Московский Государственный Университет

Сопряжённость и субдифференциалы

Выполнил: Курцев Д.В.

Группа: 517

Факультет Вычислительной математики и кибернетики

Кафедра Математических методов прогнозирования

Май 2024

Часть 1. Сопряжённость. Двойственность. ККТ

Задача 1

a) $f(x) = -\left(\prod_{i=1}^{d} x_i\right)^{1/d}$, где $x \in \mathbb{R}^d_+$.

Найдём сопряжённую функцию $f^*(y)$ по определению:

$$f^*(y) = \sup_{x \in \mathbb{R}^d_+} \left(\langle y, x\rangle - f(x)\right)$$

Тогда получим, что сопряжённая функция будет иметь следующий вид:

$$f^*(y) = \sup_{x \in \mathbb{R}^d_+} \left(\sum_{i=1}^{d} y_i x_i + \left(\prod_{i=1}^{d} x_i\right)^{1/d}\right)$$

Найдём производную по x и приравняем её к нулю:

$$\frac{\partial(\sum_{i=1}^{d} y_i x_i + \left(\prod_{i=1}^{d} x_i\right)^{1/d})}{\partial x_i} = y_i + \frac{1}{d}\frac{1}{x_i}\left(\prod_{i=1}^{d} x_i\right)^{1/d} = y_i - \frac{1}{d}\frac{1}{x_i} f(x).$$

Отсюда получим, что $y_i = \frac{f(x)}{x_i d}$ и $y_i < 0 \quad \Rightarrow \quad y \in \mathbb{R}^d_{--}$. Тогда сопряжённая функция будет равна:

$$f^*(y) = \left(\sum_{i=1}^{d} \frac{f(x)}{x_i d} x_i - f(x)\right) = 0$$

Если $y \in \mathbb{R}^d \backslash \mathbb{R}^d_{--}$, то $\exists y_i \geq 0$. Тогда

$$\left(\sum_{i=1}^{d} y_i x_i + \left(\prod_{i=1}^{d} x_i\right)^{1/d}\right) \underset{x_i \to +\infty}{\longrightarrow} +\infty$$

Таким образом, получим, что сопряжённая функция будет равна:

$$f^*(y) = \begin{cases} 0, & y \in \mathbb{R}^d_{--} \\ +\infty, & \text{иначе} \end{cases}$$

b) $f(x) = \max_{i=1,\ldots,d}\{x_i\}$, где $x \in \mathbb{R}^d$.

Тогда сопряжённая функция по определению будет равна:

$$f^*(y) = \sup_{x \in \mathbb{R}^d} \left(\sum_{i=1}^{d} y_i x_i - \max_{i=1,\dots,d}\{x_i\} \right)$$

Обозначим $M = \max_{i=1,\dots,d}\{x_i\}$, тогда:

$$\sum_{i=1}^{d} y_i x_i - \max_{i=1,\dots,d}\{x_i\} \leq \sum_{i=1}^{d} y_i M - M = M(\sum_{i=1}^{d} y_i - 1)$$

Это выражение равняется 0, если $\sum_{i=1}^{d} y_i = 1$. Тогда, если $\sum_{i=1}^{d} y_i = 1$ и $y_i \geq 0$, то супремум достигается и равен нулю. Пусть $x_i = M$, возьмём вектор y из нулей и с координатой $y_i = M$. Тогда:

$$\sum_{i=1}^{d} y_i x_i - M = 1 \cdot M - M = 0$$

Покажем, что в остальных случаях супремума не существует.

