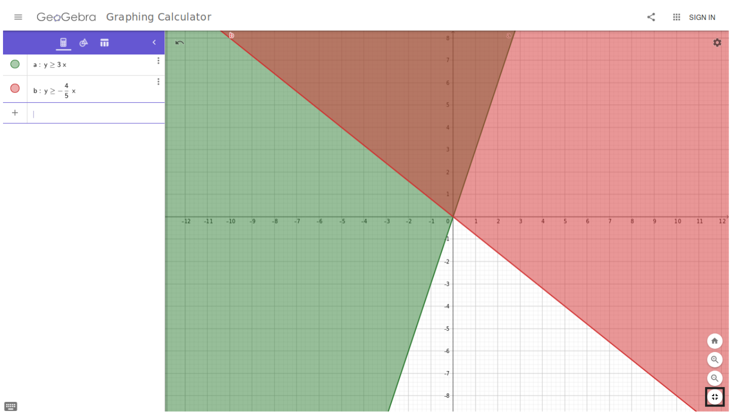
ДЗ-2 по Методам Оптимизации
Попов Николай, 776 группа

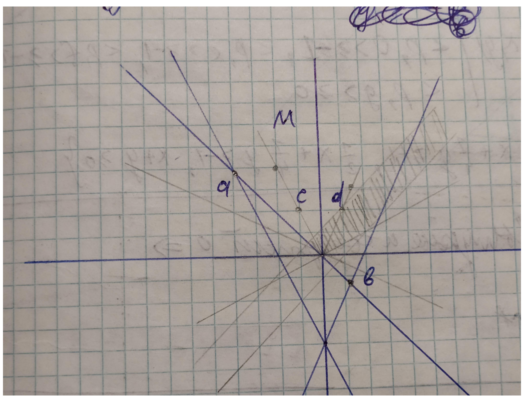
Conjugate sets

(1)
$S = \{(-3,1),(2,3)(4,5)\}$

$$S^* = \begin{cases} -3x + y \geq 0 \\ 2x + 3y \geq 0 \\ 4x + 5y \geq 0 \end{cases} \Rightarrow S^* = \begin{cases} y \geq 3x \\ y \geq -\dfrac{4}{5}x \end{cases}$$



(2)
Представлю данное множество точек плоскости как сумму Минковского выпуклой оболочки точек a и b и конической оболочки точек c и d (точки подписаны на рисунке). $S = conv(a,b) + cone(c,d)$



Найду координаты этих точек:

$$\text{a)}\begin{cases} x + y = 0 \\ 2x + y = -4 \end{cases} \Rightarrow a = (-4, 4)$$

$$\text{b)}\begin{cases} x + y = 0 \\ -2x + y = -4 \end{cases} \Rightarrow b = (\frac{4}{3}, -\frac{4}{3})$$

c) Соответсвует направлению вверх прямой $2x+y = -4 \Rightarrow c = (-1, 2)$

d) Соответсвует направлению вверх прямой $-2x+y = -4 \Rightarrow d = (1, 2)$

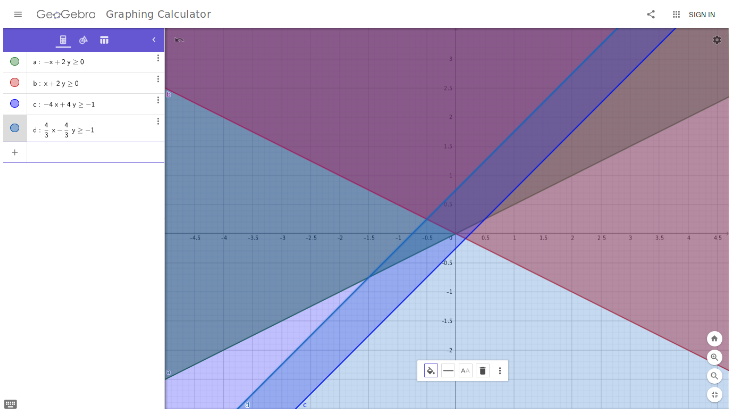
Тогда по теореме с семинара о сопряженном множестве для многогранного множества:

$$S^* = \{p = (x, y)| < c, p >\geq 0, < d, p >\geq 0, < a, p >\geq -1, < b, p >\geq -1\}$$

или

$$S^* = \{(x, y)| - x + 2y \geq 0, x + 2y \geq 0, -4x + 4y \geq -1, \frac{4}{3}x - \frac{4}{3}y \geq -1\}$$

По свойству с семинара в силу того, что множество S является замкнутым, выпуклым и содержит ноль, то $S^{**} = S$.
Из последнего получаем, что $S^{***} = S^*$.



