Домашняя работа №3

Бредихин Александр

General optimization problems

Задача 1

Give an explicit solution of the following LP.

$$c^\top x \to \min_{x \in \mathbb{R}^n}$$
$$\text{s.t. } Ax = b$$

Запишем Лагранжжиан этой задачи:

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

Условия ККТ:

$$\begin{cases} \nabla_x L: \ c + A^T \lambda = 0 \\ \nabla_\lambda L: \ Ax = b \end{cases}$$

Заметим, что данная задача - выпуклая, так как функцию, которую минимизируем - линейная (выпуклая) и функция, задающая ограничения типа равенств - аффинная.
Если у системы $Ax = b$ нет решений, то бюджетное множество пустое и решения нет.
Если решение есть, то в общем виде его можно записать через псевдообратную матрицу:

$$x^* = A^\dagger b$$

10/10

И будет выполнено условие Слейтера: есть решение системы, то есть есть допустимая точка из относительной внутренности бюджетного множества. Следовательно, условия ККТ - достаточные и мы получили решение задачи:
$x^* = A^\dagger b$, оптимальное значение $p^* = c^\top A^\dagger b$
если система $Ax = b$ несовместна, то бюджетное множество пустое и $p^* = \infty$

Задача 2

Give an explicit solution of the following LP.

$$c^\top x \to \min_{x \in \mathbb{R}^n}$$
$$\text{s.t. } 1^\top x = 1$$
$$x \succeq 0$$

This problem can be considered as a simplest portfolio optimization problem.

Запишем Лагранжиан:

$$L(x, \lambda, \mu) = c^\top x + \lambda \left(1^\top x - 1\right) - \mu^\top x$$

Запишем условия ККТ:

$$\begin{cases} \nabla_x L: \; c + \lambda \cdot 1^\top - \mu = 0 \\ \nabla_\lambda L: \; 1^T x = 1 \\ \mu_j \geqslant 0, \quad j = [1, n] \\ \mu_j x_j = 0, \quad j = [1, n] \\ x \succeq 0 \end{cases}$$

Заметим, что наша задача выпуклая: так как линейная функция выпуклая (функция для ограничения типа неравенств тоже линейная), функция для ограничения типа равенства - линейная. А также выполнены условия Слейтера: существует допустимая точка - все компоненты вектора x равны $\frac{1}{n}$. Для этой точки ограничения типа равенства выполнены, а ограничения типа неравенства выполнены строго (то есть такая точка из относительной внутренности бюджетного множества), значит условия ККТ являются достаточными.

10/10

Возьмём $x^* = (0, 0, \dots 0, 1, 0 \dots, 0)$ где 1 стоит на месте, где у вектора c минимальная компонента - c_i. Тогда для $\lambda^* = -c_i$ и $\mu_i^* = 0$ и $\mu_j^* = c_j - c_i$ (так как $c_i = \min\limits_j c_j \to c_j - c_i \geq 0 \; \forall j = [1, n]$). Выполнены все условия ККТ, следовательно, такие $x^*, \; \mu^*, \; \lambda^*$ - решение системы, следовательно (так как условия ККТ - достаточные), x^* - решение исходной задачи. Оптимальное значение: $p^* = c_i$

Задача 3

Give an explicit solution of the following LP.

$$\begin{aligned} & c^\top x \to \min_{x \in \mathbb{R}^n} \\ \text{s.t. } & 1^\top x = \alpha \\ & 0 \preceq x \preceq 1 \end{aligned}$$

where α is an integer between 0 and n. What happens if α is not an integer (but satisfies $0 \leq \alpha \leq n$)? What if we change the equality to an inequality $1^\top x \leq \alpha$?