1) Пусть $\sum_{i=1}^{d} y_i > 1$. Тогда $\exists x : x_i = M > 0$ для $\forall i = \overline{1, d}$. В итоге получим, что:

$$\sum_{i=1}^{d} y_i x_i - \max_{i=1,\dots,d}\{x_i\} = M(\sum_{i=1}^{d} y_i - 1) \underset{M \to +\infty}{\longrightarrow} +\infty$$

2) Пусть $\sum_{i=1}^{d} y_i < 1$. Тогда $\exists x : x_i = M < 0$ для $\forall i = \overline{1, d}$. Аналогично имеем:

$$\sum_{i=1}^{d} y_i x_i - \max_{i=1,\dots,d}\{x_i\} = M(\sum_{i=1}^{d} y_i - 1) \underset{M \to -\infty}{\longrightarrow} +\infty$$

3) Пусть теперь $\exists y_i < 0$. Тогда возьмём вектор x, состоящий из нулей и с координатой $x_i = M < 0$. Отсюда получим:

$$\sum_{i=1}^{d} y_i x_i - \max_{i=1,\dots,d}\{x_i\} = M y_i - 0 = M y_i \underset{M \to -\infty}{\longrightarrow} +\infty$$

Таким образом, получим, что сопряжённая функция будет равна:

$$f^*(y) = \begin{cases} 0, & \sum_{i=1}^{d} y_i = 1 \text{ и } y_i \geq 0 \\ +\infty, & \text{иначе} \end{cases}$$

c) $f(X) = \mathrm{tr} X^{-1}$, где $X \in S^d_{++}$

По определению сопряжённая функция будет следующей:

$$f^*(Y) = \sup_{X \in S^d_{++}} (\langle Y, X \rangle - \mathrm{tr} X^{-1}) = \sup_{X \in S^d_{++}} (\mathrm{tr}(YX) - \mathrm{tr} X^{-1}) = \sup_{X \in S^d_{++}} \mathrm{tr}(YX - X^{-1})$$

В предыдущем домашнем задании было показано, что функция $\mathrm{tr} X^{-1}$ является выпуклой. Функция $\langle Y, X \rangle$ линейна, а значит сопряжённая функция $f^*(Y)$ будет вогнутой по X, т.е. её стационарная точка будет точкой максимума и супремума. Чтобы её найти, вычислим производную и приравняем её нулю:

$$d_X \mathrm{tr}(YX - X^{-1}) = \mathrm{tr}(d(YX) - d(X^{-1})) = \mathrm{tr}(YdX + X^{-1} dX X^{-1}) = \mathrm{tr}(YdX + X^{-2} dX)$$

Таким образом,

$$\frac{\partial\, \mathrm{tr}(YX - X^{-1})}{\partial X} = Y + X^{-2}$$

То есть $Y = -X^{-2}$ и $X = (-Y)^{-\frac{1}{2}}$, и $Y \in S^d_{--}$. Таким образом сопряжённая функция будет равна:

$$f^*(Y) = \mathrm{tr}\left(Y(-Y)^{-\frac{1}{2}} - (-Y)^{\frac{1}{2}}\right)$$

Если обозначить $-Y = ZZ$, то тогда получим:

$$f^*(Y) = \mathrm{tr}\left(-ZZZ^{-1} - Z\right) = -2\mathrm{tr} Z = -2\mathrm{tr}(-Y)^{\frac{1}{2}}$$

Таким образом, получим, что сопряжённая функция будет равна:

$$f^*(Y) = \begin{cases} -2\mathrm{tr}(-Y)^{\frac{1}{2}}, & Y \in S^d_{--} \\ +\infty, & \text{иначе} \end{cases}$$

Задача 2

$$\min_{x \in \mathbb{R}^d} - \sum_{i=1}^{m} \log(b_i - a_i^T x)$$

с ограничением $\{x \in \mathbb{R}^d | a_i^T x < b_i, \quad \forall i = 1, \dots, m\}$

Введем дополнительные переменные $y_i = b_i - a_i^T x$.

Таким образом, преобразованная задача принимает вид:

$$\min_{y \in \mathbb{R}^d} - \sum_{i=1}^{m} \log y_i$$

с ограничениями

$$y_i = b_i - a_i^T x$$

$$y_i > 0 \quad \forall i = 1, \dots, m$$

По определению двойственная задача это максимум двойственной функции - инфимум функции Лагранжа: $g(\lambda, \mu) = \inf_{x,y} L(x, y, \lambda, \mu)$. Тогда запишем лагранжиан нашей задачи:

$$L(x, y, \lambda, \mu) = -\sum_{i=1}^{m} \log y_i + \sum_{i=1}^{m} \mu_i (y_i - b_i + a_i^T x) - \sum_{i=1}^{m} \lambda_i y_i$$