(3)
S^n_+−конус положительно полуопределенных матриц.

$$(S^n_+)^* = \{Y \in S^n | \; \forall X \in S^n_+ \hookrightarrow < X, Y >= tr(XY) \geq 0\}$$

Утверждаю, что $(S^n_+)^* = S^n_+$.
Докажем вложения этих множеств одно в другое:
1)$Y \in (S^n_+)^* \Rightarrow Y \in S^n_+ \Leftrightarrow Y \notin S^n_+ \Rightarrow Y \notin (S^n_+)^*$
$Y \notin S^n_+ \Leftrightarrow u \in R^n : u^T Y u < 0$.
Т.к. $u^T Y u = \sum_i \sum_j Y_{ij} u_i u_j$ и для матрицы $C = uu^T : C_{ij} = u_i u_j \hookrightarrow tr(CY) = \sum_i [CY]_{ii} = \sum_i \left(\sum_j u_i u_j Y_{ji} \right) = \sum_i \sum_j Y_{ij} u_i u_j = u^T Y u \qquad (*)$,
то имеем $tr(CY) =< C, Y >< 0$.
Но матрица С - симметричная (очевидно) и положительно полуопределенная, что легко понять из критерия положительной полуопределенности (из Википедии) : матрица положительно полуопределена, если любой её главный минор неотрицателен. Главный минор - это определитель подматрицы в исходной матрице с одними и теми же номерами строк и столбцов. Т.е. $C \in S^n_+$. Тогда из $< C, Y >< 0 \Rightarrow Y \notin (S^n_+)^*$. Доказано в одну сторону.

2) $Y \in S^n_+ \Rightarrow Y \in (S^n_+)^*$. Рассмотрим произвольную матрицу $X \in S^n_+$. По теореме из линала любая симметричная матрица имеет n разных собственных векторов, образующих ортонормированный базис, т.е. может быть диагонализуема:

$X = RDR^T, R = [u_1, \cdots, u_n], u_i -$ собственный вектор, соответсвующий $\lambda_i : \lambda_i \geq 0$ (из положительной определенности матрицы X).

Вводя обозначение $C^{(k)} = u_k u_k^T$, получим строку матрицы $A = RD$:

$A_{i[1,n]} = (\lambda_1 u_1^i; \lambda_2 u_2^i; \cdots; \lambda_n u_n^i) \Rightarrow$

$X_{ij} = \sum_{k=1}^n \lambda_k u_k^i u_k^j = \sum_{k=1}^n \lambda_k C_{ij}^{(k)} = \sum_{k=1}^n \lambda_k u_k u_k^T$

Тогда

$$tr(YX) = tr\left(Y \sum_{k=1}^n \lambda_k u_k u_k^T\right) = tr\left(\sum_{k=1}^n \lambda_k Y u_k u_k^T\right) = \sum_{k=1}^n \lambda_k tr\left(Y u_k u_k^T\right) =$$

$$= \sum_{k=1}^n \lambda_k tr\left(u_k u_k^T Y\right) = \sum_{k=1}^n \lambda_k u_k^T Y u_k \geq 0,$$

поскольку матрицы X и У положительно полуопределены и, значит, имеют неотрицательные собственные значения и представляют положительно полуопределенную квадратичную форму.

$\forall X \in S_+^n \hookrightarrow tr(YX) \geq 0 \Rightarrow Y \in (S_+^n)^*$. Утверждение доказано.

(4)

$K = \{(x, y, z)|\ y > 0, ye^{\frac{x}{y}} \leq z\} \Rightarrow z > 0 \qquad (*)$

