Решение домашнего задания № 2

Роман Тарасов, группа 876

1. Задача первая

1.1 Пункт 1

$$\min_{(x,y)\in\mathbb{R}^2} |x| + |y|$$

$$\text{s.t.}\ \ (x-1)^2 + (y-1)^2 \leq 1$$

Обозначим множество $(x-1)^2 + (y-1)^2 \leq 1$ как $\mathcal{Y}$

Посчитаем субдифференциал

$$\partial_{\mathcal{Y}} f(\mathbf{x}) = \partial f(\mathbf{x}) + N(\mathbf{x} \mid \mathcal{Y})$$

Здесь $\mathbf{x} = (x\ y)^\top$

$$\partial f(\mathbf{x}) = \begin{pmatrix} \begin{cases} 1 & x > 0 \\ -1 & x < 0 \\ [-1, 1] & x = 0 \end{cases} \\ \begin{cases} 1 & y > 0 \\ -1 & y < 0 \\ [-1, 1] & y = 0 \end{cases} \end{pmatrix} \tag{1}$$

Множество $\mathcal{Y}$ есть круг с центорм в точке $(1, 1)$ и радиусом 1. Если $\mathbf{x} \in \partial\mathcal{Y}$, то нормальный конус в этой точке — нулевой вектор или любой вектор, который перпендикулярен касательной к окружности в точке $\mathbf{x}$, и который, если его отложить от точки $\mathbf{x}$, направлен в сторону от окружности. Поэтому

$$N(\mathbf{x} \mid \mathcal{Y}) = \begin{cases} k\left(\mathbf{x} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right), k \geq 0 & \mathbf{x} \in \partial\mathcal{Y} \\ \mathbf{0} & \mathbf{x} \in \operatorname{int}(\mathcal{Y}) \end{cases} \tag{2}$$

Теперь рассмотрим различные значения, которые может принимать $\mathbf{x}$ и найдём, на каком из них в субдифференциале содержится 0.

Заметим, что на нашем множестве $x, y \geq 0$.

Пусть $\mathbf{x} \in \partial\mathcal{Y}$, $x, y > 0$. Тогда

$$\partial f(\mathbf{x}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{3}$$

$$N(\mathbf{x} \mid \mathcal{Y}) = k\left(\mathbf{x} - \begin{pmatrix}1\\1\end{pmatrix}\right), k \geq 0 \tag{4}$$

Из равенства $\partial f(\mathbf{x}) + N(\mathbf{x} \mid \mathcal{Y}) = 0$ получаем систему

$$\begin{cases}1 + kx - k = 0\\ 1 + ky - k = 0\end{cases} \tag{5}$$

Отсюда $x = y$ и $k = \frac{1}{1-x}$. Так как $k \geq 0$, то $x, y < 1$.

На границе множества $\mathcal{Y}$ выполнено равенство $(x-1)^2 + (y-1)^2 = 1$. Подставим сюда $y = x$, получим $2(x-1)^2 = 1$. Отсюда $x = 1 \pm \frac{1}{\sqrt{2}}$. А так как выше мы получили, что $x < 1$, то окончательно получаем $x = 1 - \frac{1}{\sqrt{2}}$.

Точка $(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ лежит в множестве $\mathcal{Y}$ и также $0 \in \partial f\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. Значит, в этой точке достигается минимум.

$$|x| + |y| = 2 - \sqrt{2}$$

1.2 Пункт 2

$$\min_{\mathbf{x} \in \mathbb{R}^n} ||\mathbf{x}||_2 + \mathbf{c}^\top \mathbf{x}$$

$$\text{s.t. } x_i \geq 0, \quad i = 1, \ldots, n$$

Обозначим множество всех таких $\mathbf{x}$, что $x_i \geq 0, \quad i = 1, \ldots, n$ как $\mathcal{Y}$. А также $\mathbf{1}_i$ — вектор, i-тая компонента которого равна 1 а все остальные 0.

$$\partial f(\mathbf{x}) = \begin{cases}\frac{\mathbf{x}}{||\mathbf{x}||_2} & \mathbf{x} \neq \mathbf{0}\\ \{\mathbf{y} \mid ||\mathbf{y}||_2 \leq 1\} & \mathbf{x} = \mathbf{0}\end{cases} \tag{6}$$

Найдём нормальный конус $N(\mathbf{x} \mid \mathcal{Y})$. Если все $x_i > 0$, то $N(\mathbf{x} \mid \mathcal{Y}) = \mathbf{0}$. Если же какая-то компонента $x_i = 0$ равна 0, то $\forall k_i \leq 0, \forall \mathbf{y} \in \mathcal{Y} \langle k_i \mathbf{1}_i, \mathbf{y} - \mathbf{0}\rangle = k_i y_i \leq 0$. Поэтому вектор $k_i \mathbf{1}_i$ лежит в нормальном конусе. А если несколько компонент вектора $\mathbf{x}$ равны 0, то в нормальном конусе лежат соответственно все линейные комбинации векторов с неположительными коэффициентами $\mathbf{1}_i$ таких что $x_i = 0$. Поэтому нормальный конус к множеству $\mathcal{Y}$ есть

$$N(\mathbf{x} \mid \mathcal{Y}) = \begin{cases}\mathbf{0} & \forall i \ x_i > 0\\ \sum\limits_{\{i:\ x_i = 0\}} k_i \mathbf{1}_i & \exists i:\ x_i = 0\end{cases} \tag{7}$$

1) Пусть сумма квадратов отрицательных компонент вектора $\mathbf{c}$ не больше 1. Покажем, что в этом случае минимум достигается.

Пусть $\mathbf{x} = \mathbf{0}$. Тогда

$$\partial_{\mathcal{Y}} f(\mathbf{0}) = \{\mathbf{y} + \mathbf{c} + \sum_{i=1}^n k_i \mathbf{1}_i \mid ||\mathbf{y}||_2 \leq 1, k_i \leq 0\}$$

Пусть сумма квадратов отрицательных компонент вектора $\mathbf{c}$ не больше 1. Выберем такой вектор $\mathbf{y}$ и такие значения k_i, что субградиент в точке $\mathbf{x} = \mathbf{0}$ будет равен 0.

Если $c_i \geq 0$, то можем выбрать $k_i = -c_i$ и $y_i = 0$. Тогда i-тая компонента вектора $\mathbf{y} + \mathbf{c} + \sum_{i=1}^{n} k_i \mathbf{1}_i$ будет равна 0.

Если же $c_i < 0$, то выберем $y_i = -c_i$, а $k_i = 0$. Тогда i-тая компонента вектора $\mathbf{y} + \mathbf{c} + \sum_{i=1}^{n} k_i \mathbf{1}_i$ также обратится в 0. Заметим, что $||\mathbf{y}||_2^2 = \sum\limits_{i=1}^{n} y_i^2 = \sum\limits_{\{i:\ c_i<0\}} c_i^2 \leq 1$, поэтому выбранный нами вектор $\mathbf{y}$ удовлетворяет необходимым условиям. А значит, вектор $\mathbf{y} + \mathbf{c} + \sum_{i=1}^{n} k_i \mathbf{1}_i$ принадлежит субдифференциалу $\partial_{\mathcal{Y}} f(\mathbf{0})$.