Запишем Лагранжиан для нашей задачи (в выражении $\lambda \in R; \quad \mu, \theta \in R^n$):

$$L(x, \lambda, \mu, \theta) = c^T x + \lambda \left(1^T x - \alpha\right) + \theta^T (x - 1) - \mu^T x$$

Снова записываем условия ККТ:

$$\begin{cases} \nabla_x L: \ c + \lambda \cdot 1^\top - \mu + \theta = 0 \\ \nabla_\lambda L: \ 1^T x = \alpha \\ \mu_j \geqslant 0, \quad j = [1, n] \\ \theta_j \geqslant 0, \quad j = [1, n] \\ \mu_j x_j = 0, \quad j = [1, n] \\ \theta_j (x_j - 1) = 0, \quad j = [1, n] \\ 0 \preceq x \preceq 1 \end{cases}$$

10/10

Условия ККТ снова будут достаточными, так как наша задача - выпуклая: функция, задающая ограничения типа равенств - аффинная (линейная), а функции ту, которую мы оптимизируем и те, которые задают ограничения типа неравенств - линейные, следовательно выпуклые. Также выполнено условие Слейтера: возьмём вектор x с компонентами $\frac{\alpha}{n}$ (по условию $\alpha \leq n$). Для этой точки все ограничения типо неравенств выполнены строго и выполнены ограничения типа равенств. Следовательно, она принадлежит относительной внутренности бюджетного множества. Значит условия ККТ - достаточные.

Б.о.о. считаем, что вектор c имеет вид

$$c_1 \leqslant c_2 \leqslant \ldots \leqslant c_\alpha \leqslant \ldots \leqslant c_n$$

+

Если это не так, то переставим компоненты вектора c и соответсвующие им компоненты вектора x и получим то, что нужно.

Покажем, что решением исходной задачи будет такой x^*, что первые его α координат равны 1, а остальные - 0. Для этого возьмём $\lambda^*, \mu^*, \theta^*$ такие, что $x^*, \lambda^*, \mu^*, \theta^*$ - решение исходной системы.

+

Аналогично предыдущей задаче заметим, что если взять $\lambda = -c_\alpha$, и взять μ такое, что первые α координат равны 0, а $\mu_i^* = c_i + \lambda^*, i = [\alpha+1, n]$ (условие, что $\mu_j \geq 0 \ \forall j = [1, n]$ выполнено, так как $c_i - c_\alpha \geq 0$ компоненты вектора c упорядочены по возрастанию).
А $\theta_i^* = 0 \ \forall i = [1, \alpha]$ и $\theta_i^* = -c_i - c_\alpha \ \forall i = [\alpha + 1, n]$ тогда условие, что $\theta_j (x_j - 1) = 0$ и $\theta_j \geqslant 0, \quad j = [1, n]$ - выполнены (аналогично условиям для μ).

Для выбранных $x^*, \lambda^*, \mu^*, \theta^*$ все условия системы выполнены, следовательно (так как условия ККТ - достаточные), x^* - решение исходной задачи, а оптимальное значение: $p^* = c_1 + \ldots + c_\alpha$

В случае, если $\alpha \notin Z$:
Аналогично предыдущему случаю только теперь $x_i^* = 1$ для $i = [1, [\alpha] - 1]$ И $x_{[\alpha]}^* = 1 + \alpha - [\alpha]$. $\lambda^*, \mu^*, \theta^*$ из предыдущего пункта. Получаем решение системы, следовательно и решение нашей задачи.
Оптимальное значение: $p^* = c_1 + \ldots + c_{[\alpha]-1} + c_{[\alpha]}\left(1 + \alpha - [\alpha]\right)$

В случае $1^\top x \leq \alpha$.
Если все компоненты вектора c неотрицательны, то понятно, что минимум достигается при $x^* = 0$ (так как отрицательное число мы получить не можем, а 0 достигается в этой точке).
Если же у c есть отрицательные компонеты, то понятно, что минимумом будет сумма первых α самых отрицательных компонент (в качестве x^* возьмём вектор с 1 в компонентах, где у вектора c самые отрицательные компоненты, а все остальные - 0)
Если отрицательных компанент меньше чем α то возьмём первые α по возрастанию

