Optimization methods

Makhmood Sodikov

Содержание

Задание № 1

Задача 1

Условие: Доказать, что множество $M = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 \leq x_2\}$ является выпуклым.

Решение: $\square$

Будем доказывать по определению [1].

Достаточно показать, что для любых двух точек $\mathbf{x_1}, \mathbf{x_2} \in M$ и $\forall \theta \in [0;1]$ выполнено $\theta \mathbf{x_1} + (1-\theta)\mathbf{x_2} \in M$. Сделаем явную проверку. Пусть для $\mathbf{x_1}$ верно $x_{11}^2 \leq x_{12}$, где x_{11} и x_{12} – это, соответственно, координаты 1-ая и 2-ая координаты вектора двумерной плоскости. Аналогично и для $\mathbf{x_2}$ – это x_{21} и x_{22} и для них верно аналогичное выражение. Покажем теперь, что для нового вектора $\mathbf{x_3} \in \mathbb{R}^2, \mathbf{x_3} = \theta \mathbf{x_1} + (1-\theta)\mathbf{x_2}$ верно, что $x_{31}^2 \leq x_{32}$:

$$\mathbf{x_3} = \theta \mathbf{x_1} + (1-\theta)\mathbf{x_2} = \{\theta x_{11} + (1-\theta)x_{21}; \theta x_{12} + (1-\theta)x_{22}\}$$

Для любого $\theta \in [0;1]$. Получается, что искомые координаты $\mathbf{x_3} : x_{31} = \theta x_{11} + (1-\theta)x_{21}, x_{32} = \theta x_{12} + (1-\theta)x_{22}$. Нам же надо доказать, что $x_{31}^2 \leq x_{32}$ вне зависимости от значения θ:

$$x_{31}^2 = (\theta x_{11} + (1-\theta)x_{21})^2 = \theta^2 x_{11}^2 + 2\theta(1-\theta)x_{11}x_{21} + (1-\theta)^2 x_{21}^2$$

$$x_{32} = \theta x_{12} + (1-\theta)x_{22}$$

Теперь выберем заранее $\hat{x} := \max\{x_{11}; x_{21}\}$ и учтем, что $x_{11}^2 \leq x_{12}$ (1) и $x_{21}^2 \leq x_{22}$ (2):

$$x_{31}^2 = \theta^2 x_{11}^2 + 2\theta(1-\theta)x_{11}x_{21} + (1-\theta)^2 x_{21}^2 =$$

$$= (\theta - \theta + \theta^2)x_{11}^2 + 2\theta(1-\theta)x_{11}x_{21} + (1 - 2\theta + \theta^2)x_{21}^2 =$$

$$= (\theta - \theta + \theta^2)x_{11}^2 + 2\theta(1-\theta)x_{11}x_{21} + (1 - \theta - \theta + \theta^2)x_{21}^2 =$$

$$= \theta x_{11}^2 + (-\theta + \theta^2)x_{11}^2 + 2(\theta - \theta^2)x_{11}x_{21} + (1-\theta)x_{21}^2 + (-\theta + \theta^2)x_{21}^2 =$$

$$= \theta x_{11}^2 - (\theta-\theta^2)x_{11}^2 + 2(\theta-\theta^2)x_{11}x_{21} + (1-\theta)x_{21}^2 - (\theta-\theta^2)x_{21}^2 = (*)$$

Теперь заменим выражение $2(\theta-\theta^2)x_{11}x_{21}$ при условии, что $x_{11}x_{21} \leq \hat{x}^2$ и, конечно, $x_{11}^2 \leq \hat{x}^2, x_{21}^2 \leq \hat{x}^2$:

$$(*) \leq \theta x_{11}^2 - \underline{(\theta-\theta^2)\hat{x}^2} + \underline{2(\theta-\theta^2)\hat{x}^2} + (1-\theta)x_{21}^2 - \underline{(\theta-\theta^2)\hat{x}^2} =$$

$$= \theta x_{11}^2 + (1-\theta)x_{21}^2.$$

А получившееся выражение, очевидно, меньше, чем $\theta x_{12}+(1-\theta)x_{22} = x_{32}$ в силу (1) и (2). ■

Задача 2

Условие: Пусть $S_1, \ldots, S_k$ – произвольные непустые множества в $\mathbb{R}^n$. Доказать, что:

$$\mathbf{cone}\left(\bigcup_{i=1}^{k} S_i\right) = \sum_{i=1}^{k} \mathbf{cone}(S_i)$$

$$\mathbf{conv}\left(\bigcup_{i=1}^{k} S_i\right) = \sum_{i=1}^{k} \mathbf{conv}(S_i)$$

Решение: □

Рассмотрим только первое выражение, второе полностью аналогично.

