Рогачев Александр. Заочно
Домашнее задание №1.
Введение/Выпуклые множества/Выпуклые функции. Векторное дифференцирование.

ВВЕДЕНИЕ (3 балла)

1. Для успешного выступления на соревнованиях спортсмен должен потратить как минимум K килокалорий. Вечером перед соревнованиями он идёт в магазин, чтобы купить продукты себе на ужин. В магазине представлено m наименований товаров, цена единицы каждого наименования равна p_i, $i = 1, ..., m$. Также известно, что i-ая единица каждого товара придаёт студенту энергию равную k_i килокалорий. Поставьте задачу определения содержимого корзины минимальной стоимости для успешного выступления. Является ли поставленная задача выпуклой и почему?

   Решение:
   m - количество различных товаров;
   p_i - стоимость i-го товара;
   k_i - энергетическая ценность i-го товара;
   Пусть x_i - количество приобритаемого i-го товара, тогда задача сводится к минимизации величины:
   $$\sum_{i=1}^{m} p_i x_i \to min,$$
   при этом ограничения имеют вид:
   $$\sum_{i=1}^{m} k_i x_i \geq K.$$
   Так как $\forall p_i \in P : p_i > 0$ (будем считать, что бесплатных товаров, увы, нет), то целевая функция представляет из себя положительную взвешенную сумму выпуклых функций, т.е. является выпуклой.

2. Что такое норма вектора? Что такое эквивалентность норм? Показать эквивалентность l_1, l_2, l_∞ норм.

   Решение:

   (a) Поставим каждому элементу (в нашем случае - вектору $x \in V$) в соответсвие некоторое значение функции $p(x)$, такое что:
      i. $p(x) = 0 \Rightarrow x = 0_V$;
      ii. $\forall x, y \in V, p(x + y) \leqslant p(x) + p(y)$;

iii. $\forall \alpha \in \mathbb{C}, \forall x \in V, p(\alpha\, x) = |\alpha| p(x)$.

Функционал p, удовлетворяющий указанным выше условиям, называется нормой.

(b) Две нормы a и b на пространстве V называются эквивалентными, если $\exists C_1, C_2 \forall x \in V$:

$$C_1 a(x) \leqslant b(x) \leqslant C_2 a(x)$$

(c) В контексте геометрической интерпретации в двумерном случае, у норм будет вид квадрата("повернутого на $\pi/4$"), сферы и, опять же, квадрата, соответственно. Поэтому задача сводится к тому, чтобы вписать и описать сферу квадратами в случаях норм 1 и 2. Учитвая, что будем рассматривать сферы и кубы в n-мерном пространстве, то получим:

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n}\, \|x\|_2$$

Аналогично:

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\, \|x\|_\infty$$

Для случая из двух кубов получим:

$$\|x\|_\infty \leq \|x\|_1 \leq n\, \|x\|_\infty$$

3. Что такое норма матрицы (линейного конечномерного оператора)? Чему равны $\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_\infty$ нормы матрицы? Что такое Фробениусова норма матрицы и какой аналог среди векторных норм векторов она имеет?

Решение:

(a) Аналогично ситуации в предыдущем пункте, поставим в соответсвие матрице A вещественное число, такое что:

i. $\|\alpha A\| = |\alpha| \|A\|$

ii. $\|A + B\| \leq \|A\| + \|B\|$

iii. $\|A\| \geq 0$

iv. $\|A\| = 0 \Rightarrow A = 0$ (в смысле матрицы такой же размерности, как и A)

Функция $\|\cdot\|$, удовлетворяющая свойстам выше называется нормой матрицы.

(b) $\|A\|_1 = \max\limits_{1 \leq j \leq n} \sum\limits_{i=1}^{m} |a_{ij}|$

$\|A\|_2 = \sup\limits_{\|x\|_2 = 1} \|Ax\|_2 = \sup\limits_{(x,x)=1} \sqrt{(Ax, Ax)} = \sqrt{\lambda_{max}(A^* A)}$

$\|A\|_\infty = \max\limits_{1 \leq i \leq m} \sum\limits_{j=1}^{n} |a_{ij}|$

(c) Норма Фробениуса является расширением Евклидовой нормы.

$$\|A\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{\operatorname{trace}(A^* A)} = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2(A)},$$