Задача 4

Give an explicit solution of the following QP.

$$\begin{aligned} c^\top x &\to \min_{x \in \mathbb{R}^n} \\ \text{s.t. } x^\top A x &\leq 1 \end{aligned}$$

where $A \in \mathbb{S}^n_{++}, c \neq 0$.
What is the solution if the problem is not convex $\left(A \notin \mathbb{S}^n_{++}\right)$ (Hint: consider eigendecomposition of the matrix: $A = Q \operatorname{diag}(\lambda) Q^\top = \sum_{i=1}^n \lambda_i q_i q_i^\top$) and different cases of $\lambda > 0, \lambda = 0, \lambda < 0$?

В случае, если матрица $A \in \mathbb{S}^n_{++}$
Лагранжиан для данной задачи ($\mu \in R$):

$$L(x, \mu) = c^T x + \mu\left(x^T A x - 1\right)$$

Условия ККТ:

$$\begin{cases} \nabla_x \mathrm{L}(x, \mu): \ c + \mu\left(A + A^\top\right) x = 0 \\ \mu \geqslant 0 \\ \mu\left(x^\top A x - 1\right) = 0 \\ x^\top A x - 1 \leqslant 0 \end{cases}$$

В этом случае условия ККТ будут достаточными, так как наша задача - выпуклая: функция та, которую мы оптимизируем - линейная (выпуклая), функция, которая задаёт ограничения типа неравенств - выпуклая (так как матрица положительно определена, то задаёт параболу). Также выполнено условие Слейтера: возьмём произвольный вектор x. Для него $l = x^T Ax \geq 0$, так как матрица положительно определена. Можем взять вектор $x' = \frac{x}{\sqrt{2\|l\|}}$, тогда $x'^T Ax' = \frac{1}{2}$. Для этой точки все ограничения типо неравенств выполнены строго. Следовательно, она принадлежит относительной внутренности бюджетного множества. Значит условия ККТ - достаточные.

Из первого условия и того, что $c \neq 0$ следует, что $\mu \neq 0$, значит, из 3го условия, $x^\top Ax = 1$. $A \in S_+^n \to A^T = A$, следовательное, первое условие можем переписать как: $c + 2\mu Ax = 0$. Выражаем отсюда x и подставляем в преобразованное 3ие условие:

$$x = -\frac{A^{-1}c}{2\mu}$$

$$x^\top Ax = \frac{c^\top A^{-1}c}{4\mu^2} = 1$$

10/10

Получаем значение μ:

$$\mu = \frac{1}{2}\sqrt{c^T A^{-1}c}$$

Таким образом решение нашей задачи в этом случае (так как ККТ - достаточные условия):

$$x^* = -\frac{A^{-1}c}{\sqrt{c^T A^{-1}c}}$$

А оптимальное значение: $p^* = -\sqrt{c^T A^{-1}c}$

В случае, если матрица A не положительно определена, тогда уже KKT не достаточное условие (так как задача перестаёт быть выпуклой). Из подсказки: воспользуемся спектральным разложением матрицы $A = \sum_{i=1}^{n} \lambda_i q_i q_i^\top$

Если все собственные значения больше 0, то есть $\lambda_i > 0,\ i = [1.n]$, то матрица положительно определена (по определению) и этот случай мы уже рассмотрели.

Если одно из собсвенных чисел равно 0 или отрицательно, то возьмём соответствующую компоненту вектора x бесконечно отрицательной (или

+

положительной, смотря какой c_i) и тогда $x^T A x \leq 1$, а минимум функции $-\infty$

+ !