Найдём производные по переменным x и y и приравняем их к нулю:

$$\frac{\partial L}{\partial y_i} = -\frac{1}{y_i} - \lambda_i + \mu_i = 0 \quad \Rightarrow \quad y_i = \frac{1}{\mu_i - \lambda_i}, \mu_i > \lambda_i \geq 0$$

$$\frac{\partial L}{\partial x} = \sum_{i=1}^{m} \mu_i a_i = 0$$

Тогда подставляя, получим:

$$g(\lambda, \mu) = \sum_{i=1}^{m} \log(\mu_i - \lambda_i) + \sum_{i=1}^{m} \frac{\mu_i}{\mu_i - \lambda_i} - \sum_{i=1}^{m} \mu_i b_i + \sum_{i=1}^{m} \mu_i a_i^T x - \sum_{i=1}^{m} \frac{\lambda_i}{\mu_i - \lambda_i} =$$

$$= \sum_{i=1}^{m} \log(\mu_i - \lambda_i) + \sum_{i=1}^{m} \frac{\mu_i - \lambda_i}{\mu_i - \lambda_i} - \sum_{i=1}^{m} \mu_i b_i + \left(\sum_{i=1}^{m} \mu_i a_i^T \right) x =$$

$$= \sum_{i=1}^{m} \log(\mu_i - \lambda_i) - \sum_{i=1}^{m} \mu_i b_i + m$$

Таким образом, двойственная задача будет следующей:

$$\max_{\mu_i > \lambda_i \geq 0} \sum_{i=1}^{m} \log(\mu_i - \lambda_i) - \sum_{i=1}^{m} \mu_i b_i$$

с ограничением $\sum_{i=1}^{m} \mu_i a_i = 0$

Задача 3

1) LP-релаксация

$$\min_x c^T x$$

с ограничениями

$$Ax \leq b$$

$$0 \leq x_i \leq 1, \quad \forall i = 1, \ldots, d$$

Лагранжиан данной задачи с двойственными переменными $\lambda \geq 0, \mu \geq 0, \nu \geq 0$ имеет следующий вид:

$$L(x, \lambda, \mu, \nu) = c^T x + \lambda^T (Ax - b) - \mu^T x + v^T (x - \mathbf{1})$$

где $\mathbf{1}$ — вектор из единиц соответствующей размерности.

По определению двойственной функции:

$$g(\lambda, \mu, \nu) = \inf_x L(x, \lambda, \mu, \nu)$$

Найдём частную производную лагранжиана по x и приравняем её к нулю:

$$\frac{\partial L}{\partial x} = c + A^T \lambda - \mu + \nu = 0$$

Тогда получим, что:

$$g(\lambda, \mu, \nu) = \inf_x L(x, \lambda, \mu, \nu) = \inf_x (c^T + \lambda^T A - \mu^T + \nu^T) x - \lambda^T b - \nu^T \mathbf{1} = -\lambda^T b - \nu^T \mathbf{1}$$

Таким образом, двойственная задача будет следующей:

$$\min_{\lambda, \mu, \nu} \lambda^T b + \nu^T \mathbf{1}$$

с ограничениями

$$c + A^T \lambda - \mu + \nu = 0$$

$$\lambda \geq 0, \mu \geq 0, \nu \geq 0$$

2)

$$\min_x c^T x$$

с ограничениями

$$Ax \leq b$$

$$x_i (1 - x_i) = 0, \quad \forall i = 1, \ldots, d$$

Лагранжиан данной задачи с двойственными переменными $\lambda \geq 0, \mu$ имеет следующий вид:

$$L(x, \lambda, \mu) = c^T x + \mu^T (x - x^2) + \lambda^T (Ax - b)$$

Найдём частную производную лагранжиана по x и приравняем её к нулю:

$$\frac{\partial L}{\partial x} = c + A^T \lambda + \mu - 2\mu \odot x = 0 \quad \Rightarrow \quad x = \frac{c + \mu + A^T \lambda}{2\mu} = \frac{y}{2\mu}, \mu \neq 0$$

где $\odot$ - поэлементное умножение, а дробь - поэлементное деление.

