Простое доказательство леммы Неймана о седловой точке

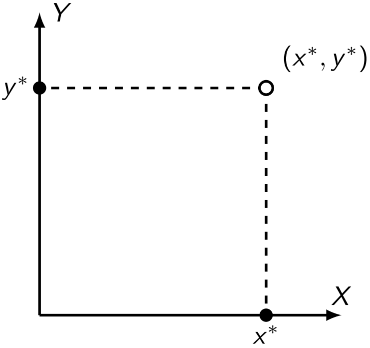
Шлезингер М.И., Крыгин В.М.
Отдел распознавания образов МНУЦ ИТиС

Семинар «Образный компьютер», 15 января 2019

Содержание

Двойственные задачи математического программирования

X, Y — множества, $f : X \times Y \to \mathbb{R}$ — функция.
$x^* = \arg\min_{x \in X} \max_{y \in Y} f(x, y)$, $y^* = \arg\max_{y \in Y} \min_{x \in X} f(x, y)$.



Теорема 1

$$\min_{x \in X} \max_{y \in Y} f(x, y) \geq \max_{y \in Y} \min_{x \in X} f(x, y)$$

Доказательство.

$$\min_{x \in X} \max_{y \in Y} f(x, y) = \max_{y \in Y} f(x^*, y) \geq f(x^*, y^*)$$

$$\max_{y \in Y} \min_{x \in X} f(x, y) = \min_{x \in X} f(x, y^*) \leq f(x^*, y^*)$$

□

Седловая точка

Определение 1

Точка $(x^*, y^*) \in X \times Y$, *такая что* $x^* = \arg\min_{x \in X} f(x, y^*)$ *и* $y^* = \arg\max_{y \in Y} f(x^*, y)$, *называется седловой точкой* f.

Теорема 2

$\min_{x \in X} \max_{y \in Y} f(x, y) = \max_{y \in Y} \min_{x \in X} f(x, y)$ *тогда и только тогда, когда для функции* f *существует седловая точка.*

Доказательство.

$$\max_{y \in Y} \min_{x \in X} f(x, y) \geq \min_{x \in X} f(x, y^*) = f(x^*, y^*) = \max_{y \in Y} f(x^*, y) \geq \min_{x \in X} \max_{y \in Y} f(x, y)$$

$\square$

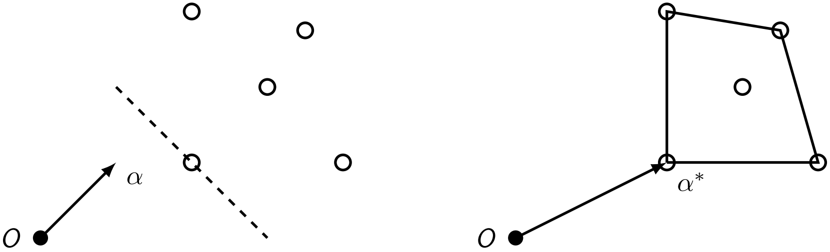
Лемма Неймана о седловой точке

Пусть $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$ — замкнутые ограниченные выпуклые подмножества, а $f : X \times Y \to \mathbb{R}$ функция, такая что

- $f(\cdot, y) : \mathbb{R}^m \to \mathbb{R}$ выпукла,
- $f(x, \cdot) : \mathbb{R}^n \to \mathbb{R}$ вогнута.

В таком случае для f существует седловая точка.

Обучение



$$\alpha^* = \arg\max_{\alpha} \min_{x \in X} \left\langle \frac{\alpha \cdot x}{\|\alpha\|} \right\rangle = \arg\min_{\alpha \in \tilde{X}} \|\alpha\|$$

Лемма Неймана в распознавании образов

- Задача Неймана-Пирсона
- Минимаксная задача
- Задача А. Вальда
- Распознавание сложных гипотез
- Дихотомия байесовских и непригодных стратегий обучения и распознавания

Субмодулярные задачи структурного распознавания

X — конечное множество.

$$\overline{x} = (x_1, x_2, \ldots, x_n)\,, x_i \in X, \overline{x} \in X^n.$$

$$\overline{x}^* = \arg\min_{\overline{x} \in X^n} \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}\left(x_i, x_j\right)$$

$$\min_{x_1,\ldots,x_n} \sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}\left(x_i, x_j\right) \mapsto \sum_{i=1}^{n} \sum_{j=1}^{n} \min_{x,y} f^*_{ij}\left(x, y\right)$$

Одностороннее неравенство Чебышева

Если $x \in \mathbb{R}$ — случайное число с математическим ожиданием 0 и дисперсией 1, то

$$P\left\{|x| \geq 3\right\} \leq \frac{1}{9}.$$

Что можно сказать о точной верхней грани вероятности $P\left\{x \geq 3\right\}$?