$K^* = \{(u, v, w)|\forall(x, y, z) \in K \hookrightarrow ux + vy + wz \geq 0\}$

Найдем ограничения на (u, v, w), изучая знак выражения $Q = ux + vy + wz$ при допустимых значениях чисел x,y,z из (*).
0) При $z \to +\infty$ можем получить отрицательное Q, поэтому $w \geq 0$.
1) При $w = 0$ нужно $ux + vy \geq 0$, поэтому $u = 0$ и $v \geq 0$.
2) При $w > 0$: подставим в Q $z = ye^{\frac{x}{y}}$, тогда если $A(x, y) = ux + vy + wye^{\frac{x}{y}} \geq 0$, то $Q \geq 0$.
a) Если u>0 при $x \to -\infty$, A не ограничена снизу.
b) Если u=0, то при $x \to -\infty \hookrightarrow A \to vy$, и $vy \geq 0 \Rightarrow v \geq 0$.
c) Если u<0, то минимизуем значение A(x,y) по x:
$0 = \dfrac{\partial A(\tilde{x}, y)}{\partial x} = u + we^{\frac{\tilde{x}}{y}} \Rightarrow \tilde{x} = y \log\left(-\dfrac{u}{w}\right) \Rightarrow$
$\Rightarrow A(\tilde{x}, y) = y(u \log\left(-\dfrac{u}{w}\right) + v - u)$.
Поскольку y>0, то $A(\tilde{x}, y) \geq 0$ при $(u \log\left(-\dfrac{u}{w}\right) + v - u) \geq 0$

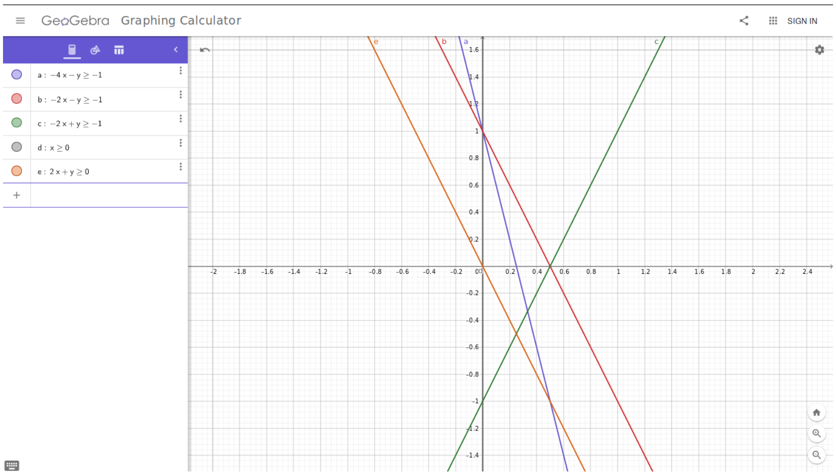
В итоге, собираем полученные точки в множество:
$K^* = \{(u, v, w)|w \geq 0, u = 0, v \geq 0\} \cup \{(u, v, w)|w > 0, u < 0, u \log\left(-\dfrac{u}{w}\right) + v - u) \geq 0\}$

(5)

$$S = conv\{(-4, -1), (-2, -1), (-2, 1)\} + cone\{(1, 0), (2, 1)\}$$

По теореме с семинара о сопряженном множестве для многогранного множества:

$S^* = \{(x, y)| -4x - y \geq -1, -2x - y \geq -1, -2x + y \geq -1, x \geq 0, 2x + y \geq 0\}$



Дуальное множество это четырехугольник, ограниченный на рисунке осью ОУ, оранжевой, зеленой и синей прямыми.

(6)
Нужно доказать: $S = S^* \Leftrightarrow S \equiv B$, где $B = \{x \in R^n | \; ||x|| \leq 1\}$.
Предположим, что для какого-то множества S верно самосопряжение $S = S^*$ по определению из условия:

$$S^* = \{y \in R^n | \forall x \in S \hookrightarrow < x, y > \leq 1\} = S \Rightarrow \forall x, y \in S \hookrightarrow < x, y > \leq 1 \Rightarrow$$

$$\forall x \in S \hookrightarrow < x, x > = ||x||^2 \leq 1 \Rightarrow \forall x \in S \hookrightarrow ||x|| \leq 1$$

Теперь докажем в другую сторону, т.е. покажем, что такой шар B самосопряженный. $B^* = \{y \in R^n | \; \forall x \in B \hookrightarrow \langle x, y \rangle \leq 1\}$. Переберем всевозможные y. Если $y \in B \Rightarrow ||y|| \leq 1 \Rightarrow \langle x, y \rangle \leq |\langle x, y \rangle| \leq ||x|| \cdot ||y|| \leq 1 \; (x \in B) \Rightarrow y \in B^*$. Иначе $y \notin B \Rightarrow ||y|| > 1 \Rightarrow \exists x \in B \; \& \; \exists C > 1 : ||x|| = 1 \; \& \; y = Cx$ (x - пересечение сферы, заданной шаром и отрезком между y и центром шара). Тогда $\langle x, y \rangle = \langle x, Cx \rangle = C\langle x, x \rangle = C||x||^2 = C > 1 \Rightarrow y \notin B^*$

