126. P. 365–392.

13. Janin R. Directional derivative of the marginal function in nonlinear programming // Math. Program. Stud. 1984. V. 21. P. 110–126.

В [12] вопрос о поведении проекций изучался при выполнении так называемого условия постоянного ранга, состоящего в том, что для любых подмножеств $\mathcal{J} \subset \{1, \ldots, l\}$ и $\mathcal{I} \subset A(\overline{x})$ величина

$$\mathrm{rank} \begin{pmatrix} h'_{\mathcal{J}}(x) \\ g'_{\mathcal{I}}(x) \end{pmatrix}$$

постоянна для всех $x \in R^n$, достаточно близких к $\overline{x}$. Это условие регулярности отлично от условия Мангасаряна–Фромовица и в некотором смысле дополняет его. Заметим, что это условие автоматически выполняется в случае линейных ограничений.

Если допускать возмущения отображений h и g, т.е. рассматривать множества вида (1.1):

$$D(\sigma) = \{x \in R^n \mid h(\sigma, x) = 0, \quad g(\sigma, x) \leq 0\}, \qquad (1.1)$$

то, в отличие от условия Мангасаряна–Фромовица, условие постоянного ранга не гарантирует выполнение (1.4)

$$\text{dist}(\overline{x}, D(\sigma)) \to 0 \ \text{ при } \ \sigma \to \overline{\sigma}. \qquad (1.4)$$

В частности, $D(\sigma)$ может быть пустым при $\sigma \in R^s$, сколь угодно близких к $(\overline{\sigma}, \overline{x})$, даже если предполагать, что для любых подмножеств $\mathcal{J} \subset \{1, \ldots, l\}$ и $\mathcal{I} \subset A(\overline{x})$ величина

$$\text{rank} \begin{pmatrix} h'_{\mathcal{J}}(\sigma, x) \\ g'_{\mathcal{I}}(\sigma, x) \end{pmatrix}$$

постоянна для всех $(\sigma, x) \in R^s \times R^n$, достаточно близких к $(\overline{\sigma}, \overline{x})$.

Тем не менее, как показано в [12], если определить множество

$$\Sigma = \{\sigma \in R^s \mid D(\sigma) \neq \emptyset\}$$

то при этом вводимое согласно

$$\Pi(\sigma, y) = \{\pi \in D(\sigma) \mid ||\pi - y|| = \mathsf{dist}(y, D(\sigma))\} \quad (1.2)$$

отображение $\Pi \ : \ \Sigma \times R^n \to 2^{R^n}$ вблизи $(\overline{\sigma}, \overline{x})$ имеет одноточечные значения $\{\pi(\sigma, y)\}$, причем отображение $\pi(\cdot)$ является кусочно-гладким: существуют окрестность $\mathcal{U}$ точки $(\overline{\sigma}, \overline{x})$ и конечный набор непрерывно дифференцируемых отображений $\pi_k \ : \ \mathcal{U} \to R^n, \quad k = 1, \ldots, K$ такие, что отображение π непрерывно на $\mathcal{U}$ и

$$\pi(\sigma, y) \in \{\pi_1, \ldots, \pi_K\} \quad \forall (\sigma, y) \in (\Sigma \times R^n) \cap \mathcal{U}.$$

Соответственно, в случае постоянного допустимого множества, задаваемого в (5.1)

$$D(\sigma) = \{x \in R^n \mid h(x) = 0, \quad g(x) \leq 0\}, \qquad (5.1)$$

из условия постоянного ранга вытекает, что отображение $\Pi \; : \; \Sigma \times R^n \to 2^{R^n}$, вводимое согласно (5.2)

$$\Pi(y) = \{\pi \in D \mid ||\pi - y|| = \mathsf{dist}(y, D)\} \qquad (5.2)$$

имеет вблизи $\overline{x}$ одноточечные значения $\pi(y)$, причем отображение $\pi(\cdot)$ является кусочно-гладким на некоторой окрестности $\overline{x}$, а значит, и липшицевым на некоторой окрестности $\overline{x}$.

Дякую за увагу!

tbardadym@gmail.com