Чтобы построить коническую оболочку [2] объединения k множеств, в самом общем случае, когда эти множества могут и не пересекаться, можем воспользоваться следующим приемом: будем просто выбирать для нашей итоговой суммы с θ-ми иксы из системы различных представителей (а множества, которые пересекаются можем идентифицировать как одно множество и выбирать по представителю из пересечения этих множеств). Тогда понятно, что нам надо просто перебрать все разные элементы из каждого класса эквивалентности – сумма этого перебора и будет по определению коническая оболочка для объединения k множеств. Переобозначим теперь уже наши классы эквивалентности через и их количество через . Тогда получаем следующую сумму:

$$\sum_{i=1}^{\hat{k}} \sum_{j=1}^{|\hat{S}_i|} \theta_i x_i \text{ , где все } \theta \in \mathbb{R}_+ \cup \{0\}$$

Но это по определению и есть $\sum_{i=1}^{k} \mathbf{cone}(S_i)$, но в котором просто "съедены"ненужные зависимости от других комбинаций (которые ими же и балансируются). ■

Задача 3

Условие: Доказать, что множество $S \subseteq \mathbb{R}^2$ выпукло $\iff$

$$(\alpha + \beta)S = \alpha S + \beta S \quad \forall \alpha, \beta \in \mathbb{R}_+ \cup \{0\}$$

.

Решение: □

Справа налево доказательство очевидное. Так как верно для любых α и β, то возьмем $\alpha = \theta$, $\beta = 1 - \theta$, где $\theta \in [0; 1]$. То есть $\alpha + \beta = 1$, а это значит, что получаем просто само множество, при попытке взять сумму всех векторов с коэффициентом α и коэффициентом β. Отсюда напрямую следует, что множество выпукло по определению.

Обратное следствие: пусть наше множество выпукло. Для всех неотрицательных α и β множество $(\alpha + \beta)S$ тоже выпукло. Так как $S \subseteq \mathbb{R}^2$, то верно, что при увеличении всех векторов этого множества скалярно в $\alpha + \beta$ раз, мы получим то же самое, что и при суммировании векторов после последовательного скалярного умножения на α и β. А значит, тождество верно для всех выпуклых множеств из $\mathbb{R}^2$. $\blacksquare$

Задача 4

Условие: Доказать, что множество выпукло:

$$M = \{x \in \mathbb{R}^2 | \exp(x_1) < x_2\}$$

Решение: $\square$

Как обычно, будем рассматривать некоторое фиксированное $\theta \in [0; 1]$ и некоторые $x, y \in M$. Докажем, что для любого z, образованного выпуклой комбинацией x, y, будет верно, что $z \in M$.

Так как x, y в нашем множестве, знаем, что для них выполнено условие принадлежности множеству, то есть $\exp(x_1) < x_2$ и $\exp(y_1) < y_2$ **(*)**.

Рассмотрим тогда z_1, z_2:

$$\begin{cases} z_1 = \theta x_1 + (1 - \theta) y_1 \\ z_2 = \theta x_2 + (1 - \theta) y_2 \end{cases}$$

Тогда, рассмотрим $\exp(z_1) = \exp(\theta x_1 + (1 - \theta) y_1) = \exp(\theta x_1) + \exp((1 - \theta) y_1) = (e^{x_1})^{\theta} (e^{y_1})^{(1-\theta)}$. Вспоминаем про **(*)**: $\exp(z_1) \leq x_2^{\theta}\ y_2^{(1-\theta)}$. А это, в свою очередь, не более $\theta x_2 + (1 - \theta) y_2 = z_2$. $\blacksquare$

Задача 5

Условие:

Решение: $\square$

$\blacksquare$

Задача 6

Условие: Являются ли выпуклыми следующие множества:

а) Полоса, $\{x \in \mathbb{R}^n \mid \alpha \leq a^T x \leq b\}$

б) Прямоугольник, $\{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq b_i \ , \ i = 1, \ldots, n\}$

в) Клин $\{x \in \mathbb{R}^n \mid a_1^T x \leq b_1 \ , \ a_2^T x \leq b_2\}$

г) Множество точек, более близких к заданной точке, чем к заданному множеству, не содержащему точку, $\{x \in \mathbb{R}^n \mid ||x - x_0||_2 \leq ||x - y||_2, \forall y \in S \subset \mathbb{R}_n\}$

