Второе задание.

Conjugate sets

№1

$$x \in S \iff x = (c, d) = \theta_1(-3, 1) + \theta_2(2, 3) + \theta_3(4, 5) \forall \theta_i \geq 0$$

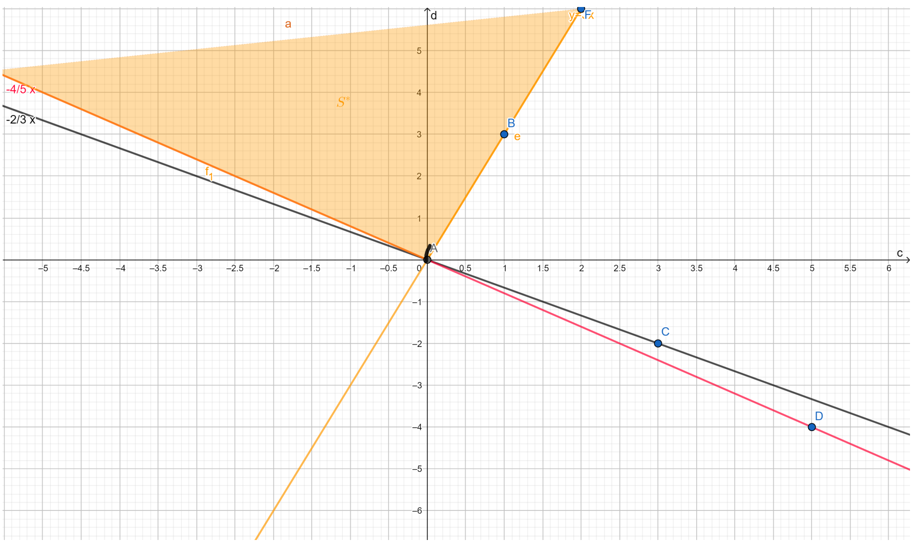
$$y = (c, d) \in S^* \iff (x, y) \geq 0 \forall x \in S$$

$$\iff \theta_1(-3c + d) + \theta_2(2c + 3d) + \theta_3(4c + 5d) \geq 0, \forall \theta_i$$

$$\iff d \geq 3c, d \geq -\frac{2}{3}c, d \geq -\frac{4}{5}c$$

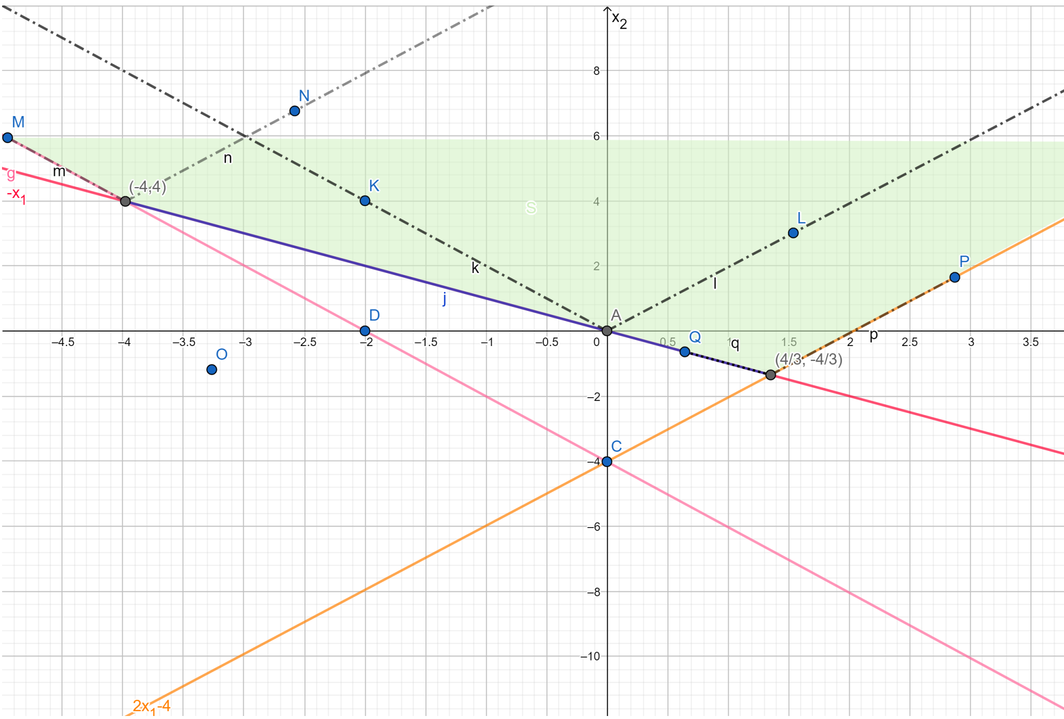
$$S^* = \{y = (c, d) | d \geq 3c, d \geq -\frac{2}{3}c, d \geq -\frac{4}{5}c\}$$