Задача 5

Give an explicit solution of the following QP.

$$c^\top x \to \min_{x \in \mathbb{R}^n}$$
$$\text{s.t. } (x - x_c)^\top A (x - x_c) \leq 1$$

where $A \in \mathbb{S}^n_{++}, c \neq 0, x_c \in \mathbb{R}^n$

1ый способ:
Сделаем замену: $z = x - x_c$. Покажем, что задача

Хорошо!

$$c^\top z \to \min_{z \in \mathbb{R}^n}$$
$$\text{s.t. } z^\top A z \leq 1$$

эквивалентна исходной, действительно: $c^\top x = c^\top z + c^\top x_c$, следовательно, $c^\top x$ - минимально, когда $c^\top z$ - минимально. Если посмотреть на полученную задачу, то поймём, что это задача 4, которую уже решили и для неё можем записать ответ:

$$z^* = -\frac{A^{-1}c}{\sqrt{c^T A^{-1} c}}$$

а оптимальное значение (учитывая константу $c^\top x_c$)

$$p^* = c^T x_c - \sqrt{c^T A^{-1} c}$$

2ой способ:
Можно теми же преобразованиями, что и в 4ой задачи честно решить через ККТ
В этом случае Лагранжиан записывается как:

$$L(x, \mu) = c^T x + \mu \left((x - x_c)^T A (x - x_c) - 1 \right)$$

Условия ККТ (сразу учитывая, что в этой задаче $A \in \mathbb{S}^n_{++}$)

$$\begin{cases} \nabla_x \mathrm{L}(x, \mu) : \ c + 2\mu A (x - x_c) = 0 \\ \mu \geqslant 0 \\ \mu \left((x - x_c)^T A (x - x_c) - 1 \right) = 0 \\ (x - x_c)^T A (x - x_c) \leqslant 1 \end{cases}$$

Условия ККТ будут достаточными, такие же рассуждения, как и в задаче 4 с положительно определённой матрицей, при которой функция $(x-x_c)^\top A(x-x_c)-1$ - выпуклая.
Из 1го условия выражаем x:

$$x = x_c - \frac{1}{2\mu}A^{-1}c$$

Их 3го условия, так как $\mu \neq 0$ (иначе из 1го условия получим, что $c=0$, а это по условию не так), получаем:

$$(x-x_c)^T A(x-x_c) = 1$$

Подставляем сюда выраженный x и находим значение μ:

$$\mu = \frac{1}{2}\sqrt{c^T A^{-1} c}$$

10/10

Так как ККТ - достаточные условия, получаем решения задачи:

$$x^* = x_c - \frac{A^{-1}c}{\sqrt{c^T A^{-1} c}}$$

А оптимальное значение:

$$p^* = c^T x_c - \sqrt{c^T A^{-1} c}$$

Что совпадает с решением, полученным 1ым способом!

Задача 6

Give an explicit solution of the following QP.

$$x^\top Bx \to \min_{x\in\mathbb{R}^n}$$
$$\text{s.t. } x^\top Ax \le 1$$
$$\text{where } A \in \mathbb{S}^n_{++}, B \in \mathbb{S}^n_{+}.$$

По условию $B \in \mathbb{S}^n_+$, то есть положительно полуопределена. Это по определению значит, что для любого $x \in \mathbb{R}^n \to x^\top Bx \ge 0$. Понятно, что минимальное значение - 0 достигается при $x^* = 0$ и это значение лежит в бюджетном множестве: $x^{*\top}Ax^* = 0 < 1$, значит, мы решили задачу: оптимальное значение $p^* = 0$ при $x = 0$

10/10

Задача 7

Consider the equality constrained least-squares problem

$$\|Ax-b\|_2^2 \to \min_{x\in\mathbb{R}^n}$$
$$\text{s.t. } Cx = d$$

where $A \in \mathbb{R}^{m\times n}$ with $\operatorname{rank} A = n$, and $C \in \mathbb{C}^{k\times n}$ with $\operatorname{rank} C = k$. Give the KKT conditions, and derive expressions for the primal solution x^* and the dual solution λ^*.