Подставим, найденное значение x в лагранжиан:

$$g(\lambda, \mu) = \inf_x L(x, \lambda, \mu) = (c^T + \mu^T + \lambda^T A)x - \mu^T x^2 - \lambda^T b = \frac{y^T y}{2\mu} - \frac{\mathbf{1}^T y^2}{4\mu} - \lambda^T b =$$

$$= \frac{\mathbf{1}^T y^2}{4\mu} - \lambda^T b$$

Таким образом, двойственная задача будет следующей:

$$\max_{\lambda \geq 0, \mu \neq 0} \sum_{i=1}^{d} \frac{(c + \mu + A^T \lambda)_i^2}{4\mu_i} - \lambda^T b$$

Задача 4

$$f(x) = \sum_{i=1}^{r} x_{[i]}$$

Значение данной функции является решением следующей оптимизационной задачи:

$$\max_y x^T y = \min -x^T y$$

с ограничениями

$$0 \leq y \leq 1$$

$$\mathbf{1}^T y = r$$

Действительно, можно заметить, что все y_i неотрицательны и не превосходят единицу. Приэтом их сумма обязательна должна равняться r. Тогда выгоднее всего всю массу вектора y положить в компоненты, которые соответствуют наибольшим значениям x_i. Но из-за ограничения $y_i \leq 1$. Мы не можем приравнять какую-то координату $y_j = r$. Поэтому необходимо выбрать наибольшие r координат вектора x.

Более строго это можно показать так. Пусть множество I - это множество индексов вектора x, соответствующих r наибольшим координатам, а множество J остальные. Тогда рассмотрим вектора y^* и y. Определим вектор $y_i^* = 1$, если $i \in I$ и $y_i^* = 0$, если $i \in J$. А y произвольно, который удовлетворят ограничениям. Тогда:

$$x^T y^* = \sum_{i=1}^{d} x_i y_i^* = \sum_{i \in I} x_i = \sum_{i=1}^{r} x_{[i]}$$

То есть при y^* значение нашей максимизируемой функции совпадает с $f(x)$. В то же время:

$$x^T y = \sum_{i=1}^{d} x_i y_i = \sum_{i \in I} x_i y_i + \sum_{j \in J} x_j y_j \leq \sum_{i \in I} x_i y_i + \sum_{j \in I} x_j y_j = \sum_{i \in I} x_i y_i^* = x^T y^*$$

Таким образом, мы показали, что решение y^* нашей оптимизационной задачи совпадает с искомым значением функции $f(x)$.

Аналогично предыдущим задачам запишем лагранжиан данной системы с двойственными переменными λ, μ, ν:

$$L(y, \lambda, \mu, \nu) = -x^T y - \lambda^T y + \mu^T (y - \mathbf{1}) + v^T (\mathbf{1}^T y - r)$$

Найдём частную производную лагранжиана по x и приравняем её к нулю:

$$\frac{\partial L}{\partial x} = -x - \lambda + \mu + \nu \mathbf{1} = 0$$

Тогда получим, что:

$$g(\lambda, \mu, \nu) = \inf_y L(y, \lambda, \mu, \nu) = \inf_y (-x^T - \lambda^T + \mu^T + \nu \mathbf{1}) y - \mu^T \mathbf{1} - \nu r = -\mu^T \mathbf{1} - \nu r$$

Таким образом, двойственная задача будет следеющей:

$$\min_{\lambda, \mu, \nu} \mu^T \mathbf{1} + \nu r$$

с ограничениями

$$x + \lambda - \mu + \nu \mathbf{1} = 0$$

$$\lambda \geq 0, \mu \geq 0$$

Задача 5

$$\min_{x\in\mathbb{R}^d} \|Ax-b\|^2$$

с ограничением:

$$Gx = h,$$

где $A \in \mathbb{R}^{m\times d},\ \mathrm{rank}A = d,\ G \in \mathbb{R}^{n\times d},\ \mathrm{rank}G = n$