Решение: $\square$

а) Является, так как любые две точки полосы M можно соединить отрезком, лежащим в полосе. Формально:

$$\forall x, y \in M \hookrightarrow \alpha \leq a^T x \leq \beta, \quad \alpha \leq a^T y \leq \beta$$

$$\begin{cases} \theta a^T x + (1 - \theta) a^T y \leq \theta\beta + (1 - \theta)\beta = \beta \\ \theta a^T x + (1 - \theta) a^T y \geq \theta\alpha + (1 - \theta)\alpha = \alpha \end{cases}$$

б) Является, так как любые две точки в прямоугольнике можно соединить отрезком, лежащим в прямоугольнике. (Формально аналогично предыдущему пункту).

в) Является, так как любые две точки в клине можно соединить отрезком, лежащим в клине. (Формально аналогично первому пункту).[1]

г)

[1]Вообще, клин выглядит как гиперполуплоскость, лежащая ниже двух непараллельных гиперпрямых. Поэтому в $\mathbb{R}^2$ интуитивно понятно, почему это выпуклое множество. Формально – аналогично 1-му пункту.

Задание № 2

В этом разделе необходимо для всех функций $f(x)$ ниже: найти $\nabla f(x)$, $f''(x)$. Проверить $f''(x)$ на (полу)определенность (положительную, отрицательную).

Задача 1

Условие: $f(x) = ||x||_2$

Решение:

$\square$

1) **Найдем градиент (набла).**

Мы знаем, что $||x||_2 = \sqrt{\langle x, x\rangle} = \langle x, x\rangle^{\frac{1}{2}}$

Для нахождения градиента продифференцируем:

$$df(x) = d(\langle x, x\rangle^{\frac{1}{2}})$$

Получается дифференциал сложной функции – $g(y) = y^{1/2}$, а $t(x) = \langle x, x\rangle$. Тогда наша исходная функция получается равной $f(x) = g(t(x)) = \langle x, x\rangle^{\frac{1}{2}}$

Тогда применим chain rule для композиции функций и вычислим с помощью него дифференциал.

$$dt(x) = d(\langle x, x\rangle) = 2\langle x, dx\rangle$$
$$dg(y) = d(y^{1/2}) = \frac{1}{2}y^{-1/2}dy$$

Подставим теперь **явно** вместо y в $dg(y)$ значение $t(x)$, а вместо dy в $dg(y)$ значение $dt(x)$:

$$df(x) = \frac{1}{2}\langle x, x\rangle^{-1/2}2\langle x, dx\rangle = \langle x, x\rangle^{-1/2}\langle x, dx\rangle = \frac{\langle x, dx\rangle}{||x||}$$

Таким образом, исходя из того, что $df(x) = \langle \nabla f(x), dx\rangle$, получаем, что для нашей исходной функции $\nabla f(x) = \frac{x}{||x||}$

2) **Найдем вторую производную $f''(x)$.**

Обозначим из прошлого пункта $df(x) = g(x) = \frac{\langle x, dx\rangle}{||x||}$ и примем $dx = dx_1$. Тогда: $g(x) = \frac{\langle x, dx_1\rangle}{||x||}$

Теперь для поиска второй производной, нам необходимо продифференцировать $g(x)$ как функцию от x, приняв dx_1 за константный вектор.

$$d^2f(x) = dg(x) = d\Big(\frac{\langle x, dx_1\rangle}{||x||}\Big)$$

В последнем выражении цепочки можем увидеть применение формулы дифференциала отношени двух функций. Тогда:

$$d\Big(\frac{\langle x, dx_1\rangle}{||x||}\Big) = \frac{d(\langle x, dx_1\rangle)||x|| + d(||x||)\langle x, dx_1\rangle}{||x||^2} \qquad (*)$$

Посчитаем компоненты отдельно:

$$||x||^2 = \langle x, x\rangle$$

$$d(\langle x, dx_1\rangle) = \langle dx_2, dx_1\rangle$$

$$d(||x||) = df(x) = \frac{\langle x, dx_2\rangle}{||x||}$$

Теперь подставим обратно и приведем к каноническому виду:

$$(*) = \frac{\langle dx_2, dx_1\rangle||x|| + \frac{\langle x, dx_2\rangle}{||x||}\langle x, dx_1\rangle}{||x||^2} =$$

$$= \langle\frac{dx_2||x||^2 + \langle x, dx_2\rangle}{||x||^3}, x_1\rangle = \langle\frac{||x||^2 + x}{||x||^3}dx_2, x_1\rangle = \langle\frac{\frac{1}{}||x||^2 + x}{||x||^3}dx_2, x_1\rangle$$

■

Задача 2

Условие: $f(x) = ||(A \cdot x + b)_+||_2$, где $()_+$ это покомпонентная положительная срезка (положительные компоненты остаются, а отрицательные заменяются нулем).