ВЫПУКЛЫЕ/КОНИЧЕСКИЕ МНОЖЕСТВА (2 балла)

1. Доказать по определению, что гиперболическое множество

$$\left\{ x \in \mathbb{R}^n_+ \;\middle|\; \prod_{i=1}^{n} x_i \geq 1 \right\} \text{– выпуклое.}$$

   Решение:

$$\ln \prod_{i=1}^{n}(\alpha x_i + (1-\alpha)y_i) = \sum_{i=1}^{n}(\ln(\alpha x_i + (1-\alpha)y_i) \geq \sum_{i=1}^{n}\alpha \ln x_i + (1-\alpha)\sum_{i=1}^{n}\ln y_i =$$

$$= \ln((\prod_{i=1}^{n} x_i)^{\alpha} \cdot (\prod_{i=1}^{n} y_i)^{1-\alpha}) \geq ln(1)$$

   Т.е. выполняется характеристическое свойство множества

$$\prod_{i=1}^{n} z_i \geq 1,$$

$$z_i = \alpha x_i + (1-\alpha)y_i$$

   чтд.

2. Пусть $\mathrm{P} \in \mathbb{S}^n_{++}$, $c \in \mathbb{R}^n$. Покажите, что множество

$$\left\{ x \in \mathbb{R}^n \mid \langle \mathrm{P}x, x\rangle \leq \langle c, x\rangle^2,\ \langle c, x\rangle \geq 0 \right\} \text{– выпуклое.}$$

   Решение:

$$\langle P(\alpha x_1 + (1-\alpha)x_2), (\alpha x_1 + (1-\alpha)x_2)\rangle = \alpha^2\langle Px_1, x_1\rangle + (1-\alpha)^2\langle Px_2, x_2\rangle + 2\alpha(1-\alpha)\langle Px_1, x_2\rangle (1)$$

$$\langle (c, (\alpha x_1 + (1-\alpha)x_2)\rangle^2 = (\alpha\langle c, x_1\rangle + (1-\alpha)\langle c, x_2\rangle)^2 = \alpha^2\langle c, x_1\rangle^2 + (1-\alpha)^2\langle c, x_2\rangle^2 +$$

$$+2\alpha(1-\alpha)\langle c, x_1\rangle\langle c, x_2\rangle (2)$$

   Первые два элемента первой суммы соответственно меньше первых двух элементов из второй суммы исходя из условия. Осталось показать, что это верно и для последних элементов сумм.

$$\langle c, x_1\rangle \cdot \langle c, x_2\rangle \geq \sqrt{\langle Px_1, x_1\rangle} \cdot \sqrt{\langle Px_2, x_2\rangle} = \sqrt{\left\|\sqrt{P}x_1\right\|^2 \cdot \left\|\sqrt{P}x_2\right\|^2} =$$

$$= \left\|\sqrt{P}x_1\right\| \cdot \left\|\sqrt{P}x_2\right\| \geq \langle \sqrt{P}x_1, \sqrt{P}x_2\rangle = \langle Px_1, x_2\rangle$$

   Использовали условие, неравенство КБШ. Таким образом, каждый элемент суммы в (1) меньше соответствующего элемента в (2), чтд.

~~ВЫПУКЛЫЕ ФУНКЦИИ (2 балла)~~

1. Докажите выпуклость следующих функций:

   (a) $f(x) = \left(\sum\limits_{i=1}^{n} e^{x_i}\right)$;

   (b) $f(x) = \frac{\|\mathrm{A}x-b\|^2}{1-x^\top x},\ \mathrm{X} = \{x \mid \|x\|^2 \leq 1\}$;

   (c) $F(x) = \frac{f^2(x)}{g(x)}$ при условии, что $f(x)$ является выпуклой и принимает только неотрицательные значения, $g(x)$ вогнута и принимает только положительные значения;

   (d) $f(x,t) = -\log(t^2 - x^T x),\ \mathrm{E} = \{(x,t)\} \in \mathbb{R}^{n\times 1} \mid \|x\|_2 < t\}$;

   (e) $f(x) = \sum\limits_{i=1}^{n} w_i \ln\left(1 + \exp(a_i^\top x)\right) + \frac{\mu}{2}\|x\|_2^2,\ \ \mu, w_1, \ldots, w_n > 0,\ a_i \in \mathbb{R}^n$;

   (f) $f(x) = \ln \sum\limits_{i=1}^{n} \exp\left(\max\{0, x_i\}^2\right)$.