Лагранжиан для данной задачи:

$$L(x,\lambda) = \|Ax-b\|_2^2 + \lambda^T(Cx-b)$$

Условия ККТ:

$$\begin{cases} \nabla_x \mathrm{L}(x,\lambda) : 2A^T(Ax-b) + C^T\lambda = 0 \\ \nabla_\lambda \mathrm{L}(x,\lambda) : Cx = d \end{cases}$$

Из 1го условия выражаем x:

$$2A^TAx = 2A^Tb - C^T\lambda \to x = \frac{1}{2}\left(A^TA\right)^{-1}\left(2A^Tb - C^T\lambda\right)$$

Из 2го условия выражаем $x = C^{-1}d$ и подставляем в полученное выражение, находим λ:

$$C^{-1}d = \left(A^TA\right)^{-1}A^Tb - \frac{1}{2}\left(A^TA\right)^{-1}C^T\lambda$$

$$\left(A^TA\right)^{-1}C^T\lambda = 2\left(\left(A^TA\right)^{-1}A^Tb - C^{-1}d\right)$$

Следовательно:

$$\lambda = 2\left(C\left(A^TA\right)^{-1}C^T\right)^{-1}\left(C\left(A^TA\right)^{-1}A^Tb - d\right)$$

Тогда, подставляя полученное λ в выражение для x, получим:

$$x = \frac{1}{2}\left(AA^T\right)^{-1}\left(2A^Tb - 2C^T\left(C\left(A^TA\right)^{-1}C^T\right)^{-1}\left(C\left(A^TA\right)^{-1}A^Tb - d\right)\right)$$

Данная задача выпуклая, так как функция, задающая ограничения типа равенств - линейная, а функцию, которую оптимизируем - выпуклая.

+

10/10

Если система $Cx = d$ несовместна, то бюджетное множество пустое и оптимальное значение принимаем за $p^* = \infty$
Иначе решение системы можно записать через псевдообратную матрицу $x = C^\dagger d$ и будет выполнено условие Слейтера (будет существовать допустимая точка, в которой все ограничения типа равенств выполнены, то есть она принадлежит относительной внутренности бюджетного множества), значит, в этом случае условия ККТ - достаточные и полученное ранее

$$x^* = \left(AA^T\right)^\dagger \left(A^T b - C^T \left(C\left(A^T A\right)^\dagger C^T\right)^\dagger \left(C\left(A^T A\right)^\dagger A^T b - d\right)\right)$$

-решение задачи (вместо обратных нужно брать псевдообратные матрицы, так как из условий задачи обратные могут не существовать. Все выкладки с псевдообратными совпадают с просто обратной)

Задача 8

Derive the KKT conditions for the problem

$$\begin{aligned} &\operatorname{tr} X - \log\det X \to \min_{X \in \mathbb{S}^n_{++}} \\ &\text{s.t. } Xs = y \end{aligned}$$

where $y \in \mathbb{R}^n$ and $s \in \mathbb{R}^n$ are given with $y^\top s = 1$. Verify that the optimal solution is given by

$$X^* = I + yy^\top - \frac{1}{s^\top s} ss^\top$$

Лагранжиан для данной задачи:

$$L(X, \lambda) = \operatorname{tr} X - \log\det X + \lambda^T (Xs - y)$$

Условия ККТ:

$$\begin{cases} \nabla_x \mathrm{L}(x, \lambda) : I - X^{-1} + \lambda s^T = 0 \\ \nabla_\lambda \mathrm{L}(x, \lambda) : Xs = y \end{cases}$$

Условия ККТ в нашем случае являются достаточными, так как задача выпуклая: ограничения типа равенств - линейные, а функцию, которую минимизируем - выпуклая (из первого задания). Также выполнено условие Слейтера: можем подобрать такую матрицу $X \in \mathbb{S}^n_{++}$, что для неё выполнено $Xs = y$, значит есть допустимая точка.