Известные доказательства леммы Неймана

- John von Neumann, Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 1928.
- Борис Френкин Теорема Неймана о минимаксе — общеизвестная и неизвестная. *Математическое просвещение*, 2005.
- Maurice Sion, On general minimax theorems. *Pacific Journal of Mathematics*, 1958.
- Jürgen Kindler, A Simple Proof of Sion's Minimax Theorem. *The American Mathematical Monthly*, 2005.
- Hichem Ben-El-Mechaiekh and Robert Dimand, A Simpler Proof of the Von Neumann Minimax Theorem. *American Mathematical Monthly*, 2011.
- Hidetoshi Komiya, Elementary proof for Sion's minimax theorem. *Kodai Mathematical Journal*, 1988.

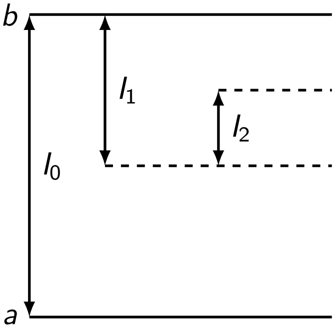
Топологическое пространство называется **компактным**, если любое его открытое покрытие содержит конечное подпокрытие.

Лемма Больцано (1817)

Любая бесконечная ограниченная последовательность чисел

$$\overline{x} = \left(x_1, \dots, x_i, x_{i+1}, \dots\right), x_i \in \mathbb{R}, a \le x_i \le b,$$

содержит в себе бесконечную сходящуюся подпоследовательность.



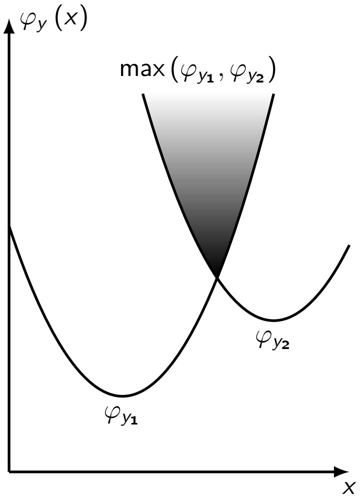
$$I_0 \supset I_1 \supset \dots \supset I_t \supset I_{t+1} \supset \dots$$

$$\left|I_{t+1}\right| = \frac{1}{2} \cdot \left|I_t\right|$$

Существует единственное

$$x^* \in \mathbb{R} : x^* \in \bigcap_{t=0}^{\infty} I_t.$$

Формальные свойства выпукых и вогнутых функций



$X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$, $f : X \times Y \to \mathbb{R}$.

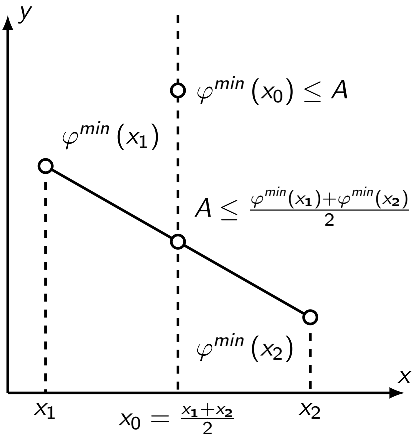
Лемма 1

Если X — выпуклое множество и $\varphi_{y_0} : x \mapsto f(x, y_0)$ выпуклая для любого $y_0 \in Y$, то $\varphi^{max} : x \mapsto \max\limits_{y \in Y} f(x, y)$ также выпуклая.

Лемма 2

Если Y — выпуклое множество и $\psi_{x_0} : y \mapsto f(x_0, y)$ вогнутая для любого $x_0 \in X$, то $\psi^{min} : y \mapsto \min\limits_{x \in X} f(x, y)$ также вогнутая.