Условия ККТ по определению выглядят следующим образом:

$$\begin{cases} f'(x) + \lambda^T(Gx-h)' = 0 \\ Gx = h \end{cases}$$

$f'(x) = 2A^T(Ax-b)$. Тогда получим, что ККТ примет вид:

$$\begin{cases} 2A^T(Ax-b) + G^T\lambda = 0 \\ Gx = h \end{cases}$$

Из первого уравнения получим оптимальное значение x^*:

$$2A^TAx^* - 2A^Tb + G^T\lambda = 0 \ \Rightarrow\ x^* = (A^TA)^{-1}A^Tb - \frac{1}{2}(A^TA)^{-1}G^T\lambda$$

Подставляя найденное решение x во второе уравнение получим:

$$G(A^TA)^{-1}(2A^Tb - G^T\lambda) = 2h \ \Rightarrow\ 2G(A^TA)^{-1}A^Tb - G(A^TA)^{-1}G^T\lambda = 2h$$

Введём обозначение $C = G(A^TA)^{-1}G^T \in \mathbb{R}^{n\times n}$. Тогда:

$$\lambda^* = 2C^{-1}(G(A^TA)^{-1}A^Tb - h)$$

Заметим, что матрица A^TA - квадратная полного ранга, а следовательно к ней существует обратная. То же можно сказать про матрицу C. $A^TA \in \mathbb{S}_+ \Rightarrow (A^TA)^{-1} \in \mathbb{S}_+ \Rightarrow \exists B : B^2 = (A^TA)^{-1}$. Тогда $C = G(A^TA)^{-1}G^T = GBBG^T = D^TD$, где ранг матрицы D равен n, так как $n \le d$, так как $\mathrm{rank}G = n$. А следовательно и матрица C будет матрицей полного ранга и с обратной матрицей. Тогда получим ответ нашей задачи:

$$x^* = (A^TA)^{-1}A^Tb - \frac{1}{2}(A^TA)^{-1}G^T\lambda^*$$

$$\lambda^* = 2(G(A^TA)^{-1}G^T)^{-1}(G(A^TA)^{-1}A^Tb - h)$$

Часть 2. Субградиент и субдифференциал

Задача 1

$f(x) = \max\{-x, x, x^2\} = \max\{f_1, f_2, f_3\}.$

$$f = \begin{cases} -x, & x \in [-1, 0] \\ x, & x \in [0, 1] \\ x^2, & \text{иначе} \end{cases}$$

По свойству субдифференциала: $\partial f = \operatorname{Conv} \cup_{i \in J(x)} \partial f_i$ получим:

$$\partial f = \begin{cases} \partial f_1, & x \in (-1, 0) \\ \partial f_2, & x \in (0, 1) \\ \partial f_3 & x \in (-\infty, -1) \cup (1, +\infty) \\ \operatorname{Conv}(\partial f_1 \cup \partial f_3), & x = -1 \\ \operatorname{Conv}(\partial f_2 \cup \partial f_3), & x = 1 \\ \operatorname{Conv}(\partial f_1 \cup \partial f_2), & x = 0 \end{cases} =$$

$$= \begin{cases} -1, & x \in (-1, 0) \\ 1, & x \in (0, 1) \\ 2x & x \in (-\infty, -1) \cup (1, +\infty) \\ \operatorname{Conv}(\{-1\} \cup \{-2\}), & x = -1 \\ \operatorname{Conv}(\{1\} \cup \{2\}), & x = 1 \\ \operatorname{Conv}(\{-1\} \cup \{1\}), & x = 0 \end{cases}$$

Таким образом, получим, что субдифференциал нашей функции будет следующим:

$$\partial f = \begin{cases} 2x & x \in (-\infty, -1) \\ [-2, -1], & x = -1 \\ -1, & x \in (-1, 0) \\ [-1, 1], & x = 0 \\ 1, & x \in (0, 1) \\ [1, 2], & x = 1 \\ 2x & x \in (1, +\infty) \end{cases}$$