Из второго условия ККТ: $s = X^{-1} y$. Первое условие можем переписать как:

$$I - X^{-1} + \frac{1}{2}\left(\lambda s^T + s\lambda^T\right) = 0$$

Домножая его справа на y и пользуясь тем, что $s^T y = 1$, получим:

$$y - s + \frac{1}{2}\left(\lambda + s\lambda^T y\right) = 0$$

Домножим полученное выражение слева на y^T и снова используя условие, что $s^T y = 1$, получим:

$$1 = y^T y + \lambda^T y \rightarrow \lambda^T y = 1 - y^T y$$

Подставляем в предыдущие уравнение и выражаем λ:

$$\lambda = -2y + \left(1 + y^T y\right) s$$

Подставляем в 1ое условие ККТ и находим выражение для X^{-1}:

$$X^{-1} = I + \left(1 + y^T y\right) ss^T - ys^T - sy^T$$

Чтобы матрица X^* задавала решение (так как ККТ - достаточные условия в этой задачи), то нужно проверить, что $X^{-1}X^* = \mathrm{I}$. Делаем это подстановкой и раскрытием скобок:

$$X^{-1}X^* = \left(\mathrm{I} - \mathrm{ys^T} - \mathrm{sy^T} + \left(1 + \mathrm{y^T y}\right)\mathrm{ss^T}\right)\left(\mathrm{I} + \mathrm{yy^T} - \left(\frac{1}{\mathrm{s^T s}}\right)\mathrm{ss^T}\right) =$$

$$= \left(\mathrm{I} + \mathrm{yy^T} - \left(\frac{1}{\mathrm{s^T s}}\right)\mathrm{ss^T}\right) + \left(1 + \mathrm{y^T y}\right)\mathrm{ss^T} + \left(1 + \mathrm{y^T y}\right)\mathrm{sy^T} - \left(1 + \mathrm{y^T y}\right)\mathrm{ss}^\top -$$

$$-ys^\top - yy^\top + ys^\top - sy^\top - \left(y^\top y\right)sy^\top + ss^\top\left(\frac{1}{s^\top s}\right) = I$$

Ещё нужно проверить, что X^* - положительно определена. По определению $\forall x$:

$$x^\top\left(\mathrm{I} + yy^\top - \left(\frac{1}{s^\top s}\right)ss^\top\right)x = x^\top x + \left(y^\top x\right)^\top\left(y^\top x\right) - \left(s^\top x\right)^\top\left(s^\top x\right)\left(\left(\frac{1}{s^\top s}\right)\right)$$

$$= \|x\|^2 + (y, x)^2 - \frac{(s, x)^2}{\|s\|^2}$$

Понятно, что $\|x\|^2 - (s, x)^2/\|s\|^2 \geq 0$, следовательно, показали, что $\forall x$: $x^\top X^* x \geq 0$, значит X^* положительно определена и следовательно является решением поставленной задачи.

10/10

Задача 9

Supporting hyperplane interpretation of KKT conditions. Consider a convex problem with no equality constraints

$$
\begin{array}{c}
f_0(x) \to \min_{x \in \mathbb{R}^n} \\
\text{s.t. } f_i(x) \leq 0, \quad i = [1, m]
\end{array}
$$

Assume, that $\exists x^* \in \mathbb{R}^n, \mu^* \in \mathbb{R}^m$ satisfy the KKT conditions

$$
\begin{array}{l}
\nabla_x L\left(x^*, \mu^*\right) = \nabla f_0\left(x^*\right) + \sum_{i=1}^m \mu_i^* \nabla f_i\left(x^*\right) = 0 \\
\mu_i^* \geq 0, \quad i = [1, m] \\
\mu_i^* f_i\left(x^*\right) = 0, \quad i = [1, m] \\
f_i\left(x^*\right) \leq 0, \quad i = [1, m]
\end{array}
$$

Show that

$$
\nabla f_0\left(x^*\right)^\top \left(x - x^*\right) \geq 0
$$

for all feasible x. In other words the KKT conditions imply the simple optimality criterion or $\nabla f_0\left(x^*\right)$ defines a supporting hyperplane to the feasible set at x^*