Решение:

$\square$

in process...

$\blacksquare$

Задача 3

Условие: $f(x) = ln(1 + x^T \cdot A \cdot x)$, $A \geq 0$.

Решение:

$\square$

Уже не будем досконально расписывать все переходы и преобразования, а просто будем фиксировать конечный результат.

По тому же правилу, что и в первом задании:

Найдем $\nabla f(x)$:

$$f(x) = ln(1 + \sum_{i,j} a_{i,j} \cdot x_i \cdot x_j)$$

Преобразуем в вид поэлементной записи (чтобы было удобнее дифференцировать).

$$(\nabla f)_i = \frac{\partial f(x)}{\partial x_i} = \frac{1}{1 + \sum_{i,j} a_{i,j} \cdot x_i \cdot x_j} \cdot (\sum_j a_{ij} \cdot x_j + \sum_j a_{ji} \cdot x_k)$$

Значит, $\nabla f(x)$:

$$\nabla f = \frac{1}{1 + x^T \cdot A \cdot x} \cdot ((A + A^T) \cdot x)$$

Можно было также по правилам дифференцирования постепенно получить конечный ответ сначала в виде векторного произведения, а оттуда уже вычленять дивергенцию и гессиан.

$\blacksquare$

Задача 4

Условие:$f(x) = \sum_{i=1}^{n} x_i \cdot ln(x_i)$, $x \geq 0$

Решение:

$\square$

Найдем $\nabla f(x)$:

$$(\nabla f)_i = \frac{\partial f(x)}{\partial x_i} = 1 + ln(x_i)$$

Значит, $\nabla f(x)$:

$$\nabla f = (1) + ln(x)$$

Найдем $f''(x)$:

$$f''(x) = \frac{d(\nabla f)}{dx^T}$$

Найдем (i, j) компоненту матрицы:

$$f''_{ij}(x) = \frac{\partial(\nabla f)_i}{\partial x_j} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{1}{x_j}(i = j)$$

Значит,

$$f''(x) = Diag(1/x)$$

где $1/x = (\frac{1}{x_i})$ и $Diag(t) = (a_{ii}), i \in \{1, n\}, a_{ii} = t_i$. Докажем, что матрица положительно определена:

Для любого ненулевого вектора y верно, что $y^T \cdot f'' \cdot y = Diag(y_i^2)$.

А значит, для любого ненулевого получается сумма квадратов, а значит матрица положительно определена.

$\blacksquare$

Задача 5

Условие: $f(x) = -\frac{1}{1+x^T \cdot x}$.

Решение:

$\square$

Найдем $\nabla f(x)$:

$$f(x) = \frac{1}{1 + x^T \cdot x} = \frac{1}{1 + x_1^2 + ... + x_n^2}$$

$$(\nabla f)_i = \frac{\partial f(x)}{\partial x_i} = \frac{2 \cdot x_i}{(1 + x_1^2 + ... + x_n^2)^2} = \frac{2 \cdot x_i}{(1 + x^T \cdot x)^2}$$

Значит, $\nabla f(x)$:

$$\nabla f = \frac{2}{(1 + x^T \cdot x)^2} \cdot x$$

Найдем $f''(x)$:

$$f''(x) = \frac{d(\nabla f)}{dx^T}$$

Найдем (i, j) компоненту матрицы:

$$f''_{ij}(x) = \frac{\partial (\nabla f)_i}{\partial x_j} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{2}{(1 + x_1^2 + ... + x_n^2)^2}(i = j) - \frac{8 \cdot x_i \cdot x_j}{(1 + x_1^2 + ... + x_n^2)^3} =$$

$$= \frac{2}{(1 + x^T \cdot x)^2}(i = j) - \frac{8 \cdot x_i \cdot x_j}{(1 + x^T \cdot x)^3}$$

Значит,

$$f''(x) = \frac{2}{(1 + x^T \cdot x)^2} \cdot E - \frac{8}{(1 + x^T \cdot x)^3} \cdot x \cdot x^T$$

$\blacksquare$

Задание № 3

Задача 1

Условие: Проверить на выпуклость $f(x,y) = \frac{x^2}{y}, dom f = \{(x,y) \in R^2 | y > 0\}$.