S^* область между лучами $d = 3x$ и $d = -4/5c$, проведенными из нуля (на рисунке заштрихована оранжевым)



№2

Предствим исходное множесвто, как (на рисунке показано как это можно сделать, зеленая област распрастраняется и выше(не ограничена)):



$$S = conv\left\{(-4,4),(\frac{4}{3},-\frac{4}{3})\right\} + cone\{(-1,2,1,2)$$

Тогда по теореме: $S^* = \{(y,x)|2y-x \geq 0, 2y+x \geq 0, y \leq x+\frac{3}{4}, y \geq x-\frac{1}{4}\}$
S, S^* выпуклы, замкнуты и содержат нуль. $\Rightarrow S^{**} = S, S^{***} = S^*$

№3

Пусть A, B положительно полуопределнные матрицы ($\in S^n_+$)), тогда их можно представить как $A = U^T U, B = V^T V$. Тогда $\langle A, B\rangle = tr(U^T U V^T V)$

$||UV^T||_F^2 \geq 0$. Если A положительно определена ($\in S^n_{++}$)), то она и положительная полуопределена, т.е. : $(A, B) \geq 0 \forall B \in S^n_+$.

$$\Rightarrow S^n_+ \subset (S^n_{++})^*, S^n_+ \subset (S^n_+)^*$$

Пусть $C \notin S^n_+, C \in (S^n_{++})^*$. Тогда: $\exists x \equiv 0 \in \mathbb{R}^n : x^T C x < 0, \Rightarrow (x^T x, C) \leq 0$, где $X = x^T x \in S^n_{++}$. Но это значит, что $C \notin (S^n_{++})^*$. Противоречие, значит такого C не существует. Т. е. $(S^n_{++})^* = S^n_+$. Т. к. $S^n_{++} \subset S^n_+$, то $S^n_+ \subset (S^n_+)^* \subset (S^n_{++})^* = S^n_+$. $\Rightarrow (S^n_+)^* = S^n_+$.
Ответ: $(S^n_{++})^* = (S^n_+)^* = S^n_+$

№4

$$S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_2 > 0, x_1 \geq x_2 e^{x_3/x_2}\}$$

Покажем, что $S^* = E = \{(y_1, y_2, y_3) \in \mathbb{R}^3 | y_3 < 0, y_1 \geq -y_3 e^{y_2/y_3 - 1}\} \cup \{(y_1, y_2, 0) \in \mathbb{R}^3 | y_1, y_2 \geq 0\}$:
Пусть $y \in E, y_3 > 0$, тогда:

$$\forall x \in S \to (y, x) = y_1 x_1 + y_2 x_2 + y_3 x_3 \geq -x_2 e^{x_3/x_2} y_3 e^{y_2/y_3 - 1} + y_2 x_2 + y_3 x_3 =$$

$$= -x_2 y_3 e^{\frac{x_3 y_3 + x_2 y_2}{x_2 y_3} - 1} + y_2 x_2 + y_3 x_3 \underbrace{\geq}_{\text{т.к.} e^x \geq 1 + x} -x_2 y_3 \frac{x_3 y_3 + x_2 y_2}{x_2 y_3} + + y_2 x_2 + y_3 x_3 = 0$$