Возьмём x - допустимое и используя условия ККТ покажем, что $\nabla f_0\left(x^*\right)^\top \left(x - x^*\right) \geq 0$: буду сводить к 1му дифференциальному критерию выпуклой функции (а у нас выпукаля функция f_0 по условию), который записывается как:

$$
f_0(x) \geq f_0(x^*) + \nabla f_0^T(x^*)(x - x^*)
$$

Из условия, что градиент по x от Лагранжиана равен нулю, получаем:

$$
\nabla f_0\left(x^*\right)^\top \left(x - x^*\right) = -\sum_{i=1}^m \mu_i^* \nabla f_i^T\left(x^*\right)\left(x - x^*\right)
$$

так как $f_i(x) \leq 0, \quad i = [1, m]$ (по постановке задачи), а из условий ККТ:
$\mu_i^* \geq 0, \quad i = [1, m]$
$\mu_i^* f_i\left(x^*\right) = 0, \quad i = [1, m]$, то можно записать следующие неравенство:

$$
\nabla f_0\left(x^*\right)^\top \left(x - x^*\right) =
$$

$$
= -\sum_{i=1}^m \mu_i^* \nabla f_i^T\left(x^*\right)\left(x - x^*\right) \geqslant -\sum_{i=1}^m \mu_i^* \nabla f_i^\top\left(x^*\right)\left(x - x^*\right) + \sum_{i=1}^m \mu_i^*\left(f_i(x) - f_i\left(x^*\right)\right) =
$$

10/10

$$= \sum_{j=1}^{m} \mu_i^* \left((f_i(x) - f_i(x^*)) - \nabla f_i^\top (x^*) (x - x^*) \right) \geqslant 0$$

Так как как раз в скобках получили дифференциальный критерий 1го порядка для выпуклой функции. Следовательно, доказали, что для любой допустимой x выполнено:

$$\nabla f_0 (x^*)^\top (x - x^*) \geq 0$$

90/90

Duality

Задача 1

Fenchel + Lagrange = ♡. Express the dual problem of

$$c^\top x \to \min_{x\in\mathbb{R}}$$
$$\text{s.t. } f(x) \leq 0$$

with $c \neq 0$, in terms of the conjugate function f^*. Explain why the problem you give is convex. We do not assume f is convex.

Лагранжиан для этой задачи:

$$L(x,\mu) = c^T x + \mu f(x)$$

Записываем по определению двойственную функцию и сводим к сопряжённой (в следующей задаче вывел для более общего случая):

$$g(\mu) = \inf_x \left(c^T x + \mu f(x)\right) = -\mu \sup \left(-\frac{c^T x}{\mu} - f(x)\right) = -\mu f^* \left(-\frac{c}{\mu}\right)$$

+

Область определения двойственной функции задаётся областью определения сопряжённой, следовательно dual problem записываем так:

$$\mu f^* \left(-\frac{c}{\mu}\right) \to \min_\mu$$
$$\text{s.t. } \mu \geq 0,\ -\frac{c^\top}{\mu} \in \operatorname{dom} f^*$$

хорошо!

10/10

Знаем, что двойственная функция всегда вогнута. При домножении на неотрицательное число $\mu \geq 0$ (из области определения) вогнутость не потеряется, значит, получаем выпуклую задачу (вогнутая функция со знаком минус - выпуклая и мы её максимизируем). Ограничения типа неравенств - линейные, то есть выпуклые.