ТЕСТ/ВЫПУКЛЫЕ ФУНКЦИИ (2 балла)

1. Какое максимальное количество точек минимума может быть у выпуклой функции?

   Ответ: (d) Несчётное множество
   Комментарий: В случае, если речь НЕ идет про строгую выпуклость и глобальный минимум,.

2. Какие из следующих функций выпуклы?
   Ответ:
   (a) $f(x) = \sum_{i=1}^{r} x_{[i]}$, где $x_{[1]} \geq x_{[1]} \geq \ldots \geq x_{[n]}$ — упорядоченные по убыванию элементы вектора x

   (d) $f(A) = x^\top A x$, где $A \in \mathbb{R}^{n\times n}$, $x \in \mathbb{R}^n$ — фиксированный вектор
   (e) $\|x\|_{1/3} = (\sum_{i=1}^{n} |x_i|^{1/3})^3$

3. Чем является касательная к графику выпуклой функции в каждой точке области определения?

   Ответ: (b) Глобальной оценкой снизу

4. На каком множестве выпукла функция $f = x_2^3 + x_2(x_1^2 + x_3^2)$?

   Ответ: (c) $\{(x_1, x_2, x_3) \mid x_2 \geq 0,\ 3x_2^2 - x_1^2 \geq x_3^2\}$
   Комментарий: получаем

$$\begin{vmatrix} 2x_2 & 2x_1 & 0 \\ 2x_1 & 6x_2 & 2x_3 \\ 0 & 2x_3 & 2x_2 \end{vmatrix}$$

   Угловые миноры:

$$3x_2^2 - x_1^2 \geq 0$$

$$-8x_2(x_1^2 - 3x_2^2 + x_3^2) \geq 0$$

В скобке имеем предыдущий минор(который неотрицателен), умноженный на -1.

5. Может ли область определения выпуклой функции быть невыпуклым множеством?

   Ответ: (b) Нет

6. Отметьте все композиции $f(x) = h(g(x))$, такие что $f(x)$ — выпуклая функция.

   Ответ: (c) $g(x) = Ax + b$ для некоторой матрицы A и вектора b, h – выпуклая функция

7. Чему равна константа сильной выпуклости у функции $f(x) = \frac{1}{2}\|x\|_2^2$?

   Ответ: (b) 2

ТЕСТ/ВЕКТОРНОЕ ДИФФЕРЕНЦИРОВАНИЕ (2 балла)

1. Посчитайте $df(x)$ и $\nabla f(x)$ для функции $f(x) = \log(x^\top \mathrm{A} x)$ и выберите правильный вариант ответа.
   (b) $\nabla f(x) = \frac{\left(\mathrm{A}+\mathrm{A}^\top\right)x}{x^\top \mathrm{A} x}$, $df(x) = \frac{x^\top\left(\mathrm{A}+\mathrm{A}^\top\right)dx}{x^\top \mathrm{A} x}$

2. Посчитайте $df(x)$, $d^2f(x)$ и $\nabla f(x)$, $\nabla^2 f(x)$ для функции $f(x) = \frac{1}{p}\|x\|_2^p$, $p > 1$ и выберите все правильные результаты.
   (a) $\nabla f(x) = \|x\|_2^{p-1} x$
   (d) $\nabla^2 f(x) = (p-1)\|x\|_2^{p-3} x x^\top + \|x\|_2^{p-1} I$
   (f) $d^2 f(x) = dx^\top \left((p-1)\|x\|_2^{p-3} x x^\top + \|x\|_2^{p-1} I\right) dx$

3. Верно ли, что у функции $f(x) = \frac{1}{n}\sum\limits_{i=1}^{n} \log\left(1 + \exp(a_i^\top x)\right) + \frac{\mu}{2}\|x\|_2^2$, $a_i \in \mathbb{R}^n$, $\mu > 0$, гессиан $\nabla^2 f(x) = \frac{1}{n}\sum\limits_{i=1}^{n}\left(-\frac{\exp(a_i^\top x)}{(1+\exp(a_i^\top x))^2} a_i a_i^\top + \frac{\exp(a_i^\top x)}{1+\exp(a_i^\top x)} a_i a_i^\top\right) + \mu \mathrm{I}$?
   (a) Да