Пусть теперь $y \in E, y_3 = 0$:

$$\forall x \in S \to (y, x) = y_1 x_1 + y_2 x_2 \geq x_2 e^{x_3/x_2} y_1 + y_2 x_2 \geq 0,$$

т.к. $y_2, y_1, x_2 \geq 0$.

$$\Rightarrow E \subset S^*.$$

Рассмотрим призвольный $y : y_1 < -y_3 e^{y_2/y_3 - 1}$, т. е. $y_3 = -y_3 e^{y_2/y_3 - 1} - C, C > 0$.. Тогда $\exists x \in S : x_2 > 0, x_3 = \frac{y_3 - y_2}{y_3} x_2, x_1 = x_2 e^{x_3/x_2} = x_2 e^{\frac{y_3 - y_2}{y_3}}$ и:

$$(y, x) = -x_2 e^{\frac{y_3 - y_2}{y_3}} y_3 e^{y_2/y_3 - 1} + y_2 x_2 + y_3 x_2 \frac{y_3 - y_2}{y_3} - C x_2 e^{x_3/x_2} =$$

$$= -x_2 y_3 + x_2 y_3 - \underbrace{C x_2 e^{x_3/x_2}}_{<0} < 0, \Rightarrow y \notin S^*$$

$$\Rightarrow y \in S^*, y_3 \equiv 0 \longrightarrow y_1 \geq -y_3 e^{y_2/y_3 - 1}$$

Рассмотрим призвольный $y : y_3 = 0, y_1 < 0, y_2 < 0$. Тогда возмем $x \in S : x_1 = x_2 e^{x_3/x_2}, x_2 > 0, x_3 = 0 :$

$$(y, x) = x_2 e^{0/x_2} y_1 + y_2 x_2 = \underbrace{x_2}_{>0} \underbrace{(y_1 + y_2)}_{<0} < 0. \Rightarrow y \notin S^*$$

Рассмотрим призвольный $y : y_3 = 0, y_2 < 0, y_1 > 0$. Тогда возмем $x \in S : x_1 = x_2 e^{x_3/x_2}, x_2 > 0, x_3 = x_2 \ln(\frac{-y_2}{2y_1}) :$

$$(y, x) = x_2 e^{\ln(\frac{-2y_2}{y_1})} y_1 + y_2 x_2 = x_2(-y_2/2 + y_2) = \underbrace{x_2}_{>0} (y_2/2) < 0. \Rightarrow y \notin S^*$$

Аналогично, при $y_3 = 0, y_1 < 0. \Rightarrow y \in S^*, y_3 = 0 \longrightarrow y_1, y_2 \geq 0$
Пусть теперь $y : y_3 = C, C < 0, y_1 \geq -y_3 e^{y_2/y_3 - 1}$ Возьмем $x \in S, x_1 = x_1 = x_2 e^{x_3/x_2}, x_3 = x_2 t$. Тогда:

$$(y, x) = x_2 e^{x_3/x_2} y_1 + y_2 x_2 + C x_3 = \underbrace{x_2}_{>0} (e^t y_1 + y_2 + \underbrace{C}_{<0} t) < 0$$

