

HW1: выпуклые множества, матричное дифференцирование, выпуклые функции

Семинарист: Курузов Илья

Дедлайн: 23:59, 26.09.2020

1 Выпуклые множества

1. Пусть множество $S \subset \mathbb{R}^n$ выпукло, $\|\cdot\|$ есть норма в \mathbb{R}^n , $a > 0$ - некоторое положительное число.

а) [1] Определим расширение множества S , как

$$S_a = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \text{dist}(\mathbf{x}, S) \leq a \right\}.$$

Напомним, что расстояние от точки до множества определяется как $\text{dist}(\mathbf{x}, S) = \inf_{\mathbf{y}} \|\mathbf{x} - \mathbf{y}\|$. Докажите, что S_a является выпуклым.

б) [2] Определим сужение множества S , как

$$S_{-a} = \left\{ \mathbf{x} \in \mathbb{R}^n \mid B_{a, \|\cdot\|}(\mathbf{x}) \subseteq S \right\},$$

где $B_{a, \|\cdot\|}(\mathbf{x})$ - замкнутый шар с центром в \mathbf{x} и радиусом a по норме $\|\cdot\|$. Докажите, что S_{-a} выпукло.

2. [2] Функция $f: X \rightarrow \mathbb{R}$, где X есть некоторое линейное пространство, называется квази-выпуклой, если $\forall \alpha \in [0, 1], \forall \mathbf{x}_1, \mathbf{x}_2 \in X$ выполнено

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \leq \max(f(\mathbf{x}_1), f(\mathbf{x}_2)).$$

Определим множество подуровня функции f : $L_c = \left\{ \mathbf{x} \in X \mid f(\mathbf{x}) \leq c \right\}$. Докажите, что множество L_c является выпуклым для любого $c \in \mathbb{R}$ тогда и только тогда, когда f - квази-выпуклая функция.

3. [2] Пусть $K \subseteq \mathbb{R}_+^n$ есть конус. Докажите, что K - выпуклый конус тогда и только тогда, когда множество $\{\mathbf{x} \in K \mid \sum_{i=1}^n x_i = 1\}$ есть выпуклое множество.

2 Матричное дифференцирование

Постарайтесь представить ответы в следующих номерах в матричном виде.

4. [3] Пусть дана некоторая последовательность $\{\mathbf{x}_k\}_{k=1}^N \subset \mathbb{R}^n$ и некоторый постоянный вектор $\mathbf{y}_0 \in \mathbb{R}^m$. Рассмотрим следующую рекурсивную модель:

$$\mathbf{y}_j(W, V, \mathbf{b}) = \sigma(W\mathbf{x}_j + V\mathbf{y}_{j-1} + \mathbf{b}), j = \overline{1, N}$$

в которой $W \in \mathbb{R}^{m \times n}$, $V \in \mathbb{R}^{m \times m}$ и $\mathbf{b} \in \mathbb{R}^m$ есть параметры модели, $\sigma(x) = \frac{1}{1+e^{-x}}$ есть сигмоида (к вектору данная функция применяется поэлементно $[\sigma(\mathbf{x})]_i = \sigma(x_i)$). Рассмотрим функцию:

$$f(W, V, \mathbf{b}) = \|\mathbf{y}_N - \mathbf{Y}\|_2^2,$$

где $\mathbf{Y} \in \mathbb{R}^m$ есть некоторый константный вектор.

Найдите градиенты функции f по W, V, \mathbf{b} , как функцию, зависящую только от этих параметров, $\mathbf{x}_n, \mathbf{y}_n$ и градиента от \mathbf{y}_{n-1} по этим параметрам.

Для упрощения записи может помочь $\sigma'(x) = \sigma(x)(1 - \sigma(x))$. Для обозначения поэлементного матричного умножения (произведение Адамара) используйте символ \circ (`\circ`).

5. [2] Пусть $Y \in \mathbb{R}^{m \times n}$, $U \in \mathbb{R}^{m \times k}$, $V \in \mathbb{R}^{k \times n}$. Найдите градиенты по U и V следующей функции

$$J(U, V) = \|UV - Y\|_F^2 + \frac{\lambda}{2} (\|U\|_F^2 + \|V\|_F^2),$$

где $\|\cdot\|_F$ есть норма Фробениуса для матриц.

6. [2] Найдите градиент и гессиан по параметру $\mathbf{w} \in \mathbb{R}^n$ функции $f(\mathbf{w}) = \ln(1 + \langle \mathbf{x}, \mathbf{w} \rangle)$ для постоянного вектора \mathbf{x} .

3 Выпуклые функции

7. [3] Функция $f(x, t) = -\ln(t^2 - \mathbf{x}^\top \mathbf{x})$, определённая на множестве $E = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 < t\}$ называется логарифмическим барьером для конуса второго порядка. Она используется при переходе от условной оптимизации с ограничением на конус второго порядка к безусловной с помощью введения штрафа за выход из этого конуса. Поэтому такие функции называются барьерными. Докажите, что функция f выпукла.

8. [2] Докажите, что функция $f(\mathbf{x}) = \left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}}$ является вогнутой на \mathbb{R}_{++}^n для любого $p < 1, p \neq 0$.

9. [2] Докажите, что функция $f(X) = (\det X)^{\frac{1}{n}}$ является вогнутой на множестве \mathbb{S}_{++}^n .

10. Пусть величина x есть вещественная случайная величина, принимающая вещественные значения $a_1 \dots a_n \in \mathbb{R}, a_1 \leq \dots \leq a_n$ с вероятностями $p_1 \dots p_n$ соответственно. Пусть $P = \left\{ \mathbf{p} \in \mathbb{R}_+^n \mid \sum_k p_k = 1 \right\}$ есть n -мерный симплекс, т.е. все возможные вероятностные распределения для величины с n значениями. Являются следующие функции, как функции от \mathbf{p} , выпуклыми/вогнутыми?

(a) [1] $\mathbb{E}x$

(b) [1] $\text{var}(x^2)$