Решение:

$\square$

Понятно, что $dom f$ - выпуклое множество. Докажем выпуклость f по критерию: $f(x)$ - выпукла $\Leftrightarrow \nabla^2 f(x) \geq 0$

Найдем первые производные:

$$\frac{\partial f(x)}{\partial x} = \frac{2 \cdot x}{y}$$

$$\frac{\partial f(x)}{\partial y} = -\frac{x^2}{y^2}$$

Найдем вторые производные:

$$\frac{\partial^2 f(x)}{\partial x^2} = \frac{2}{y}$$

$$\frac{\partial^2 f(x)}{\partial y^2} = \frac{2 \cdot x^2}{y^3}$$

$$\frac{\partial^2 f(x)}{\partial y \cdot \partial x} = \frac{\partial^2 f(x)}{\partial x \cdot \partial y} = -\frac{2 \cdot x}{y^2}$$

Значит, наша искомая матрица $\nabla^2 f(x)$:

$$\begin{pmatrix} \frac{2}{y} & -\frac{2 \cdot x}{y^2} \\ -\frac{2 \cdot x}{y^2} & \frac{2 \cdot x^2}{y^3} \end{pmatrix}$$

Проверим её на положительную полуопределенность:

$\frac{2}{y} > 0$ в $dom f$;

$\det(\nabla^2 f(x)) = 0$; Значит, так как главные миноры неотрицательны, то матрица положительно полуопределена в $dom f$, а значит по критерию выпуклости функции, $f(x)$ – выпукла.

Задача 2

Условие: Показать, что функция вогнута $f(x) = (\prod\limits_{i=1}^{n} x_i)^{\frac{1}{n}}$, $dom f = R^n_{++}$

Решение:

$\square$

Рассмотрим первую производную:

$$\frac{\partial f(x)}{\partial x_i} = \frac{1}{n} \cdot (\prod_{j=1}^{n} x_i)^{\frac{1-n}{n}} \cdot (\prod_{j=1, j\neq i}^{n} x_i)$$

По утверждению $f(x)$ - вогнута $\Leftrightarrow f(y) \leq f(x) + < \nabla f(x), y - x >$.

$$f(y) - f(x) - < \nabla f(x), y - x > = (\prod_{i=1}^{n} y_i)^{\frac{1}{n}} - (\prod_{i=1}^{n} x_i)^{\frac{1}{n}} - \sum_{i=1}^{n} (\frac{1}{n} \cdot (\prod_{j=1}^{n} x_j)^{\frac{1-n}{n}} \cdot (\prod_{j=1, j\neq i}^{n} x_j)) \cdot (y_i - x_i) =$$

$$= (\prod_{i=1}^{n} y_i)^{\frac{1}{n}} - (\prod_{i=1}^{n} x_i)^{\frac{1}{n}} - \frac{1}{n} \cdot (\prod_{j=1}^{n} x_j)^{\frac{1}{n}} \cdot \sum_{i=1}^{n} (\frac{y_i}{x_i} - 1) = (\prod_{j=1}^{n} x_j)^{\frac{1}{n}} ((\frac{\prod\limits_{j=1}^{n} y_j}{\prod\limits_{j=1}^{n} x_j})^{\frac{1}{n}} - \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{y_i}{x_i})$$

Заметим, что $(\prod\limits_{j=1}^{n} x_j)^{\frac{1}{n}} \geq 0$ в силу области определения.

То есть осталось рассмотреть знак второй скобки. Вторую скобку можно интерпретировать как разницу между средним геометрическим и средним арифметическим. В силу неравенств, получаем, что эта скобка всегда неположительна.

Значит, произведение этих скобок не положительно. Получаем вогнутость функции $f(x)$.

$\blacksquare$

Задача 3

Условие: Выпуклы ли следующие функции: $f(x) = e^x - 1, x \in R$; $f(x_1, x_2) = x_1 \cdot x_2, x \in R^2_{++}$; $f(x_1, x_2) = \frac{1}{x_1 \cdot x_2}, x \in R^2_{++}$?

Решение:

$\square$

1. Рассмотрим $f(x) = e^x - 1, x \in R$.
   $f'(x) = e^x$
   $f''(x) = e^x$
   Значит, так как $f''(x) > 0 \ \forall x$, получаем выпуклость функции $f(x)$ в силу критерия.