Итак, для любого вектора $\mathbf{c}$, такого что $\sum\limits_{\{i:\ c_i<0\}} c_i^2 \leq 1$, минимум достигается в точке $\mathbf{x} = \mathbf{0}$ и равен $||\mathbf{0}||_2 + \mathbf{c}^\top \mathbf{0} = 0$.

2) Для векторов $\mathbf{c}$, таких что $\sum\limits_{\{i:\ c_i<0\}} c_i^2 > 1$ можно рассмотреть все случаи, посчитать $\partial_{\mathcal{Y}}(\mathbf{x})$ и показать, что ни для какого $\mathbf{x}$ число 0 не лежит в $\partial_{\mathcal{Y}}(\mathbf{x})$. А можно просто показать, что исследуемая функция не имеет минимума, приведя в пример соответствующие значения $\mathbf{x}$.

Рассмотрим вектор $\mathbf{x}$ такой, что $x_i = 0$, если $c_i > 0$, и иначе $x_i = -kc_i$ (k — некоторое положительное число). При любом $k > 0$ все компоненты вектор $\mathbf{x}$ будут неотрицательные. Тогда

$$||\mathbf{x}||_2 + \mathbf{c}^\top \mathbf{x} = \sqrt{\sum_{\{i:\ c_i<0\}} (kc_i)^2} - k \sum_{\{i:\ c_i<0\}} c_i^2 = k \sum_{\{i:\ c_i<0\}} c_i^2 \left(1 - \sum_{\{i:\ c_i<0\}} c_i^2\right)$$

$1 - \sum\limits_{\{i:\ c_i<0\}} c_i^2 < 0$, поэтому при $k \to +\infty$, мы устремим значение исследуемой функции в $-\infty$. То есть, минимум не достигается.

2. Задача вторая

2.1 Пункт 1

$$f^*(y) = \sup_{x>0} (yx - x^p)$$

1) Если $p > 1$, то возьмём производную функции $g(x) = yx - x^p$. $g'(x) = y - px^{p-1}$

$g'(x_m) = 0$ при $x_m = \sqrt[p-1]{\frac{y}{p}}$. Так как $p > 1$, то $g'(x) > 0$ при $x < x_m$ и $g'(x) < 0$ при $x > x_m$. Значит, в точке $x_m = \sqrt[p-1]{\frac{y}{p}}$ достигается супремум. $f^*(y) = y \sqrt[p-1]{\frac{y}{p}} - \left(\sqrt[p-1]{\frac{y}{p}}\right)^p = y \left(\frac{y}{p}\right)^{\frac{1}{p-1}} - \left(\frac{y}{p}\right)^{\frac{p}{p-1}}$.

2.2 Пункт 2

1) $g(\mathbf{x}) = f(\mathbf{x} + \mathbf{b})$

$$g^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x}\rangle - f(\mathbf{x}+\mathbf{b})\right) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x}+\mathbf{b}\rangle - \langle \mathbf{y}, \mathbf{b}\rangle - f(\mathbf{x}+\mathbf{b})\right) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x}+\mathbf{b}\rangle - f(\mathbf{x}+\mathbf{b})\right) - \langle \mathbf{y}, \mathbf{b}\rangle$$

Так как преобразование $\mathbf{x} \to \mathbf{x} + \mathbf{b}$ биекция, то $\sup\limits_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} + \mathbf{b}\rangle - f(\mathbf{x} + \mathbf{b})\right) - \langle \mathbf{y}, \mathbf{b}\rangle = \sup\limits_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x}\rangle - f(\mathbf{x})\right) - \langle \mathbf{y}, \mathbf{b}\rangle = f^*(\mathbf{y}) - \langle \mathbf{y}, \mathbf{b}\rangle$.

2) $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x})$, где $\mathbf{A}$ невырождена. Тогда

$$g^*(\mathbf{y}) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{A}\mathbf{x})\right) = \sup_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{A}_{-1}\mathbf{A}\mathbf{x} \rangle - f(\mathbf{A}\mathbf{x})\right) = \sup_{\mathbf{x}} \left(\langle \mathbf{A}^{-\top}\mathbf{y}, \mathbf{A}\mathbf{x} \rangle - f(\mathbf{A}\mathbf{x})\right) =$$

(преобразование $\mathbf{x} \to \mathbf{A}\mathbf{x}$ биекция)

$$= \sup_{\mathbf{x}} \left(\langle \mathbf{A}^{-\top}\mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})\right) = f^*(\mathbf{A}^{-\top}\mathbf{x})$$

3) Если же $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x})$, где $\mathbf{A}$ вырождена, то рассмотрим пространство X решений системы $\mathbf{A}\mathbf{x} = \not\vdash$. Если $\mathbf{y} \notin X^\perp$ ($X^\perp$ — оргогональное пространство), то $\exists \mathbf{x}_0 \in X$ такой что $\langle \mathbf{y}, \mathbf{x}_0 \rangle = p \neq 0$. Тогда если подставим $k\mathbf{x}_0$, получим $(\langle \mathbf{y}, k\mathbf{x}_0 \rangle - f(k\mathbf{A}\mathbf{x}_0) = kp - f(\mathbf{0})$. Тогда, устремляя $k \to +\infty$ или $k \to -\infty$, можно сделать так, что $kp - f(\mathbf{0}) \to +\infty$. И тогда $g^*(\mathbf{y}) = \sup\limits_{\mathbf{x}} \left(\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{A}\mathbf{x})\right) = +\infty$.

2.3 Пункт 3

$$f^*(\mathbf{y}) = \sup_{\mathbf{x}}(-\frac{1}{2}\mathbf{x}^\top\mathbf{A}\mathbf{x} + (\mathbf{y}^\top - \mathbf{b}^\top)\mathbf{x} - \mathbf{c})$$

Функция $g(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^\top\mathbf{A}\mathbf{x} + (\mathbf{y}^\top - \mathbf{b}^\top)\mathbf{x} - \mathbf{c}$ вогнута и $\frac{\partial^2 g}{\partial \mathbf{x}^2} = -\mathbf{A} \preceq 0$, поэтому максимум достигается в точках где $\frac{\partial g}{\partial \mathbf{x}} = \mathbf{0}$.

$$\frac{\partial g}{\partial \mathbf{x}} = -\mathbf{A}\mathbf{x} + (\mathbf{y} - \mathbf{b})$$

Если $\mathbf{A} \in \mathbf{S}^n_{++}$, то у матрицы $\mathbf{A}$ есть обратная и $\mathbf{x}_m = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})$.