Формальные свойства выпукых и вогнутых функций



$XY \subset \mathbb{R}^{m+n}$ — выпуклое множество

$$(x, y) \in \mathbb{R}^{m+n}, x \in \mathbb{R}^m, y \in \mathbb{R}^n,$$
$$X = \{x \in \mathbb{R}^m \mid \exists y \in \mathbb{R}^n : (x, y) \in XY\}.$$

Лемма 3

Если $f : XY \to \mathbb{R}$ — выпуклая (вогнутая) функция, то функции

$$\varphi^{\min} : x \mapsto \min_{y:\,(x,y)\in XY} f(x, y),$$

$$\varphi^{\max} : x \mapsto \max_{y:\,(x,y)\in XY} f(x, y)$$

выпуклые (вогнутые).

Шаг индукции для доказательства леммы Неймана

Лемма 4

Пусть m^* и n^* — два целых числа. Если лемма Неймана справедлива для всех $m \le m^*$, $n \le n^*$, то она справедлива и для всех $m \le m^* + 1$, $n \le n^* + 1$.

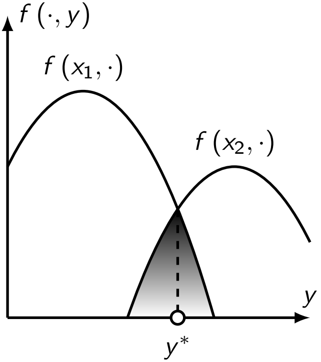
Доказательство.

$$\min_{x\in X}\ \min_{y\in Y(x)}\ \max_{z\in Z} f(x,y,z) = \min_{x\in X}\max_{z\in Z}\ \underbrace{\min_{y\in Y(x)}\ f(x,y,z)}_{\varphi(x,z)\ \ \smile} = \max_{z\in Z}\min_{x\in X}\ \min_{y\in Y(x)}\ f(x,y,z).$$

$$\min_{x\in X}\max_{y\in Y}\ \overbrace{\max_{z\in Z(y)}\ f(x,y,z)}^{\psi(x,y)\ \ \frown} = \max_{y\in Y}\min_{x\in X}\ \max_{z\in Z(y)}\ f(x,y,z) = \max_{y\in Y}\ \max_{z\in Z(y)}\ \min_{x\in X} f(x,y,z).$$

□

Лемма о двух точках



$$y^* = \arg\max_{y\in\mathbb{R}} \min_{x\in X} f(x, y)$$

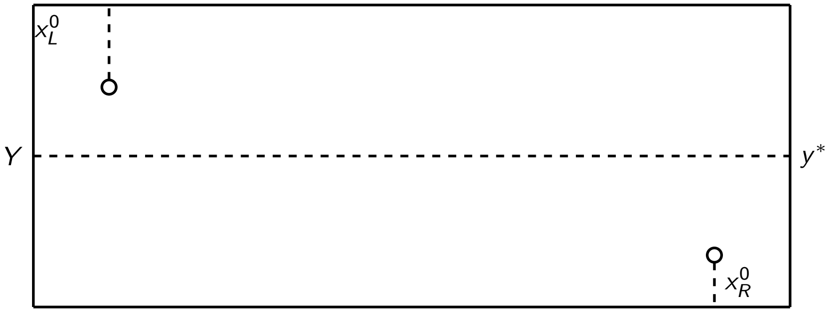
Если $f : X \times \mathbb{R} \to \mathbb{R}$ непрерывна от $x \in X$ и вогнута от $y \in \mathbb{R}$, и $y^* = \arg\max_{y\in\mathbb{R}} \min_{x\in X} f(x, y)$, то существуют $x_1, x_2 \in X$, такие что

$$\max_{y\in\mathbb{R}} \min \{f(x_1, y), f(x_2, y)\} = \\ = \max_{y\in\mathbb{R}} \min_{x\in X} f(x, y)$$

и $\arg\max_{y\in\mathbb{R}} f(x_1, y) \le y^* \le \arg\max_{y\in\mathbb{R}} f(x_2, y)$.