2. Рассмотрим $f(x_1, x_2) = x_1 \cdot x_2, x \in R^2_{++}$
   Гессиан G:
   $$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
   Значит, т.к. $\det(G) = -1 < 0$, то матрица не положительно полуопределена, а значит, функция $f(x_1, x_2)$ не является выпуклой в силу критерия.

3. Рассмотрим $f(x_1, x_2) = \frac{1}{x_1 \cdot x_2}, x \in R^2_{++}$
   $$\frac{\partial f(x)}{\partial x_1} = -\frac{1}{x_1^2 \cdot x_2}, \ \frac{\partial f(x)}{\partial x_2} = -\frac{1}{x_1 \cdot x_2^2}$$
   $$\frac{\partial^2 f(x)}{(\partial x_1)^2} = \frac{2}{x_1^3 \cdot x_2}, \ \frac{\partial^2 f(x)}{(\partial x_2)^2} = \frac{2}{x_2^3 \cdot x_1}$$
   $$\frac{\partial^2 f(x)}{\partial x_1 \cdot \partial x_2} = \frac{\partial^2 f(x)}{\partial x_2 \cdot \partial x_1} = \frac{1}{x_1^2 \cdot x_2^2}$$
   Гессиан G:
   $$\begin{pmatrix} \frac{2}{x_1^3 \cdot x_2} & \frac{1}{x_1^2 \cdot x_2^2} \\ \frac{1}{x_1^2 \cdot x_2^2} & \frac{2}{x_2^3 \cdot x_1} \end{pmatrix}$$

Рассмотрим главные миноры:

$\frac{2}{x_1^3 \cdot x_2} > 0$ в нашей области определения,

$\det(G) = \frac{3}{x_1^2 \cdot x_2^4} > 0$ на всём R^2.

Значит, так как матрица G положительно полуопределена. получаем выпуклость $f(x_1, x_2)$ в силу критерия выпуклости.

■

Задача 4

Условие: Расстоянием Кульбака – Лейблера между $p, q \in R^n_{++}$ называется:

$$D(p,q) = \sum_{i=1}^{n} (p_i \cdot \ln(\frac{p_i}{q_i}) - p_i + q_i)$$

Доказать, что $D(p,q) \geq 0, \forall p, q \in R^n_{++}$ и $D(p,q) = 0 \Leftrightarrow p = q$.

Решение:

$\square$

Рассмотрим $g(x) = x \cdot \ln x$:

$g'(x) = 1 + \ln(x)$

$g''(x) = 1/x$

Значит, $g(x)$ - выпуклая функция на R_+

Значит, $f(p)$ - выпуклая функция (как сумма выпуклых) на R^n_{++}

По утверждению, т.к. f - выпукла, получаем $f(p) - f(q) - \nabla f(q)^T(p-q) \geq 0$, а значит $D(p,q) \geq 0 \ \forall x \in R^n_{++}$.

Почему $D(p,q) = 0 \Leftrightarrow p = q$?

Геометрически дифференциальный критерий выпуклости означает, что значения функции $f(x)$ лежат выше гиперплоскости $\{(x,t) | t = f(x) + < \nabla f(x), y - x >\}$, а касание к графику происходит в точке $(y, f(y))$. Значит, только в этой точке значения и равны, т.е. в точке $p = q$ в наших обозначениях.

$\blacksquare$

Задание № 4

Задача 1

Условие:

$$(x_1 - 3) \cdot (x_2 - 2)$$

$$s.t. \; x_1 + 2 \cdot x_2 = 4$$

$$x_1^2 + x_2^2 \leq 5$$

$$x_1 \geq 0, \; x_2 \geq 0$$

Решение:

$\square$

Введем обозначения:

$$f(x) = (x_1 - 3) \cdot (x_2 - 2)$$

$$h(x) = x_1 + 2 \cdot x_2 - 4$$

$$g_1(x) = x_1^2 + x_2^2 - 5$$

$$g_2(x) = -x_1$$

$$g_3(x) = -x_2$$

Обозначим $x \in R^2$, удовлетворяющие условиям:

1. $h(x) = 0$
2. $g_1(x) \leq 0$
3. $g_2(x) \leq 0$
4. $g_3(x) \leq 0$

за $\mathcal{Q}$.