Задача 2

$f(x) = |x-2| + |x+2| + |x-1| = f_1(x) + f_2(x) + f_3(x)$

Модули можно переписать в следующем виде: $f_1(x) = \max\{x-2, 2-x\}$

Аналогично предыдущей задаче получим:

$$\partial f_1 = \begin{cases} -1, & x < 2 \\ 1, & x > 2 \\ [-1, 1], & x = 2 \end{cases}, \partial f_2 = \begin{cases} -1, & x < -2 \\ 1, & x > -2 \\ [-1, 1], & x = -2 \end{cases}, \partial f_3 = \begin{cases} -1, & x < -1 \\ 1, & x > -1 \\ [-1, 1], & x = -1 \end{cases}$$

Тогда по свойству субдифференциала: $\partial f = \partial f_1 + \partial f_2 + \partial f_3$ имеем:

$$\partial f = \begin{cases} -1-1-1, & x < -2 \\ [-1, 1] - 1 - 1, & x = -2 \\ 1-1-1, & -2 < x < 1 \\ 1 + [-1, 1] - 1, & x = 1 \\ 1+1-1, & 1 < x < 2 \\ 1 + 1 + [-1, 1], & x = 2 \\ 1+1+1, & x > 2 \end{cases} = \begin{cases} -3, & x < -2 \\ [-3, -1], & x = -2 \\ -1, & -2 < x < 1 \\ [-1, 1], & x = 1 \\ 1, & 1 < x < 2 \\ [1, 3], & x = 2 \\ 3, & x > 2 \end{cases}$$

Задача 3

$f(x) = \exp(\|Ax - b\|_p)$, где $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ и $p \in [1, +\infty)$

1) Для начала найдём субдифференциал внутренней функции $g(x) = \|x\|_p$:

$\|x\|_p = (|x_1|^p + |x_2|^p + \ldots + |x_n|^p)^{1/p}$. Чтобы найти субдифференциал этой функции, сначала обратим внимание на то, что норма p является нормальной выпуклой функцией. Для выпуклых функций, которые являются дифференцируемыми, субдифференциал в точке x содержит один единственный элемент, который совпадает с градиентом f в этой точке. Функция $\|x\|_p$ дифференцируема в любой точке $x \neq 0$. Дифференциал этой функции:

Рассмотрим несколько случаев.

1.1) $x \neq 0$:

$$\frac{\partial \|x\|_p}{\partial x_i} = \frac{1}{p}\left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}-1} \cdot p|x_i|^{p-1}\operatorname{sgn}(x_i) = |x_i|^{p-1}\operatorname{sgn}(x_i)\left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}-1}$$

Таким образом, градиент $\|x\|_p$ в точке $x \neq 0$ представляет собой вектор

$$\nabla \|x\|_p = \left(\frac{x_1}{|x_1|}|x_1|^{p-1}, \frac{x_2}{|x_2|}|x_2|^{p-1}, \ldots, \frac{x_n}{|x_n|}|x_n|^{p-1}\right)\left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}-1}$$

Это выражение можно упростить:

$$\nabla \|x\|_p = \left(\frac{|x_1|^{p-1}x_1}{|x_1|^p}, \ldots, \frac{|x_n|^{p-1}x_n}{|x_n|^p}\right)\frac{1}{\|x\|_p^{p-1}} = \left(\frac{|x_1|^{p-1}\operatorname{sgn}(x_1)}{\|x\|_p^{p-1}}, \ldots, \frac{|x_n|^{p-1}\operatorname{sgn}(x_n)}{\|x\|_p^{p-1}}\right)$$

Таким образом, субдифференциал нормы задаётся следующим образом:

$$\partial \|x\|_p = \left\{\frac{|x_1|^{p-1}\operatorname{sgn}(x_1)}{\|x\|_p^{p-1}}, \ldots, \frac{|x_n|^{p-1}\operatorname{sgn}(x_n)}{\|x\|_p^{p-1}}\right\}.$$

Заметим, что если $x_i = 0$, то $\operatorname{sgn}(x_i)$ превращается в отрезок $[-1, 1]$, для $p = 1$.