$$\Rightarrow S^* \subset E$$

$$\Rightarrow S^* = E$$

№5

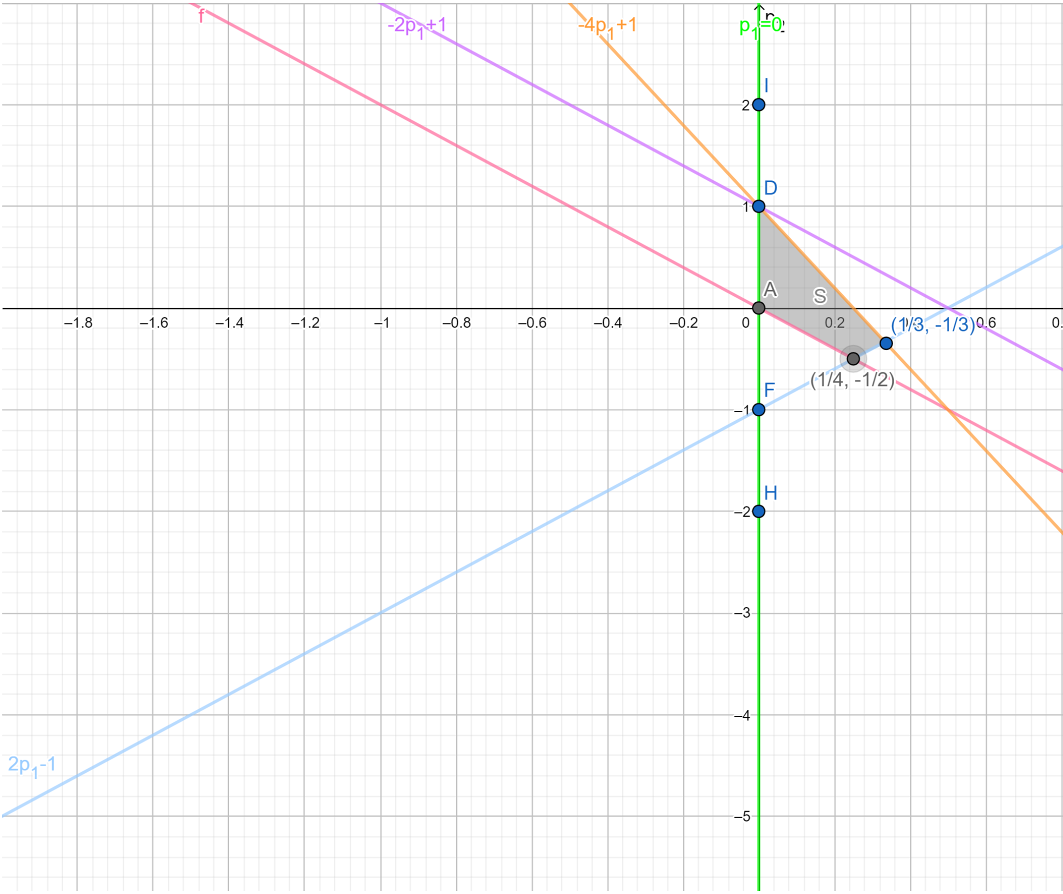
По теореме сопряженным к данному множеству является следующий полиэдр:

$$S^* = \{p \in \mathbb{R}^n \Big| -4p_1 - p_2 \geq -1, -2p_1 - p_2 \geq -1,$$

$$-2p_1 + p_2 \geq -1, p_1 \geq 0, 2p_1 + p_2 \geq 0\} =$$

$$= conv\{(0,0), (0,1), (1/3, -1/3), (1/4, -1/2)\}$$

На рисунке S^* заштриховано серым.



№6

Пусть S самосопряженный. Тогда $\forall x, y \in S (x, y) \leq 1.(1)$ А $\forall z \notin S \to \exists x \in S : (x, z) > 1.(2)$. Предположим, что $\exists z \in S : (z, z) > 1$, это противоречит (1). Значит, $S \subset B(0, 1)$. Пусть, $S \equiv B(0, 1)$. Тогда $\exists z \notin S, x \in B(0, 1)$. Значит из (2) $\exists x \in S : (x, z) > 1$. Но из неравенства Коши-Буняковского: $|(x, z)| \leq \sqrt{(x, x), (z, z)} \leq 1$, т.к. $x, z \in B(0, 1)$. Противоречие. $\Rightarrow, S = B(0, 1)$