$$f^*(\mathbf{y}) = -\frac{1}{2}\left(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})\right)^\top \mathbf{A}\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b}) + (\mathbf{y}-\mathbf{b})^\top\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b}) - \mathbf{c} = (\mathbf{y}-\mathbf{b})^\top(-\frac{1}{2}\mathbf{A}^{-\top} + \mathbf{A}^{-1})(\mathbf{y}-\mathbf{b}) - \mathbf{c}$$

Если $\mathbf{A} \in \mathbf{S}^n_{+}$, то матрица может быть и необратимой. Тогда если система $\mathbf{A}\mathbf{x} = (\mathbf{y} - \mathbf{b})$ имеет решения, то максимум достигается на любом из них. Покажем, что если решения нет, то максимума нет. Пусть $R(\mathbf{A})$ — образ $\mathbb{R}^n$ при преобразовании $\mathbf{A}$, а $R^\perp(\mathbf{A})$ — ортогональное к нему подпространство. Разложим вектор $\mathbf{y} - \mathbf{b} = \mathbf{u} + \mathbf{v}$, где $\mathbf{v} \in R(\mathbf{A})$, $\mathbf{u} \in R^\perp(\mathbf{A})$. Так как система не имеет решений, то $\mathbf{u} \neq \mathbf{0}$. Поскольку $\mathbf{A}\mathbf{u} \in R(\mathbf{A})$, $\mathbf{u} \in R^\perp(\mathbf{A})$, то $\mathbf{u}^\top\mathbf{A}\mathbf{u} = 0$. Поэтому $g(k\mathbf{u}) = -\frac{1}{2}k\mathbf{u}^\top\mathbf{A}k\mathbf{u} + (\mathbf{u} + \mathbf{v})^\top k\mathbf{u} - \mathbf{c} = k||\mathbf{u}||_2^2 - \mathbf{c} \to +\infty$ при $k \to +\infty$.

2.4 Пункт 4

$$f^*(\mathbf{y}) = \sup_{\mathbf{x}}(\mathbf{y}^\top\mathbf{x} - \max_{i=1,\dots,n} x_i)$$

Рассмотим такие $\mathbf{x}$, что $x_1 = \dots = x_n = m$ (m – просто некоторое число). Тогда $f^*(\mathbf{y}) = m\sum_{i=1}^n y_i - m = m(\sum y_i - 1)$. Поэтому если $\sum y_i \neq 1$, то устремляя $m \to +\infty$ или $m \to -\infty$, получим, что $\mathbf{y}^\top\mathbf{x} - \max\limits_{i=q,\dots,n} x_i \to +\infty$, то есть $f^*(\mathbf{y}) = +\infty$.

Если же $\sum y_i = 1$, то найдётся $y_j \neq 0$. Рассмотрим $\mathbf{x}$, такой что $x_j = m - 1$, а остальные компоненты равны m. Тогда $f^*(\mathbf{y}) = m\sum_{i=1}^n y_i - y_j - m = m((\sum y_i) - y_j - 1)$. Но теперь уже точно $(\sum y_i) - y_j - 1 \neq 0$, поэтому устремляя $m \to +\infty$ или $m \to -\infty$, получим, что $\mathbf{y}^\top\mathbf{x} - \max\limits_{i=q,\dots,n} x_i \to +\infty$, то есть $f^*(\mathbf{y}) = +\infty$.

Итак, $f^*(\mathbf{y}) = +\infty$.

3. Задача третья

3.1 Пункт 1

$$\min_{(x,y)} (x-3)^2 - (y-2)^2$$

$$\text{s.t.} \;\; y - x - 1 = 0$$

$$y + x - 3 \leq 0$$

Подставим $y = x + 1$ в целевую функцию и в неравенство и тем самым сведём задачу к одномерной:

$f(x) = (x-3)^2 - (x-1)^2 = -6x + 9 + 2x - 1 = -4x + 8$. $x + 1 + x - 3 = 2x - 2 \leq 0$, $h(x) = x - 1 \leq 0$. Задача примет вид:

$$\min_{x} -4x + 8$$

$$\text{s.t.} \; x - 1 \leq 0$$

1) $h(x, y) \leq 0$ — выпукло, функция f выпукла, поэтому условия ККТ будут достаточными.

2) Функция h линейна. Следовательно, выполнено условие регулярности.

3) Запишем условия из ККТ

$$\begin{cases} x - 1 \leq 0 & (1) \\ \mu(x-1) = 0 & (2) \\ \mu \geq 0 & (3) \\ L'(x, \mu) = -4 + \mu = 0 & (4) \end{cases} \tag{8}$$

Из последнего уравнения получаем $\mu = 4$. Это значение больше 0, поэтому условие (3) выполнено. Из (2) получаем $x = 1$ и при этом верно ограничение (1). Следовательно, Минимум достигается в точке $x = 1$. При этом $f(x) = -4 + 8 = 4$.

Значение y, при котором достигается условный минимум есть $y = x + 1 = 2$.

3.2 Пункт 2

Условие $|x_1 - 2x_2 + 3x_3| \leq 4$ равносильно системе двух неравенств

$$\begin{cases} x_1 - 2x_2 + 3x_3 \leq 4 \\ x_1 - 2x_2 + 3x_3 \geq -4 \end{cases} \tag{9}$$

Или

$$\begin{cases} h_1(\mathbf{x}) = x_1 - 2x_2 + 3x_3 - 4 \leq 0 \\ h_2(\mathbf{x}) = -x_1 + 2x_2 - 3x_3 - 4 \leq 0 \end{cases} \tag{10}$$

Заметим, что в силу линейности функций h_1 и h_2 выполнено условие регулярности.

Лагранжиан равен $L(\mathbf{x}) = x_1^2 + 2x_2^2 + x_3 + \mu_1(x_1 - 2x_2 + 3x_3 - 4) + \mu_2(-x_1 + 2x_2 - 3x_3 - 4)$.

Запишем условия ККТ:

$$
\begin{cases}
h_1(\mathbf{x}) = x_1 - 2x_2 + 3x_3 - 4 \leq 0 \\
h_2(\mathbf{x}) = -x_1 + 2x_2 - 3x_3 - 4 \leq 0 \\
\mu_1, \mu_2 \geq 0 \\
\mu_2(-x_1 + 2x_2 - 3x_3 - 4) = 0 \\
\mu_1(x_1 - 2x_2 + 3x_3 - 4) = 0 \\
2x_1 + (\mu_1 - \mu_2) = 0 \\
4x_2 - 2(\mu_1 - \mu_2) = 0 \\
1 + 3(\mu_1 - \mu_2) = 0
\end{cases} \tag{11}
$$

Последние 3 равенства есть условие того, что $L'_{\mathbf{x}} = \mathbf{0}$. Из них получаем

$$
\begin{cases}
\mu_1 - \mu_2 = -\frac{1}{3} \\
x_1 = \frac{1}{6} \\
x_2 = -\frac{1}{6}
\end{cases} \tag{12}
$$