Рассмотрим Лагранжиан:

$$\mathcal{L}(x;\lambda,\mu) = f(x) + \sum_{i=1}^{3} \lambda_i \cdot g_i(x) + \mu_1 \cdot h(x) =$$

$$= (x_1 - 3) \cdot (x_2 - 2) + \lambda_1 \cdot (x_1^2 + x_2^2 - 5) - \lambda_2 \cdot x_1 - \lambda_3 \cdot x_2 + \mu_1 \cdot (x_1 + 2 \cdot x_2 - 4)$$

Необходимое условие точки минимума задачи с ограничениями:

$$\begin{cases} x^* \in \mathcal{Q} \\ \frac{\partial \mathcal{L}}{\partial x_1^*} = (x_2^* - 2) + 2 \cdot \lambda_1 \cdot x_1^* - \lambda_2 + \mu_1 = 0 \\ \frac{\partial \mathcal{L}}{\partial x_2^*} = (x_1^* - 3) + 2 \cdot \lambda_2 \cdot x_2^* - \lambda_3 + 2 \cdot \mu_1 = 0 \\ \lambda_1^* \cdot ((x_1^*)^2 + (x_2^*)^2 - 5) = 0 \\ \lambda_2^* \cdot x_1^* = 0 \\ \lambda_3^* \cdot x_2^* = 0 \\ \lambda_i \geq 0, i \in [1,2,3] \end{cases}$$

Решая данную систему получаем ответ задачи

Задача 2

Условие:

$$\sum_{i=1}^{n} \frac{c_i}{x_i}$$

$$s.t. \sum_{i=1}^{n} a_i \cdot x_i \leq b$$

$$x_i > 0,\ b > 0,\ c_i > 0,\ a_i > 0$$

Решение:

$\square$

Перед нами выпуклая задача, а значит, достаточно проверить только необходимое условие.

$$f(x) = \sum_{i=1}^{n} \frac{c_i}{x_i}$$

$$g_1(x) = \sum_{i=1}^{n} a_i \cdot x_i - b$$

$$g_i(x) = -x_{i-1}, i \in [2, ...n+1]$$

$$b > 0,\ c_i > 0,\ a_i > 0$$

Обозначим $x \in R^2$, удовлетворяющие условиям:

1. $g_i(x) \leq 0, i \in [1, ...n+1]$

за $\mathcal{Q}$.

Рассмотрим Лагранжиан:

$$\mathcal{L}(x); \lambda, \mu) = f(x) + \sum_{i=1}^{n+1} \lambda_i \cdot g_i(x) =$$

$$= \sum_{i=1}^{n} \frac{c_i}{x_i} + \lambda_1 \cdot (\sum_{j=1}^{n} a_j \cdot x_j - b) - \sum_{i=1}^{n} \lambda_{i+1} \cdot x_i$$

Необходимое и достаточное условие точки минимума задачи с ограничениями:

$$\begin{cases} x^* \in \mathcal{Q} \\ \frac{\partial \mathcal{L}}{\partial x_i^*} = -\frac{c_i}{(x_i^*)^2} + \lambda_1 \cdot a_i - \lambda_{i+1} = 0, i \in [1, ..., n] \\ \lambda_i^* = 0, i \in [2, ..., n+1], \\ \lambda_1^* \cdot (\sum_{j=1}^n a_j \cdot x_j^* - b) = 0 \\ \lambda_i^* \geq 0, i \in [1, ..., n+1] \end{cases}$$

Рассмотрим случаи:

1. $\lambda_1^* = 0$. Значит, $-\frac{c_i}{(x_i^*)^2} = 0$. Значит, $c_i = 0$ - противоречие.

2. $\lambda_1^* \neq 0$.

$$\begin{cases} \frac{c_i}{(x_i^*)^2} = \lambda_1^* \cdot a_i, \\ \sum_{i=1}^n a_i \cdot x_i^* = b. \end{cases}$$

Значит, выразив x_i^* из первого уравнения, получим:

$$\sum_{i=1}^n a_i \cdot x_i^* = \sum_{i=1}^n \sqrt{\frac{a_i \cdot c_i}{\lambda_1^*}} = b$$

Выразив λ_1:

$$x_i^* == \frac{b}{\sum_{j=1}^n \sqrt{a_j \cdot c_j}} \cdot \sqrt{\frac{c_i}{a_i}}$$

Так как задача аналитическая, то дадим аналитический ответ: $x^* = (x_1^*, ..., x_n^*)$, где каждую компоненту вычисляем из условия выше. ■