(7) Используем факт: скалярное произведение векторов в линейном пространстве не зависит от базиса. Будем работать в двух базисах: в первом базисе $S = \{x | \sum_{i=1}^n a_i^2 x_i^2 \leq e^2\}$.Сделаем переход в базис, в котором множество S станет шаром: $\bar{y} = diag(a_i)_{i=1}^n \cdot \bar{x} \Rightarrow y_i = a_i x_i$. Тогда в этом базисе S описывается $\sum_{i=1}^n y_i^2 \leq e^2$, т.е. это шар $B(0, e)$. Из примера с семинара знаем, что $S^* = B\left(0, \dfrac{1}{e}\right)$, что описывается уравнением $\sum_{i=1}^n y_i^2 \leq \dfrac{1}{e^2}$. Выполняем переход в первый базис: $\bar{x} = diag\left(\dfrac{1}{a_i}\right)_{i=1}^n \cdot \bar{y} \Rightarrow y_i = a_i x_i$.

Тогда в исходном базисе S^* описывается уравнением $\sum_{i=1}^n \left(\frac{x_i}{a_i}\right)^2 \leq \frac{1}{e^2}$
Поскольку скалярное произведение векторов во втором базисе равно таковому в первом (оно инвариантно), то полученное множество S^* при переходе из второго базиса в первый останется сопряженным множеству S.

Conjugate function

(1)
$f(x) = -\frac{1}{x}, x > 0;\quad f^*(y) = \underset{x>0}{sup}(xy + \frac{1}{x}).$
Поскольку $f^*(y)$ неограничена при $\forall y$, то $domf^* = \emptyset$, т.е. сопряженная функция не определена.

(2)
$f(x) = -0.5 - \log x, x > 0;\quad f^*(y) = \underset{x>0}{sup}\ (xy + 0.5 + \log x) = \underset{x>0}{sup}\ g(x, y)$
$domf^* = \{y < 0\};\quad g'_x = y + \frac{1}{x} = 0 \Rightarrow x = -\frac{1}{y} \Rightarrow$
$f^*(y) = -0.5 - \log(-y)$

(3)
$f(x) = \log(\sum_{i=1}^n e^{x_i})$
$f^*(y) = \sup_x[< x, y > - \log(\sum_{i=1}^n e^{x_i})] = \sup_x\ g(x, y))$
$d_x g =< y, dx > - \frac{1}{\sum_{i=1}^n e^{x_i}} d\left(\sum_{i=1}^n e^{x_i}\right) = y^T dx - \frac{1}{e^{f(x)}} d(e^{f(x)}) = y^T dx -$
$- \frac{e^f}{e^f} df =< y - \bigtriangledown f, dx >;\quad \bigtriangledown_x g = y - \bigtriangledown f = 0 \Rightarrow y = \bigtriangledown f \Rightarrow$
$(*)\qquad y_i = \frac{\partial f}{\partial x_i} = \frac{e^{\tilde{x}_i}}{\sum_{i=1}^n e^{\tilde{x}_i}} \Rightarrow y_i = \frac{e^{\tilde{x}_i}}{e^{f(\tilde{x})}} \Rightarrow \tilde{x}_i = \log y_i + f(\tilde{x}) \Rightarrow$
$f^*(y) = \sum_i y_i(\log y_i + f(\tilde{x})) - f(\tilde{x}) = f(\tilde{x}) \sum_i y_i - f(\tilde{x}) + \sum_i y_i \log y_i =$
$= \sum_i y_i \log y_i$

Из $(*)$ получаем $domf^* = \{y | \sum_i y_i = 1, y_i > 0\ \forall\ 1 \leq i \leq n\}$

(4)
$f(x) = -\sqrt{a^2 - x^2}, |x| \leq a, a > 0;\quad f^*(y) = \sup_{|x|\leq a}(xy + \sqrt{a^2 - x^2}) =$
$= \sup_{|x|\leq a} g(x, y);\quad g'_x = y + \frac{-x}{\sqrt{a^2 - x^2}} = 0 \Leftrightarrow y = \frac{x}{\sqrt{a^2 - x^2}}\ (domf^* = R)$