Тогда можно упростить систему ККТ

$$
\begin{cases}
\frac{1}{2} + 3x_3 - 4 \leq 0 \\
-\frac{1}{2} - 3x_3 - 4 \leq 0 \\
\mu_1, \mu_2 \geq 0 \\
\mu_1 \left(\frac{1}{2} + 3x_3 - 4\right) = 0 \\
\mu_2 \left(-\frac{1}{2} - 3x_3 - 4\right) = 0 \\
\mu_1 - \mu_2 = -\frac{1}{3} \\
x_1 = \frac{1}{6} \\
x_2 = -\frac{1}{6}
\end{cases} \tag{13}
$$

Теперь будем разбирать случаи

1) $\mu_1 = 0$

Тогда $\mu_2 = \frac{1}{3} \neq 0 \Rightarrow \left(-\frac{1}{2} - 3x_3 - 4\right) = 0 \Rightarrow x_3 = -\frac{3}{2}$. При этом $\frac{1}{2} + 3x_3 - 4 < 0$. Значит, все условия ККТ выполнены и $\left(\frac{1}{6}, -\frac{1}{6}, -\frac{3}{2}\right)$ — стационарная точка.

Функция $f(\mathbf{x}) = x_1^2 + 2x_2^2 + x_3$ выпукла как сумма выпуклых функций. Также выпуклы функции h_1 и h_2. То есть, задача выпуклая и условие ККТ является достаточным и стационарная точка является точкой минимума. Значение функции в этой точке $f\left(\frac{1}{6}, -\frac{1}{6}, -\frac{3}{2}\right) = \frac{1}{36} + \frac{2}{36} - \frac{3}{2} = -\frac{17}{12}$.

2) $\mu_1 \neq 0$

Тогда

$$
\frac{1}{2} + 3x_3 - 4 = 0
$$

$$
\Downarrow
$$

$$
-\frac{1}{2} - 3x_3 - 4 \neq 0
$$

$$\Downarrow$$

$$\mu_2 = 0$$

$$\Downarrow$$

$$\mu_1 = -\frac{1}{3}$$

Но μ_1 должно быть неотрицательным. Значит, в этом случае условия ККТ не могут быть выполнены.

Стационарная точка всего одна, это точка минимума $\left(\frac{1}{6}, -\frac{1}{6}, -\frac{3}{2}\right)$.

3.3 Пункт 3

$$\min_{\mathbf{x}\in\mathbb{R}^2} (x_1 - 2)^2 + (x_2 + 1)^2$$

$$\text{s.t.} \max(x_1 + x_2 + 1, x_1 - x_2 - 1) + 1 \leq 0$$

$\max(x_1 + x_2 + 1, x_1 - x_2 - 1) = x_1 + \max(x_2 + 1, -(x_2 + x_1)) = x_1 + |x_2 + 1| \leq -1$. Поэтому это условие эквивалентно системе неравенств

$$\begin{cases} x_1 + x_2 + 2 \leq 0 & (1) \\ x_1 - x_2 \leq 0 & (2) \end{cases} \tag{14}$$

Поскольку функции ограничений линейны, то выполнено условие регулярности. Также $f(\mathbf{x}) = (x_1 - 2)^2 + (x_2 + 1)^2$ выпукла как сумма выпуклых функций, значит, задача выпукла.

Также в условия ККТ входят

$$\begin{cases} \mu_1, \mu_2 \geq 0 & (3) \\ \mu_1(x_1 + x_2 + 2) = 0 & (4) \\ \mu_2(x_1 - x_2) = 0 & (5) \end{cases} \tag{15}$$

Запишем условие $L'(\mathbf{x}, \mu_1, \mu_2) = \mathbf{0}$:

$$\begin{cases} 2x_1 + (\mu_1 + \mu_2) - 4 = 0 & (6) \\ 2x_2 + (\mu_1 - \mu_2) + 2 = 0 & (7) \end{cases} \tag{16}$$

Рассмотрим случаи

1) $x_1 = x_2 = -1$. Тогда выполнены условия $(1), (2), (4), (5)$. Подставим значения в уравнения $(6), (7)$.

$$\begin{cases} -6 + (\mu_1 + \mu_2) = 0 & (6) \\ (\mu_1 - \mu_2) = 0 & (7) \end{cases} \tag{17}$$

Отсюда $\mu_1 = \mu_2 = 3$. Все условия выполнены, то есть, точка $(-1, -1)$ стационарная. Так как задача выпукла, то это есть точка минимума. $f(-1, -1) = 9$.

2) $x_1 = x_2 \neq -1$. Тогда выполнены условия $(2), (5)$, но $x_1 + x_2 + 2 \neq 0$. Тогда $\mu_1 = 0$. Условия $(6), (7)$ примут вид

$$\begin{cases} 2x_1 + \mu_2 - 4 = 0 \quad (6) \\ 2x_1 - \mu_2 + 2 = 0 \quad (7) \end{cases} \tag{18}$$

Отсюда $x_1 = x_2 = \frac{1}{2}$. Но тогда $x_1 + x_2 + 2 > 0$ и не выполнено условие (1).

Теперь оставшиеся 2 случая, когда $x_1 \neq x_2$. Тогда чтобы было выполнено второе условие, надо чтобы $\mu_2 = 0$.

3) $\mu_2 = \mu_1 = 0$. Условия $(6), (7)$ примут вид

$$\begin{cases} 2x_1 - 4 = 0 \quad (6) \\ 2x_2 + 2 = 0 \quad (7) \end{cases} \tag{19}$$

Отсюда получаем $x_1 = 2,\ x_2 = -1$. Но, увы, $x_1 + x_2 + 2 > 0$.

4) $\mu_2 = 0$, $\mu_1 \neq 0$

Тогда $x_1 + x_2 + 2 = 0$ и из системы

$$\begin{cases} 2x_1 + (\mu_1 + \mu_2) - 4 = 0 \quad (6) \\ 2x_2 + (\mu_1 - \mu_2) + 2 = 0 \quad (7) \end{cases} \tag{20}$$

получаем $x_1 = \frac{1}{5},\ x_2 = -\frac{5}{2}$. Но не выполнено условие (1).

Итак, стационарная точка всего одна, $(-1, -1)$ и это точка минимума.