Задача 3

Условие:

$$(x_1 \cdot x_3 - 2 \cdot x_2)$$

$$s.t.\ 2 \cdot x_1 - x_2 - 3 \cdot x_3 \leq 10$$

$$3 \cdot x_1 + 2 \cdot x_2 + x_3 = 6$$

$$x_2 \geq 0$$

Решение: $\square$

Введем обозначения:

$$f(x) = x_1 \cdot x_3 - 2 \cdot x_2$$

$$g_1(x) = 2 \cdot x_1 - x_2 - 3 \cdot x_3 - 10$$

$$g_2(x) = -x_2$$

$$h(x) = 3 \cdot x_1 + 2 \cdot x_2 + x_3 - 6$$

Обозначим $x \in R^2$, удовлетворяющие условиям:

1. $g_i(x) \leq 0, i = 1, 2$
2. $h(x) = 0$

за $\mathcal{Q}$.

Рассмотрим Лагранжиан:

$$\mathcal{L}(x; \lambda, \mu) = f(x) + \sum_{i=1}^{2} \lambda_i \cdot g_i(x) + \mu \cdot h(x) =$$

$$= (x_1 \cdot x_3 - 2 \cdot x_2) + \lambda_1 \cdot (2 \cdot x_1 - x_2 - 3 \cdot x_3 - 10) - \lambda_2 \cdot x_2 + \mu \cdot (3 \cdot x_1 + 2 \cdot x_2 + x_3 - 6)$$

Необходимое условие точки минимума задачи с ограничениями:

$$
\begin{cases}
x^* \in \mathcal{Q} \\
\frac{\partial \mathcal{L}}{\partial x_1^*} = x_3^* + 2 \cdot \lambda_1 + 3 \cdot \mu = 0 \\
\frac{\partial \mathcal{L}}{\partial x_2^*} = -2 - \lambda_1 - \lambda_2 + 2 \cdot \mu = 0 \\
\frac{\partial \mathcal{L}}{\partial x_3^*} = x_1^* - 3 \cdot \lambda_1 + \mu = 0 \\
\lambda_2^* \cdot x_1 = 0 \\
\lambda_1^* \cdot (2 \cdot x_1^* - x_2^* - 3 \cdot x_3^* - 10) = 0 \\
\lambda_i^* \geq 0, i = 1, 2
\end{cases}
$$

Рассмотрим случаи: Рассмотрим случаи:

1. $\lambda_1^*, \lambda_2^* = 0$. Из первых трех условий и $h(x)$:

$$x_1^* = -1$$

$$x_2^* = 6$$

$$x_3^* = -3$$

$$\mu^* = 1$$

   Не является решением, так как не удовлетворяет $g_1(x)$

2. $\lambda_1^* = 0, \lambda_2^* \neq 0$

$$x_1^* = 1$$

$$x_2^* = 0$$

$$x_3^* = 3$$

$$\mu^* = -1$$

$$\alpha_2 = -4$$

   Как видно, $\lambda_2 < 0$ - противоречие.

3. $\lambda_2^* = 0, \lambda_1^* \neq 0$

$$x_1^* = \frac{2}{7}$$
$$x_2^* = \frac{60}{7}$$
$$x_3^* = -6$$
$$\lambda_1^* = \frac{18}{35}$$
$$\mu^* = \frac{44}{35}$$

Возможный минимум.

4. $\lambda_1^*, \lambda_2^* \neq 0$

$$x_1^* = \frac{28}{11}$$
$$x_2^* = 0$$
$$x_3^* = -\frac{18}{11}$$
$$\lambda_1^* = \frac{102}{121}$$
$$\lambda_2^* = -\frac{280}{121}$$
$$\mu^* = \frac{32}{121}$$

Не подходит, т.к.$\lambda_2 < 0$

Заметим, гессиан Лагранжиана G:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Детерминант равен 0, значит, точка $(\frac{2}{7};\ \frac{60}{7};\ -6)$ - точка минимума. ■

■

Определения и формулировки

По Заданию 1:

[1]: <u>Определение.</u> Множество M называется **выпуклым**, если $\forall x, y \in M \wedge \forall \theta \in [0;1] \hookrightarrow \theta x + (1-\theta)y \in M$.

[2]: <u>Определение.</u> Коническая оболочка множества S – это множество, которое получается объединением всех конических комбинаций элементов этого множества, то есть множество $H = \{\sum_{i=1}^{|S|} \theta_i x_i\}$, причем все $\theta \in \mathbb{R}_+ \cup \{0\}$.