Рассмотрим $S = B(0, 1)$ - единичный шар с центром в нуле. $x \in S \Leftrightarrow (x, x) \leq 1$. Из неравенства Коши-Буняковского $\forall x, y \in S$: $|(x, y)| \leq \sqrt{(x, x)(y, y)} \leq 1. \Rightarrow (x, y) \leq 1$. Значит, $S \subset S^*$. Пусть $z : z \notin S$

$(\Leftrightarrow (z,z) > 1), z \in S^*$. Тогда возьмем $y \in S$, коллинеарное zи $||y|| = 1$, (такое точно найдется так как в шаре содержатся вектора всевозможных направлений и длин до единицы включительно). В этом случае: $|(z,y)| = \sqrt{(z,z)(y,y)} > 1 \cdot 1 = 1$. Противоречие, с тем что $z \in S^*$. Значит такого z не существует. $\Rightarrow S = S^*$, т.е. единичнй шар самосопряжен.

Таким образом, самосопряженным является только единицный шар с центром в нуле.

№7

Покажем, что $S^* = P = \{p \in \mathbb{R}^n \middle| \sum_{i=1}^n \frac{p_i^2}{a_i^2} \leq \frac{1}{\varepsilon^2}\}$:

1) Пусть $p \in P$, тогда с помощью неравенства Коши-Буняковского получим $\forall x \in S \to |(x,p)|^2 = |\sum_{i=1}^n x_i p_i|^2 = |\sum_{i=1}^n (x_i a_i)(\frac{p_i}{a_i})|^2 \leq$
$\leq (\sum_{i=1}^n (x_i a_i)^2)(\sum_{i=1}^n (\frac{p_i}{a_i})^2) \leq \varepsilon^2 \cdot \frac{1}{\varepsilon^2} = 1$, т. к. $p \in P, x \in S$.
$\Rightarrow \forall p \in P, x \in S \to (x,p) \geq -1$. Значит, $P \subset S^*$.

2) Предположим, что $\exists p \notin P, p \in S^*$. $\Rightarrow \sum_{i=1}^n \frac{p_i^2}{a_i^2} = \frac{1}{\varepsilon^2} + C$, $C > 0$.. Заметим, что т. к. эллипсоид симметричен относительно начала координат, то $p \in S^* \Leftrightarrow \forall y \in S \to (y,p) \geq -1 \Leftrightarrow \forall y \in S \to (y,p) \leq 1$, (1). Возьмем $x \in S$, такой что x коллинеарно p (компоненты пропорционально) и $\sum_{i=1}^n a_i^2 x_i^2 = \varepsilon^2$ (это возможно сделать: выражаем черех x_1 и компоненты p все x_1 и подставляем в условие на сумму квадратов компонет, оттуда находим x_1, а затем через него остальные компоненты). В случае коллинеарных векторов нерво Коши-Буняковского обращается в равенство, т.е. $|(x,p)|^2 = |\sum_{i=1}^n x_i p_i|^2 = |\sum_{i=1}^n (x_i a_i)(\frac{p_i}{a_i})|^2 =$
$= (\sum_{i=1}^n (x_i a_i)^2)(\sum_{i=1}^n (\frac{p_i}{a_i})^2) = (\frac{1}{\varepsilon^2} + C)\varepsilon^2 = 1 + C\varepsilon^2 > 1$. Но это противоречит (1). Значит, такого p не существует.

$$\Rightarrow S^* = P == \{p \in \mathbb{R}^n \middle| \sum_{i=1}^n \frac{p_i^2}{a_i^2} \leq \frac{1}{\varepsilon^2}\}$$

Conjugate function

№1

$f*(y) = \sup_{x \in domf}((y,x) - f(x)) = \sup_{x \in domf} yx + \frac{1}{x}$ При $y > 0$ функция $yx + \frac{1}{x}$ неограничена сверху: можем взять $x = +\infty$, при этом дробь устремится к нулю, а первое слагаемое к бесконечности. Аналогично, при $y < 0$ можно взять $x = \infty$, и устремить выражение в бесконечность. При $y = 0$ функция также неограничена: при x стремящемся к нулю, она стремится к бесконечности. Значит $domf^* = \{\emptyset\}$. Таким образом, сопряженная функция нигде не опеделена.