$y^2 = \dfrac{x^2}{a^2 - x^2} \Rightarrow x^2 = \dfrac{a^2y^2}{y^2+1} \Rightarrow f^*(y) = \dfrac{ay^2}{\sqrt{y^2+1}} + \sqrt{a^2 - \dfrac{a^2y^2}{y^2+1}} =$
$a\dfrac{y^2}{\sqrt{y^2+1}} + \dfrac{a}{\sqrt{y^2+1}} = a\dfrac{y^2+1}{\sqrt{y^2+1}} = a\sqrt{y^2+1}$

(5)

$f(X) = -ln(detX), X \in S^n_{++};\quad f^*(Y) = \sup\limits_{X \succ 0}(< X, Y > +ln(detX))$

$g(X,Y) = (< X, Y > +ln(detX);\quad d_X g =< Y, dX > + \dfrac{d(detX)}{detX} =$
$=< Y, dX > + \dfrac{detX}{detX} \cdot < X^{-T}, dX > \Rightarrow \bigtriangledown_X(g) = Y + X^{-1}$(т.к. $X \in S^n$)
$\bigtriangledown_X(g) = 0 \Rightarrow X^{-1} = -Y \Rightarrow I = X \cdot X^{-1} = X(-Y) \Rightarrow$
$\Rightarrow X = -(Y^{-1}) = (-Y)^{-1}$.
Покажем, что для ограниченности супремума, нужно $domf^* = \{Y \prec 0\}$, что равносильно отрицательности всех собственных значений матрицы Y. Предположим противное, что $\exists \lambda \geq 0$ и соответствующий собственный вектор $v : ||v||_2 = 1$. Возьмем $X = I + tvv^T$ $(t > 0)-$ симметричная положительно определенная матрица (положительная полуопределенность матрицы vv^T показана в номере 3 из раздела про дуальные множества). Тогда имеем
$tr(XY) = tr[(I + tvv^T)Y] = tr[Y(I + tvv^T)] = tr[Y + tYvv^T] = trY + tr(t\lambda vv^T) = trY + \lambda t||v||_2 = trY + \lambda t$

$ln(detX) = ln(det(I+tvv^T)) = ln(\prod_i \lambda_i(I+tvv^T)) = ln(\prod_i (1+t\lambda_i(vv^T))$

Поскольку для $v \in R^n \hookrightarrow det(vv^T - \lambda I) = 0 \Leftrightarrow (-1)^n \lambda^{n-1}(\lambda - \sum_i v_i^2)$, то в нашем случае $\lambda_1(vv^T) = 1;\quad 1 < i \leq n \hookrightarrow \lambda_i(vv^T) = 0$.

Тогда $ln(detX) = ln(1+t)$ и $\sup\limits_{X \succ 0}(< X, Y > +ln(detX)) = sup_{t>0}(trY + \lambda t + ln(1+t)) = +\infty$

Тогда $f^*(Y) =< -Y^{-1}, Y > +ln(det((-Y)^{-1})) = tr(-Y^{-1}Y) + ln\left(\dfrac{1}{det(-Y)}\right) =$

$= -n - ln(det(-Y))$

(6)

$f(x) = g(Ax), x \in R^n, detA \neq 0;$

$f^*(y) = \sup_x(< x, y > -g(Ax)) = \sup_x(< A^{-1}Ax, y > -g(Ax)) \overset{(1)}{=}$
$= \sup_x(< Ax, A^{-T}y > -g(Ax)) = [a = Ax; b = A^{-T}y] \overset{(2)}{=}$
$= sup_a(< a, b > -g(a)) = g^*(b) = g^*(A^{-T}y)$
Переход (1) $:< Ax, y >= (Ax)^T y = x^T(A^T y) =< x, A^T y >$
Так как матрица A не вырождена, то взятие супремума по a после замены $a = Ax$ равносильно взятию супремума по x.

Subgradient and subdifferential

(1)
x_0- точка минимума функции f(x) на множестве S $\Leftrightarrow 0 \in \partial_S f(x_0)$

$(\rightarrow) \quad min\ x_0 \Rightarrow \forall x \in S \hookrightarrow f(x) \geq f(x_0) \Rightarrow$
$f(x) \geq f(x_0)+ < \bar{0}, x - x_0 >\Rightarrow \bar{0} \in \partial_S f(x_0)$

$(\leftarrow)\bar{0} \in \partial_S f(x_0) \Rightarrow \forall x \in S \hookrightarrow f(x) \geq f(x_0)+ < \bar{0}, x - x_0 >= f(x_0)$, т.е. x_0 - точка минимума функции на S.