База индукции для доказательства леммы Неймана

$$y^* = \arg\max_{y \in Y} \min_{x \in X} f(x, y)$$



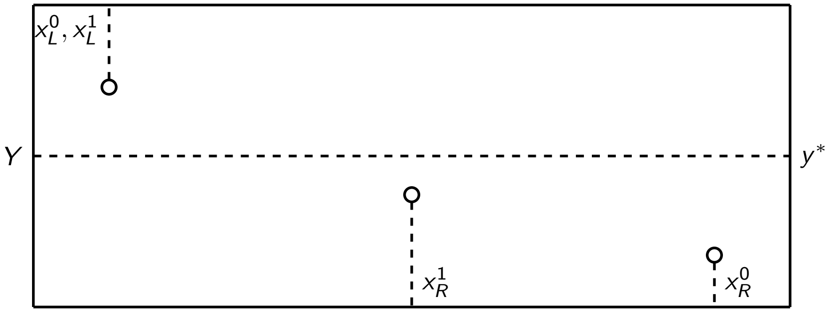
$$\left\{ x' \in X \,\middle|\, f(x', y^*) = \min_{x \in X} f(x, y^*) \right\}$$

$$x_L^0, \ldots, x_L^i, \ldots \qquad x_R^0, \ldots, x_R^i, \ldots$$

$$\lim_{i \to \infty} x_L^i = \lim_{i \to \infty} x_R^i = x^*$$

База индукции для доказательства леммы Неймана

$$y^* = \arg\max_{y \in Y} \min_{x \in X} f(x, y)$$



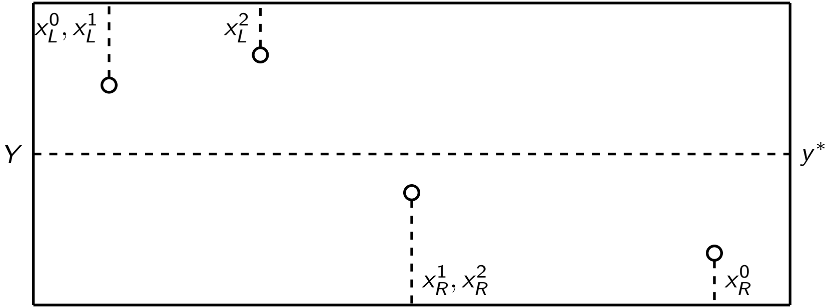
$$\left\{ x' \in X \,\middle|\, f(x', y^*) = \min_{x \in X} f(x, y^*) \right\}$$

$$x_L^0, \ldots, x_L^i, \ldots \qquad x_R^0, \ldots, x_R^i, \ldots$$

$$\lim_{i \to \infty} x_L^i = \lim_{i \to \infty} x_R^i = x^*$$

База индукции для доказательства леммы Неймана

$$y^* = \arg\max_{y \in Y} \min_{x \in X} f(x, y)$$



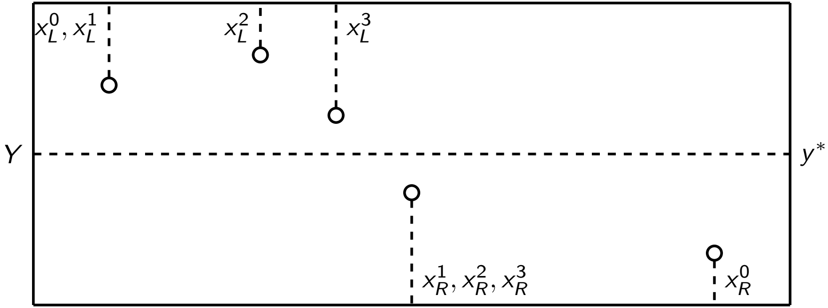
$$\left\{ x' \in X \,\middle|\, f(x', y^*) = \min_{x \in X} f(x, y^*) \right\}$$

$$x_L^0, \ldots, x_L^i, \ldots \qquad x_R^0, \ldots, x_R^i, \ldots$$

$$\lim_{i \to \infty} x_L^i = \lim_{i \to \infty} x_R^i = x^*$$

База индукции для доказательства леммы Неймана

$$y^* = \arg\max_{y \in Y} \min_{x \in X} f(x, y)$$



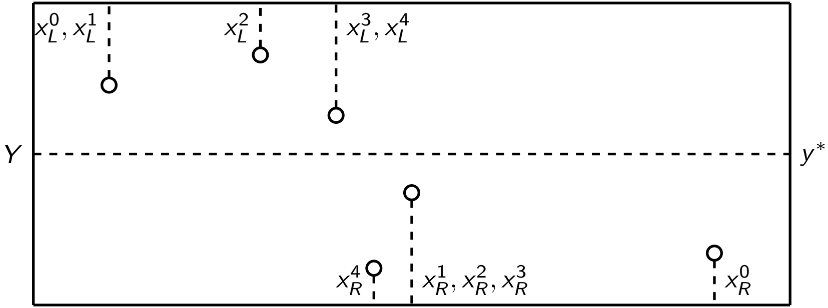
$$\left\{ x' \in X \,\middle|\, f(x', y^*) = \min_{x \in X} f(x, y^*) \right\}$$