1.2. В случае, когда $x = 0$ (нулевой вектор). По определению субдифференциал $\partial f(x)$ функции $f(x)$ в точке x — это множество векторов z, таких, что выполняется следующее неравенство для всех $\forall y$:

$$f(y) \geq f(x) + \langle z, y - x\rangle$$

Тогда для запишем это определение для нашей функции в точке $x = 0$:

$$\|y\|_p \geq \langle z, y\rangle$$

Это влечёт к условию $\|z\|_q \leq 1$, где q — сопряжённая норма к p, которая определяется как $\frac{1}{p} + \frac{1}{q} = 1$.

Таким образом, для $x = 0$, субдифференциал $\partial\|x\|_p$ — это единичный шар в норме $\|\cdot\|_q$.

2) По теореме субдифференциала для функции $k(x) = h(g(x))$, где $h(y) = Ay - b$ - линейная, монотонно возрастающая функция получим, что

$$\partial k(x) = h'(g(x))\partial g(x),$$

где $\nabla h(y) = A^T$

3) Аналогично по теореме субдифференциала для функции $f(x) = k(h(x))$, где $k(y) = \exp(y)$ - выпуклая, монотонно возрастающая функция получим, что

$$\partial f(x) = k'(h(x))\partial h(x)$$

где $k'(y) = \exp(y)$, т.е. $k'(g(x)) = \exp(\|Ax - b\|_p)$. Таким образом имеем:

$$\partial f(x) = \begin{cases} \exp(\|Ax-b\|_p) \left\{ A^T z : \|z\|_q \le 1, \langle z, Ax - b\rangle = \|Ax-b\|_p, \frac{1}{p} + \frac{1}{q} = 1 \right\}, & x = 0 \\ \exp(\|Ax-b\|_p) A^T \left\{ \frac{|(Ax-b)_1|^{p-1} \operatorname{sgn}((Ax-b)_1)}{\|(Ax-b)\|_p^{p-1}}, \dots, \frac{|(Ax-b)_n|^{p-1} \operatorname{sgn}((Ax-b)_n)}{\|(Ax-b)\|_p^{p-1}} \right\}, & x \neq 0 \end{cases}$$

Задача 4

Индикаторная функция множества $B_{\|\cdot\|_p}(0,1)$ в $\mathbb{R}^n$ задается как:

$$f(x) = \begin{cases} 0, & \text{если } \|x\|_p \le 1, \\ +\infty, & \text{иначе.} \end{cases}$$

1) Для точек внутри множества $B_{\|\cdot\|_p}(0,1)$ функция $f(x)$ определена и равна нулю. Тогда получим, что субградиаент данной функции для $\|x\|_p < 1$ равен нулю.

2) Субдифференциал функции f вне множества $B_{\|\cdot\|_p}(0,1)$ не определен, поскольку функция показывает значение $+\infty$ за пределами этого множества.

3) На границе множества $B_{\|\cdot\|_p}(0,1)$, в точках $\|x\|_p = 1$ получим:

$$f^*(y) = \sup_x(\langle x, y\rangle - f(x)) = \sup_{\|x\|_p = 1} \langle x, y\rangle = \|y\|_q, \quad \frac{1}{p} + \frac{1}{q} = 1$$

$$\partial f(x) = \{s \in S^* : \langle s, x\rangle = f^*(s) + f(x) = \|s\|_q\}.$$

Таким образом получим, что субдифференциал нашей функции имеет следующий вид:

$$\partial f(y) = \begin{cases} 0, & \|x\|_p < 1, \\ \|y\|_q, & \langle y, x\rangle = \|y\|_q, \|x\|_p = 1, \\ +\infty, & \|x\|_p > 1 \end{cases}$$

Задача 6

а) $f(X) = \lambda_{\max}(X), \quad X \in \mathbb{S}$

$f(X) = \lambda_{\max}(X) = \max_{|z|=1} z^T X z$

$\nabla f(X) = zz^T \quad \Rightarrow g \in \partial f(X) \quad \Leftrightarrow \quad g = vv^T$, v - собственный вектор, соответствующий $\lambda_{\max}$