3.4 Пункт 4

3.5 Пункт 5

$$L(\mathbf{x}, \lambda) = -3x_1^2 + x_2^2 + 2x_3^2 + 2(x_1 + x_2 + x_3) + \lambda(x_1^2 + x_2^2 + x_3^2 - 1)$$

$$L'_{\mathbf{x}} = \mathbf{0}$$

$$\begin{cases} \frac{\partial L}{\partial x_1} = -6x_1 + 2 + 2\lambda x_1 = 0 \\ \frac{\partial L}{\partial x_1} = 2x_2 + 2 + 2\lambda x_2 = 0 \\ \frac{\partial L}{\partial x_1} = 4x_3 + 2 + 2\lambda x_3 = 0 \end{cases} \tag{21}$$

Отсюда получаем

$$\begin{cases} x_1 = \frac{1}{3-\lambda} \\ x_2 = -\frac{1}{1+\lambda} \\ x_3 = -\frac{1}{2+\lambda} \end{cases} \tag{22}$$

Подставим эти значения в уравнение $x_1^2 + x_2^2 + x_3^2 - 1 = 0$. Получим $(1+\lambda)^2(2+\lambda)^2 + (\lambda-3)^2(\lambda+1)^2 + (\lambda-3)^2(\lambda+2)^2 - (1+\lambda)^2(2+\lambda)^2(\lambda-3)^2 = 0$. С помощью Wolfram найдём действительные корни этого уравнения. Их всего 4. Подставим значения λ в выражения для x_1, x_2, x_3 и получим 4 стационарные точки. Условия ККТ для данной задачи являются необходимыми, значит, минимум нужно искать среди этих 4 точек. Для этого нужно просто подставить в выражение для целевой функции и найти, в какой точке значение наименьшее. В таблице ниже приведены значения функции при каждом из 4 значений λ.

λ	$f(\mathbf{x})$
-3.14929	4.64728181
0.223509	-1.13040306
1.8919	-1.59219431
4.03523	-5.36546345

Видим, что минимум достигается при $\lambda = 4.03523$, то есть в точке $\mathbf{x} \approx (-0.966, -0.197, -0.166)$. $f_{min} \approx -5.365$.

4. Задача четвёртая

4.1 Пункт 1

a)

1) $f(x) = x^2 + 1,\ h(x) = (x-1)(x-4)$ — выпуклые функции, так как это параболы, ветви которых смотрят вверх. Также $\exists x_0 = 2 :\ h(2) = -2 < 0$. Следовательно, выполнено условие регулярности и задача выпукла.

2)

$$L(x) = x^2 + 1 + \mu(x^2 - 5x + 4)$$

ККТ:

$$\begin{cases} \mu \geq 0 \\ \mu(x-1)(x-4) = 0 \\ (x-1)(x-4) \leq 0 \\ L'(x) = 2x + 2\mu x - 5\mu = 0 \end{cases} \tag{23}$$

Заметим, что условие $(x-1)(x-4) \leq 0$ равносильно условию $x \in [1, 4]$

Рассмотрим случай, когда $x = 1$. Тогда $L'(1) = 2 + 2\mu - 5\mu = 0 \Rightarrow \mu = \frac{2}{3}$. Все условия ККТ выполнены и задача выпуклая, значит, $x = 1$ — точка минимума. $p^* = f(1) = 2$.

b)

Нарисуем целевую функцию и выделим серым цветом допустимое множество (то есть где $x \in [1, 4]$). Красной точко отметим оптимальное значение. (см рис. 1 в конце файла).

c) Нарисуем графики лагранжиана для значений $\mu = 0, \frac{2}{3}, 1, 2$ (см рис. 2, 3, 4, 5 в конце файла).

d) $g(\mu) = \inf\limits_x (x^2 + 1 + \mu(x^2 - 5x + 4)) = \inf\limits_x ((\mu+1)x^2 - 5\mu x + (1 + 4\mu)$. Минимум квадратного трёхчлена достигается в точке $x^* = \frac{5\mu}{2(\mu+1)}$. Тогда $g(\mu) = L(x^*, \mu) = \frac{29}{4} - \frac{9}{4}\mu - \frac{25}{4(\mu+1)}$. Теперь мы можем сформулировать двойственную задачу:

$$\max_{\mu} \frac{29}{4} - \frac{9}{4}\mu - \frac{25}{4(\mu+1)}$$

$$\text{s.t. } \mu \geq 0$$

Решим её.

$g'(\mu) = -\frac{9}{4} + \frac{25}{4(\mu+1)^2} = 0$. У этого уравнения один неотрицательный корень $\mu^* = \frac{2}{3}$. При этом $g'(\mu) > 0$ при $0 \leq \mu < \mu^*$, и $g'(\mu) < 0$ при $\mu > \mu^*$. Значит, μ^* — точка максимума.

$d^* = g(\mu^*) = 2 = p^*$. Зазор двойственности равен $p^* - d^* = 0$.

4.2 Пункт 2

Пусть дан орграф без циклов, в котором вершины s и t являются, соответственно источником и стоком. Обозначим E — множество ориентированных рёбер в графе, V — множество вершин, $k_{ij} \in \mathbb{N}$, $(i,j) \in E$ — максимальная пропускная способность ребра (i,j), $x_{ij} \in \mathbb{N}_0$ — поток, который мы пускаем по ребру (i,j).

Сформулируем задачу о максимальном потоке:

$$\max_{v \in \mathbb{R},\ x_{ij} \in \mathbb{N}_0} v$$

$$\text{s.t.} \sum_{j:\ (s,j) \in E} x_{sj} - v = 0$$

$$v - \sum_{j:\ (j,t) \in E} x_{jt} = 0$$

$$\sum_{j:\ (j,h) \in E} x_{jh} - \sum_{j:\ (h,j) \in E} x_{hj} = 0, \quad h \in V \backslash \{s,t\}$$

$$x_{ij} \le k_{ij}, \quad (i,j) \in E$$

Теперь попробуем сформулировать двойственную задачу ($\max\limits_{v \in \mathbb{R},\ x_{ij} \in \mathbb{N}_0} v = \min\limits_{v \in \mathbb{R},\ x_{ij} \in \mathbb{N}_0} -v$).

$$L = -v + u_s \left(\sum_{j:\ (s,j) \in E} x_{sj} - v \right) + u_t \left(v - \sum_{j:\ (j,t) \in E} x_{jt} \right) +$$

$$+ \sum_{h \in V \backslash \{s,t\}} u_h \left(\sum_{j:\ (h,j) \in E} x_{hj} - \sum_{j:\ (j,h) \in E} x_{jh} \right) + \sum_{(i,j) \in E} y_{ij}(x_{ij} - k_{ij})$$

Здесь $u_i,\ i \in V,\ y_{ij},\ (i,j) \in E$ переменные Лагранжа. $g(\mathbf{u}, Y)$ есть инфимум по всем x_{ij} и v. v учатсвует в Лагранжиане в слагаемых $-v + vu_t - vu_s = v(u_t - u_s - 1)$. Если $u_t - u_s - 1 \neq 0$, то устремив $v \to +\infty$ или $v \to -\infty$, и взяв $x_{ij} = 0$ получим, что лагранжиан устремится в $-\infty$ и $g(\mathbf{u}, Y) = -\infty$. Так как в двойственной задаче мы ищем максимальное значение g, то такие точки, где $g(\mathbf{u}, Y) = -\infty$ нам не интересны, поэтому примем $(u_t - u_s - 1) = 0$. x_{ij} участвует в слагаемых $(u_i - u_j + y_{ij})x_{ij}$. Если $u_i - u_j + y_{ij} < 0$, то устремив $x_{ij} \to +\infty$, получим, что $L \to -\infty$. Это нам не подходит. Если же $\forall (i,j) \in E\ \ u_i - u_j + y_{ij} \ge 0$, то инфимум достигается при $x_{ij} = 0$ и равен $\left(-\sum\limits_{(i,j) \in E} y_{ij} k_{ij} \right)$. Теперь можем записать двойственную задачу:

$$\max_{\mathbf{u}, Y} - \sum_{(i,j) \in E} y_{ij} k_{ij}$$

$$\text{s.t.} \ \ u_t - u_s - 1 = 0$$

$$u_i - u_j + y_{ij} \ge 0, \quad (i,j) \in E$$

$$y_{ij} \geq 0$$

Последнее ограничение взялось так как y_{ij} — множители для ограничений типа неравенств.