№2

$$f(x) = -0,5 - log x$$

$$f^*(y) = \sup_{x \in domf}((y,x) - f(x)) = \sup_{x \in domf} yx + 0.5 + \log x$$

Эта функция не оганичена сверху при $y \geq \infty$. Значит, $domf^* = \{y < 0\}$ Максимум при $x = -1/y$, $f^*(y) = -0.5 - log(-y)$
Ответ: $f^*(y) = -0.5 - log(-y)$

№3

$$f(x) = \ln \sum_{i=1}^{n} e^{x_i}$$

Рассмотрим функцию $g(x,y) = yx - \ln \sum_{i=1}^{n} e^{x_i}$. Если какие-то компоненты y равны нулю (пусть для определенности $y_1, \dots, y_k$), то функция уходит в бесконечность при $x_1 = \dots = x_k \to -\infty$, а при $k+1 \leq m \leq n : x_m = \ln y_m + \ln \sum_{i=k+1}^{n} e^{x_i}$.
Если какая-то компонента $y_i < 0$, то при $x_k = 0 \forall k \equiv i, x_i \to -\infty$, то $g(x,y) = ykx_i - \ln(n - 1 + e^{x^i}) \to \infty - \ln n \to \infty$ Таким образом, $y > 0$. Если $y > 0$, и $1^T y \equiv 1$, то возьмем $x = (C, \dots, C)$. Тогда, $g(x,y) = C \cdot 1^T y - C - \ln n = C(1^T y - 1) - \ln n \to \infty$ (при $C \to \infty$, если $1^T y > 0$ и $C \to -\infty$, если $1^T y < 0$). $\Rightarrow 1^T y = 1$ Таким образом,

$\mathrm{dom} f^* = \{y \in \mathbb{R}^{\varkappa} \Big| y > 0, 1^T y = 1\}$. Найдем градиент:

$$\nabla_i = y_i - \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}}$$

Тогда максимум при $y_i = \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}}$. Значит,

$$f^*(y) = \frac{\sum_{i=1}^{n} x_i e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} - \ln(\sum_{i=1}^{n} e^{x_i}) = \sum_{i=1}^{n} \frac{(\ln e^{x_i}) e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} - (\sum_{i=1}^{n} e^{x_i}) \frac{\ln(\sum_{j=1}^{n} e^{x_j})}{\sum_{j=1}^{n} e^{x_j}} =$$

$$= \sum_{i=1}^{n} \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \left(\ln e^{x_i} - \ln(\sum_{j=1}^{n} e^{x_j}) \right) = \sum_{i=1}^{n} \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \ln \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} = \sum_{i=1}^{n} y_i \ln y_i$$

№4

$$f(x) = -(a^2 - x^2)^{\frac{1}{2}}$$

$$g(x, y) = yx + (a^2 - x^2)^{\frac{1}{2}}$$

Т.к. x ограничена, то $g(x, y)$ ограничена при $\forall y$. Т.е. $\mathrm{dom} f^* = \mathbb{R}$. Так как x определена на компакте, а $g(x, y)$ непрерывна по x на нем, то при фиксированном y эта функция достигает своего максимума. При $x^2 \equiv a^2$ существует производная:

$$y - \frac{x}{(a^2 - x^2)^{\frac{1}{2}}} = 0 \Leftrightarrow y = \frac{x}{(a^2 - x^2)^{\frac{1}{2}}} = 0 \Leftrightarrow |x| = \frac{ay}{(1 + y^2)^{\frac{1}{2}}}$$