$$x_L^0, \ldots, x_L^i, \ldots \qquad x_R^0, \ldots, x_R^i, \ldots$$

$$\lim_{i \to \infty} x_L^i = \lim_{i \to \infty} x_R^i = x^*$$

База индукции для доказательства леммы Неймана

$$y^* = \arg\max_{y \in Y} \min_{x \in X} f(x, y)$$



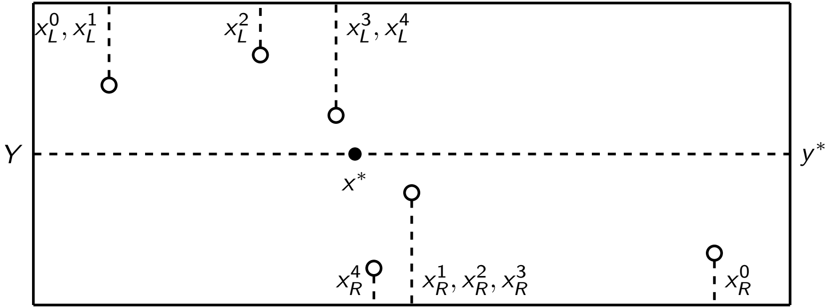
$$\left\{ x' \in X \,\middle|\, f(x', y^*) = \min_{x \in X} f(x, y^*) \right\}$$

$$x_L^0, \ldots, x_L^i, \ldots \qquad x_R^0, \ldots, x_R^i, \ldots$$

$$\lim_{i \to \infty} x_L^i = \lim_{i \to \infty} x_R^i = x^*$$

База индукции для доказательства леммы Неймана

$$y^* = \arg\max_{y \in Y} \min_{x \in X} f(x, y)$$

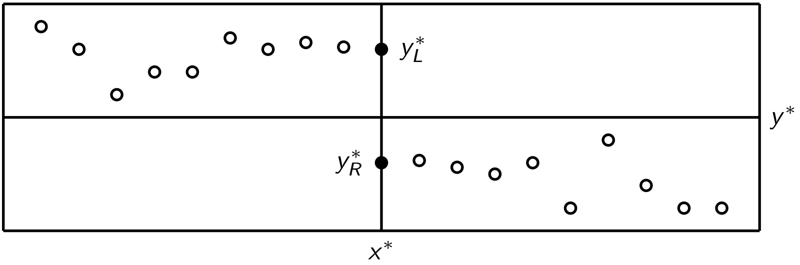


$$\left\{ x' \in X \,\middle|\, f(x', y^*) = \min_{x \in X} f(x, y^*) \right\}$$

$$x_L^0, \ldots, x_L^i, \ldots \qquad x_R^0, \ldots, x_R^i, \ldots$$

$$\lim_{i \to \infty} x_L^i = \lim_{i \to \infty} x_R^i = x^*$$

База индукции для доказательства леммы Неймана



$$\lim_{j\to\infty}\left(x_L^{i(j)}, y_L^{i(j)}\right) = (x^*, y_L^*) \qquad \lim_{j\to\infty}\left(x_R^{i(j)}, y_R^{i(j)}\right) = (x^*, y_R^*)$$

$$f(x^*, y_L^*) = \lim_{j\to\infty} f\left(x_L^{i(j)}, y_L^{i(j)}\right) = \lim_{j\to\infty} \max_{y\in Y} f\left(x_L^{i(j)}, y\right) = \max_{y\in Y} f(x^*, y)$$

$$f(x^*, y_R^*) = \lim_{j\to\infty} f\left(x_R^{i(j)}, y_R^{i(j)}\right) = \lim_{j\to\infty} \max_{y\in Y} f\left(x_R^{i(j)}, y\right) = \max_{y\in Y} f(x^*, y)$$

$$\min_{x\in X} f(x, y^*) = f(x^*, y^*) = \max_{y\in Y} f(x^*, y)$$