Рассмотрим набор значений u_i и y_{ij}, который удовлетворяет условиям двойственной задаче и, вдобавок, $u_i,\ y_{ij} \in \{0, 1\}$. Покажем, что таким способом можно задать разрез.

Вспомним, как задаётся разрез. Множество вершин делаится на 2 части: S и $\overline{S}$ (В нашем случае пусть S содержит сток, вершину t, а $\overline{S}$ — соответственно, источник, вершину s). Вес разреза есть сумма пропускный способностей всех рёбер, которые соединяют какуб-либо вершину из $\overline{S}$ с вершиной из S.

Пусть $u_i = 1$, если вершина i принадлежит множеству S, а иначе $u_i = 0$. Тогда из равенства $u_t - u_s - 1 = 0$ получаем, что $u_s = 0,\ u_t = 1$, то есть $s \in \overline{S},\ t \in S$. Если ребро принадлежит разрезу, то $u_i - u_j = -1$, и, чтобы было выполнено неравенство $u_i - u_j + y_{ij} \geq 0$, необходимо, чтобы $y_{ij} = 1$. Если же ребро не принадлежит разрезу, то $u_i - u_j \geq 0$ и так как нам нужно минимизировать сумму $\sum\limits_{(i,j)\in E} y_{ij}k_{ij}$, можно приравнять $y_{ij} = 0$ и тогда условие на y_{ij} будет выполнено и сумма будет меньше, чем при $y_{ij} = 1$. Таким образом, $y_{ij} = 1$, если ребро (i, j) принадлежит разрезу, то есть, $j \in S,\ i \notin S$, и $y_{ij} = 0$ иначе.

Поэтому если наложить ограничения $u_i,\ y_{ij} \in \{0, 1\}$, то решение двойственной задачи как раз и даст нам минимальный разрез и его вес. По теореме Форда-Фалкерсона (знаем из курса "Алгоритмы и модели вычислений") максимальный поток равен весу минимального разреза. То есть выполнена сильная двойственность.

4.3 Пункт 3

4.4 Пункт 4

$$\min_{x\in\mathbb{R}, y\in\mathbb{R}_{++}} e^{-x}$$

$$\text{s.t. } \frac{x^2}{y} \leq 0$$

Функция $h(x, y) = \frac{x^2}{y}$ является выпуклой на множестве $\mathbb{R} \times \mathbb{R}_{++}$, так как

$$f''(x, y) = \begin{pmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & \frac{2x^2}{y^3} \end{pmatrix} = \frac{2}{y} \begin{pmatrix} 1 & -\frac{x}{y} \\ -\frac{x}{y} & \frac{x^2}{y^2} \end{pmatrix} \succeq 0 \tag{24}$$

$f(x, y) = e^{-x}$ также выпукла. Поэтому, задача является выпуклой.

Найдём решение.

На множестве $\mathbb{R} \times \mathbb{R}_{++}$ неравенству $\frac{x^2}{y} \leq 0$ удовлетворяет лишь точки, где $x = 0$. Поэтому f может принимать лишь одно допустимое значение $f(0, y) = e^0 = 1$. Получаем $p^* = 1$.

Построим двойственную задачу.

$L(x, y, \mu) = e^{-x} + \mu\frac{x^2}{y}$, $\mu \geq 0$.

$$g(\mu) = \inf_{x\in\mathbb{R}, y\in\mathbb{R}_{++}} e^{-x} + \mu\frac{x^2}{y}$$

$\forall \varepsilon > 0$ мы можем выбрать такое $x_0 > -\ln \varepsilon$, что будет выполнено $e^{-x_0} < \varepsilon$. Также мы можем выбрать такое $y_0 > \mu \frac{x_0^2}{\varepsilon}$, что будет верно $\mu \frac{x^2}{y} < \varepsilon$. Поэтому $\forall \varepsilon > 0 \; \exists (x_0, y_0)$: $L(x_0, y_0, \mu) < 2\varepsilon$. А так как, очевидно, при $\mu \geq 0, \; x \in \mathbb{R}, \; y \in \mathbb{R}_{++}$ всегда $L(x, y, \mu) > 0$, то $\inf\limits_{x \in \mathbb{R}, y \in \mathbb{R}_{++}} e^{-x} + \mu \frac{x^2}{y} = 0$, то есть, $g(\mu) = 0$.

Задача двойственности будет выглядеть так:

$$\max_{\mu} \; 0$$

$$\text{s.t.} \;\; \mu \geq 0$$

Целевая функция здесь не зависит от μ, и решение задачи $d^* = 0$.

Зазор двойственности равен $p^* - d^* = 1$, а значит сильной двойственности нету, хотя исходная задача выпуклая.

Дело в том, что для исходной задачи не выполнено условие Слейтера, то есть, в заданной области $\mathbb{R} \times \mathbb{R}_{++}$ не существует точек, для которых выполнено строгое неравенство $\frac{x^2}{y} < 0$. А при невыполнении условия Слейтера сильной двойственности может и не быть даже для выпуклой исходной задачи.