Тогда, $f^*(y) = sup_{|x| \leq a} g(x, y) = \max\{sup_{|x| < a} g(x, y), g(a, y)\} =$

$$= \max\{g(|x|, y), g(-|x|, y), ya\}$$

$g(x, y) = \frac{ay^2}{(1+y^2)^{1/2}} + (a^2 - \frac{a^2 y^2}{(1+y^2)})^{1/2} = a(1 + y^2)^{1/2} > ay$

$$\Rightarrow f^*(y) = a(1 + y^2)^{1/2}.$$

№5

$$f(X) = -\ln\det X$$
$$g(X, Y) = \mathrm{tr}(XY) + \ln\det X$$

Предположим, что $\exists Z X > 0 : (Z, Y) = c_0 \geq 0$. Тогда возьмем $X = CZ, C > 0$, тогда $X \in S^n_{++}$. $D = \det X > 0$.

$$g(X, Y) = C(Y, Z) + \ln\det CX = Cc_0 + \ln(C^n D) = Cc_0 + n\ln C + \ln D \overbrace{\rightarrow}^{C\to\infty} \infty$$

Таким образом, чтобы функция была ограничена сверху необходимо, чтоюы такого Z не существовало, т. е. Y была отрицательна определена. Почситаем градиент (в первом дз он уже был посчитан для $\ln\det X$) и полуучим уравнение для максимума:

$$Y + X^T = 0 \Leftrightarrow X = (-Y)^{-1}$$

Как видно, отрицательная определенность Y является достаточным условвием, чтоб это уровнение решалось относительно X. Т. е. $\mathrm{dom} f^* = \{Y \in S^n_{-,-}\}$. Подставим найденное значение в $f^*(Y)$:

$$f^*(Y) = (Y, Y^{-1}) + \ln\det(-Y)^{-1} = n - \ln\det(-Y)$$

№6

$$f(x) = g(Ax)$$
$$f(A^{-1}x) = g(x) \Rightarrow \mathrm{dom} f = \{x_1 = A^{-1}x_2 | x_2 \in \mathrm{dom} g\}$$
$$f^*(y) = \sup_{x\in \mathrm{dom} g} (y^T A^{-1} x - g(x)) = \sup_{x\in \mathrm{dom} g} ((\underbrace{A^{-T}y}_{=t}, x) - g(x)) =$$
$$= \sup_{x\in \mathrm{dom} g} ((t, x) - g(x)) = g^*(t) = g^*(A^{-T}y)$$

Subgradient and Subdifferential

№1

Пусть $0 \in \partial f(x_0)$. Из определения: $\forall x \in S \to f(x) \geq f(x_0) + (0, x - x_0) = f(x_0)$. Значит x_0 точка минимума данной функции.
Пусть x_0 точка минимума функции, любой минимум выпуклой функции является глобальным. Значит, $\forall x \in S \to f(x) \geq f(x_0) = f(x_0) + 0 = f(x_0) + (0, x - x_0)$. $\Rightarrow 0 \in \partial f(x_0)$.

№2

$$f(x) = \max\{0, x\}$$

Из Теоремы Дубовицкого-Милютина:

$$\partial f(x) = \begin{cases} 0, x < 0 \\ conv(0, 1) = [0, 1], x = 0 \\ 1, x > 0 \end{cases}$$

Если в n-мерном пространстве, то вместо $[0, 1]$: $[0, 1]^n$.