(2)
Используем теорему Дубовитского-Милютина:

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

В следующих трех задачах используется субдифференциал нормы.
Покажем, что $(*)$ $\partial||x|| = V(x) = \{v \in R^n |\ (1) < v, x >= ||x||,\ (2)\ ||v||_* \leq 1\}$, где $||x||_* = \sup_{||u|| \leq 1} < x, u >$ - сопряженная норма, используя неравенство Гельдера (3) $|< x, y >| \leq ||x|| \cdot ||y||_*$ и функцию, сопряженную к норме (по Фенхелю) $||x||^* = \sup_u(< x, u > -||x||) = 0$ при $||x||_* \leq 1$(иначе не определена)

1)Пусть $v \in V(x)$ и y - произвольный вектор. Тогда :
$||x||+ < v, y - x >= ||x||+ < v, y > - < v, x >\overset{(1)}{=} ||x||+ < v, y > -||x|| =$
$=< v, y >\overset{(3)}{\leq} ||y|| \cdot ||v||_* \overset{(2)}{\leq} ||y|| \Rightarrow \forall y \hookrightarrow ||y|| \geq ||x||+ < v, y - x >\Rightarrow$
$v \in \partial||x|| \Rightarrow V(x) \subset \partial||x||$

2)Пусть $v \in \partial||x|| \Rightarrow \forall y \hookrightarrow ||y|| \geq ||x|| + < v, y - x > \Leftrightarrow \langle v, y\rangle - ||y|| \leq \langle v, x\rangle - ||x|| \Rightarrow ||v||^* = \sup_y(\langle v, y\rangle - ||y||) \leq \langle v, x\rangle - ||x|| \Rightarrow \langle v, x\rangle - ||x|| \geq 0$, при $||v||_* \leq 1 \Rightarrow ||v||_* \leq 1$ & $0 \leq \langle x, y\rangle - ||x|| \overset{(3)}{\leq} ||x|| \cdot ||v||_* - ||x|| = ||x|| \cdot (||v||_* - 1) \leq 0 \Rightarrow ||x|| = \langle v, x\rangle$.

$||v||_* \leq 1$ & $||x|| = \langle v, x\rangle \Rightarrow v \in V(x) \Rightarrow \partial||x|| \subset V(x)$.
В итоге, $\partial||x|| = V(x)$.

(3)
$f(x) = ||x||_p,\ p = 1, 2, \infty$

1) $(*) \Rightarrow \partial||x||_1 = \{v|\langle v, x\rangle = ||x||_1, ||v||_* = ||v||_\infty \leq 1\}$

2) $(*) \Rightarrow \partial||x||_2 = \{v|\langle v, x\rangle = ||x||_2, ||v||_* = ||v||_2 \leq 1\} \Rightarrow$

$$\Rightarrow \partial||x||_2 = \begin{cases} \dfrac{x}{||x||_2} & , x \neq 0 \\ v : ||v||_2 \leq 1 & , x = 0 \end{cases}$$

3) $(*) \Rightarrow \partial||x||_\infty = \{v|\langle v, x\rangle = ||x||_\infty, ||v||_* = ||v||_1 \leq 1\}$

(4)
$f(x) = (||Ax - b||_1)^2$
По теореме о субдифференциале сложной функции:
$\partial f(x) = 2||Ax - b||_1 \cdot \partial||Ax - b||_1$

$h(x) = ||x||_1 \overset{(*)}{\Rightarrow} \partial h(x) = \{v|\langle v, x\rangle = ||x||_1, ||v||_\infty \leq 1\}$

По свойству с семинара для выпуклой функции $|| \cdot ||_1$:
$||Ax - b||_1 = h(Ax - b) \Rightarrow$
$\partial||Ax - b||_1 = A^T\partial h(Ax - b) = A^T\{v|\ ||v||_\infty \leq\ 1, \langle v, Ax - b\rangle = ||Ax - b||_1\}$

$\partial f(x) = 2A^T||Ax - b||_1\{v|\ ||v||_\infty \leq\ 1, \langle v, Ax - b\rangle = ||Ax - b||_1\}$

(5)
Поскольку экспонента - это монотонно неубывающая дифференцируемая выпуклая функция, то теореме о субдифференциале сложной функции для $f(x) = e^{||x||}$ получаем:
$(*) \Rightarrow \partial f(x) = e^{||x||}\partial||x|| = e^{||x||}\{v|\langle v, x\rangle = ||x||, ||v||_* \leq 1\}$