Рис. 1: Целевая функция и допустимое множество

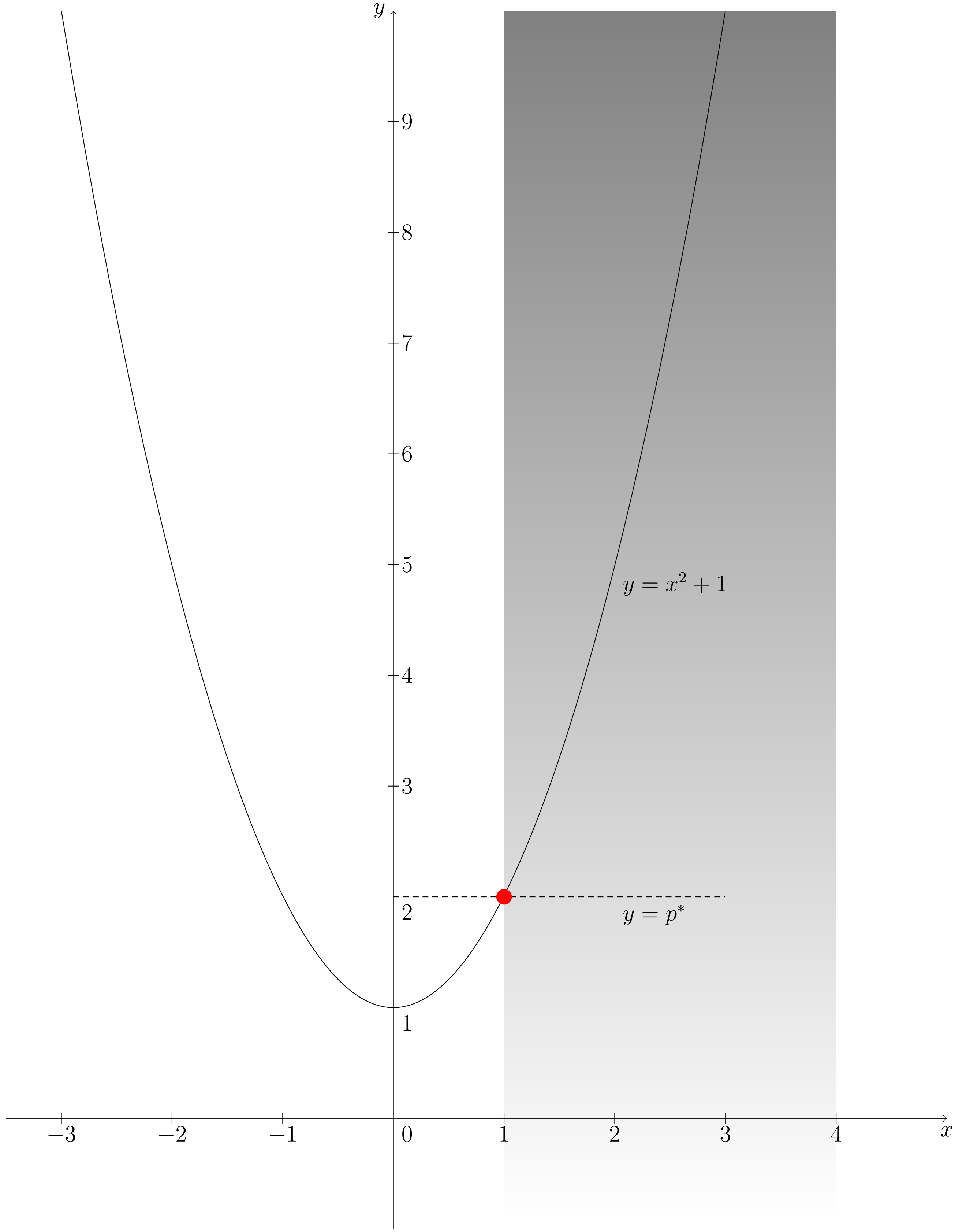

Рис. 2: $\mu = 0,\ L(x,0) = x^2 + 1$

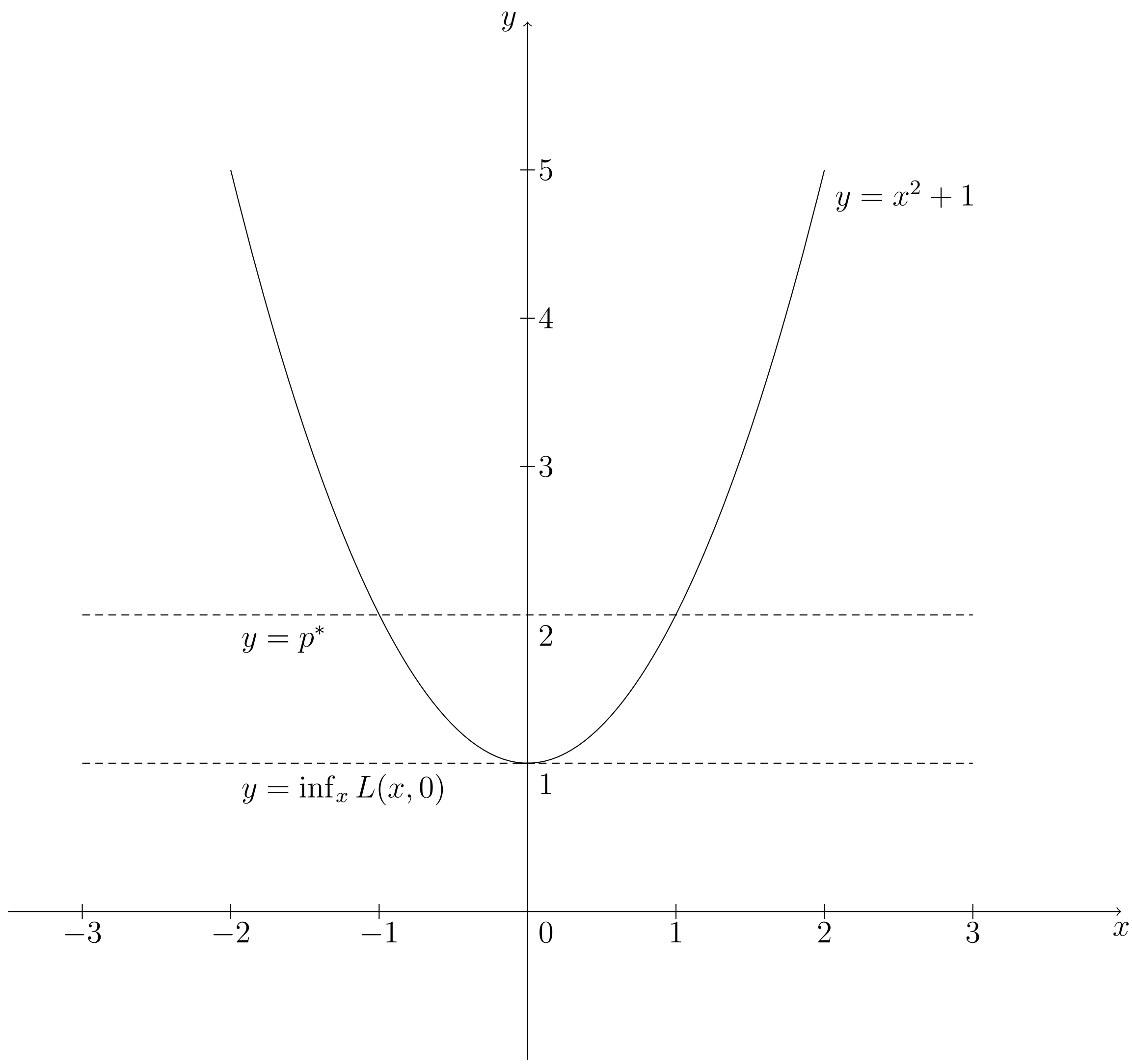

Рис. 3: $\mu = \frac{2}{3},\ L(x,\mu) = \frac{5}{3}x^2 - \frac{10}{3}x + \frac{11}{3}$