№3

$p = 1)$ $||x_1|| = \max_{s_i = \{-1, 1\}} s^T x = \max_k g_k(x)$. Каждая $g_k(x)$ однозначно определяется соответствующем набором коэффициентов (вектором s). Тогда по теореме Дубовицкого-Милютина:

$$\partial f = conv\Big(\cup_{k \in I(x_0)} \partial g_k(x) \Big)$$

В свою очередь,

$$\partial g(x) = \partial(max\{s^T x, -s^T x\}) = \begin{cases} -s, s^T x < 0 \\ conv(-s, s), s^T x = 0 \\ s, s^T x > 0 \end{cases}$$

Правило выбора «активной» функции поточечного максимума в каждой точке следующее: если координата точки отрицательна, $s_i = -1$.

Если координата точки положительна, $s_i = 1$. Если координаты имеют одинаковую точку, то они должны включать в себя оба варианта коэффициентов и соответствующие им функции, поэтому необходимо включать субстраты этих функций в объединение в Теореме Моро Рокафеллара.

$$\Rightarrow \partial f(x) = \left\{g \ : \ ||g||_\infty \le 1, \quad g^\top x = ||x||_1\right\}$$

$p = 2)$

$$||x||^2 = \sqrt{x_1^2 + \cdots + x_n^2}$$

При $x \equiv 0$ функция дифференцируема и $\Rightarrow \partial f(x) = \nabla_i f(x) = \frac{x_i}{||x||^2}$.
При $x = 0$:

$$\nu \in \partial f(x) \Leftrightarrow \forall y \in \mathbb{R}^n \to ||y||_2 \ge ||x||_2 + (\nu, y - x) = (\nu, y) \Leftrightarrow ||\nu||_2 \le 1$$

Последний переход верен в левую сторону из нер-ва Коши-Буняковскую, а в правую можно доказать, предположив обратное ($||\nu||_2 > 1$), и взяв y коллинеарное ν,тогда нер-во К-Б певратится в равенство и получим $|(\nu, y)| = ||\nu||_2 ||y||_2 > ||y||_2$, т. е. противоречие.
Значит, $\partial f(0) = \{z \in \mathbb{R}^n \Big| ||z||_2 \le 1\}$. И в итоге:

$$\partial f(x) = \begin{cases} \{z \in \mathbb{R}^n \Big| ||z||_2 \le 1\}, x = 0 \\ \frac{x}{||x||^2}, x \equiv 0 \end{cases}$$

$p = \infty)$

$$f(x) = ||x||_\infty = \max_{1 \le i \le n} |x_i| = sup_{||w||_1 \le 1} x^T w$$

При этом $\nabla f = argmax_{||w||_1 \le 1} x^T w$.

$$\Rightarrow \partial f(x) = \{||w||_1 \le 1, z^T w = ||z||_\infty\}$$

№4

$$f(x) = ||Ax - b||_1^2 = h(g(t(x))),$$

где $h(x) = x^2$, $g(x) = ||x||_1$, $t(x) = Ax - b$. Тогда по Chain rule (т.к. все три функции выпуклы, дифференцируемы и определены на откр выпуклом мн-ве):

$$\partial f(x) = 2gA^T \partial t(Ax - b) = 2||Ax - b||_1 A^T \partial(||t||_1)(Ax - b)$$

$\partial(||t||_1)$ знаем из 3. В итоге:

$$\partial f(x) = \{2||Ax - b||_1 A^T g \Big| g \in \mathbb{R}^n, ||g||_\infty \le 1, g^T(Ax - b) = ||Ax - b||_1\}$$

№5

$$f(x) = e^{||x||_2} = h(g(x)),$$

где $h(x) = e^x, g(x) = ||x||_2$. Применим Chain rule (функции очевидно выпуклы и h диффиринцируема) и воспользуемся найденным в третьем номере субградиентом для евклидовой нормы:

$$\partial f(x) = e^{||x||_2} \partial g(x) = \begin{cases} \{z \in \mathbb{R}^n \Big| ||z||_2 \le 1\}, x = 0 \\ e^{||x||_2} \frac{x}{||x||^2}, x \equiv 0 \end{cases}$$