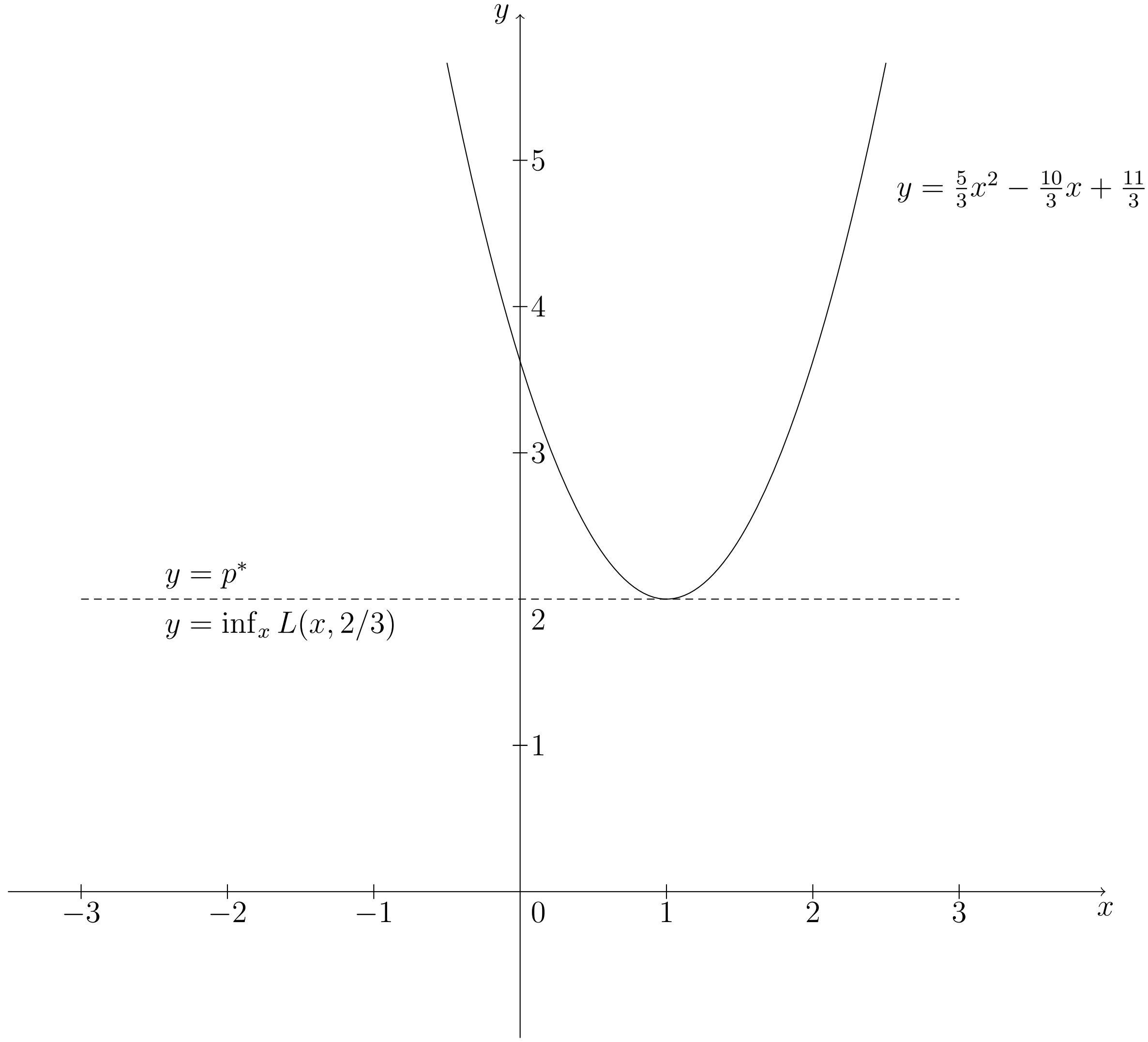

Рис. 4: $\mu = 1,\ L(x,1) = 2x^2 - 5x + 5$

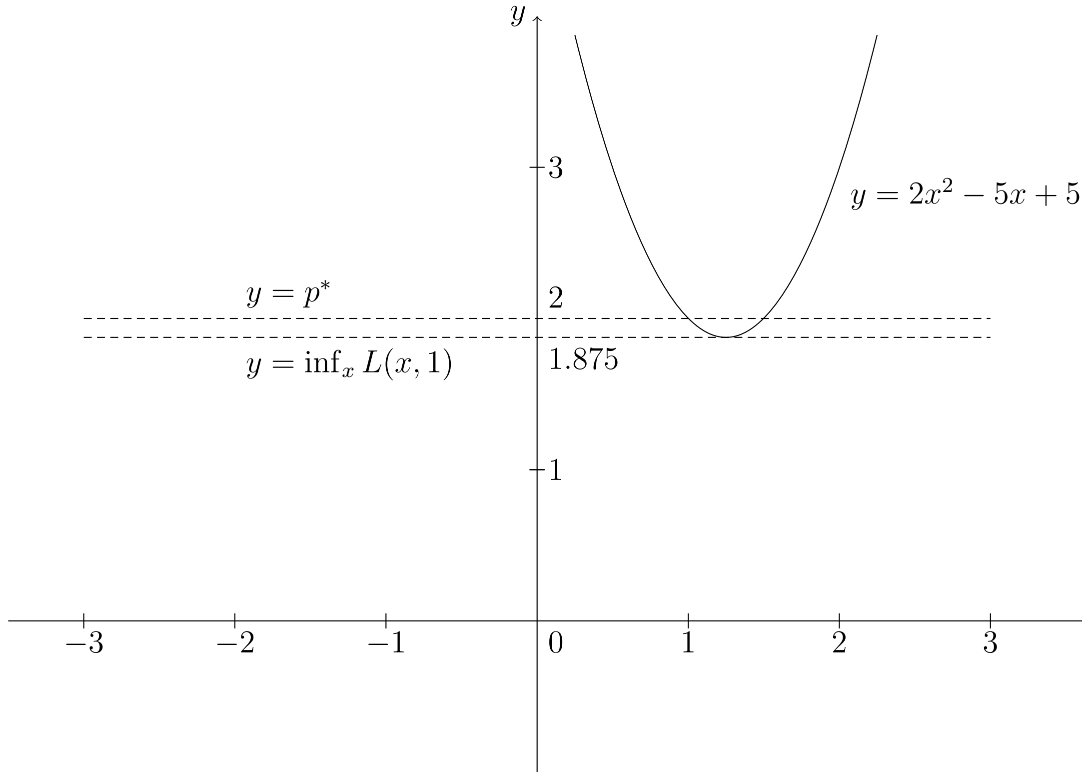

Рис. 5: $\mu = 2,\ L(x,2) = 3x^2 - 10x + 13$