Задача 2

Minimum volume covering ellipsoid.
Let we have the primal problem:

$$\ln\det X^{-1} \to \min_{X\in\mathbb{S}^n_{++}}$$
$$\text{s.t. } a_i^\top X a_i \leq 1, i = 1,\dots,m$$

1. Find Lagrangian of the primal problem
2. Find the dual function
3. Write down the dual problem
4. Check whether problem holds strong duality or not
5. Write down the solution of the dual problem

10/10

Заметим, что ограничения типа неравенств в нашей задаче аффины и их можно переписать как:

$$\operatorname{tr}\left(\left(a_i a_i^T\right) X\right) \leq 1$$

Теперь запишем Лагранжиан для поставленной задачи

$$\mathrm{L}(\mathrm{X}, \mu) = \ln\det \mathrm{X}^{-1} + \sum_{i=1}^{\mathrm{m}} \lambda_{\mathrm{i}} \left(\mathrm{a}_{\mathrm{i}}^{\mathrm{T}} \mathrm{X} \mathrm{a}_{\mathrm{i}} - 1\right)$$

На семинаре разбирали, как записывать dual function через сопряжённую функцию, а сопряжённую функцию к $\ln\det X^{-1}$ уже находили в 5ой задаче 2го домашнего задания и получили:

$$f_0^*(Y) = \log\det(-Y)^{-1} - n$$

Повторим эти небольшие выкладки из семинара для задачи в общей форме, а потом используем полученный результат к нашей проблеме:

$$\begin{array}{rl} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b \\ & Cx = d \end{array}$$

$$\begin{aligned} g(\lambda, \nu) &= \inf_x \left(f_0(x) + \lambda^T(Ax - b) + \nu^T(Cx - d)\right) \\ &= -b^T\lambda - d^T\nu + \inf_x \left(f_0(x) + \left(A^T\lambda + C^T\nu\right)^T x\right) \\ &= -b^T\lambda - d^T\nu - f_0^*\left(-A^T\lambda - C^T\nu\right) \end{aligned}$$

Область определения двойственной функции определяется областью определения сопряжённой к исходной функции:

$$\operatorname{dom} g = \left\{(\lambda, \nu) \mid -A^T\lambda - C^T\nu \in \operatorname{dom} f_0^*\right\}$$

В нашем случае оганичений типа неравенств нет, вектор $b = \mathbf{1}^T$ матрица $A = a_i a_i^\top$, dom $f_0^* = -\mathbf{S}_{++}^n$ Получается dual function:

$$g(\lambda) = \begin{cases} \log\det\left(\sum_{i=1}^m \lambda_i a_i a_i^T\right) - \mathbf{1}^T\lambda + n & \sum_{i=1}^m \lambda_i a_i a_i^T \succ 0 \\ -\infty & \text{иначе} \end{cases}$$

Следовательно двойственную задачу мы можем записать как:

$$\begin{array}{ll} \text{maximize} & \log\det\left(\sum_{i=1}^{m}\lambda_i a_i a_i^T\right) - \mathbf{1}^T\lambda + n \\ \text{s.t.} & \lambda \succeq 0 \end{array}$$

В этой задаче выполняется условие Слейтера: всегда существует $X \in \mathbf{S}^n_{++}$ для которой $a_i^T X a_i < 1$, для $i = 1, \ldots, m$ то есть есть допустимая точка. Значит, в этой задаче есть сильная двойственность между прямой и двойственной задачами

Задача 3

A penalty method for equality constraints. We consider the problem minimize

$$f_0(x) \to \min_{x\in\mathbb{R}^n}$$
$$\text{s.t. } Ax = b$$

where $f_0(x) : \mathbb{R}^n \to \mathbb{R}$ is convex and differentiable, and

10/10

$$\alpha \in \mathbb{R}^{m\times n}$$

with $\operatorname{rank} A = m$. In a quadratic penalty method, we form an auxiliary function $\phi(x) = f_0(x)+\alpha\|Ax-b\|_2^2$ where $\alpha > 0$ is a parameter. This auxiliary function consists of the objective plus the penalty term $\alpha\|Ax-b\|_2^2$. The idea is that a minimizer of the auxiliary function, $\tilde{x}$, should be an approximate solution of the original problem. Intuition suggests that the larger the penalty weight α, the better the approximation $\tilde{x}$ to a solution of the original problem. Suppose $\tilde{x}$ is a minimizer of $\phi(x)$. Show how to find, from $\tilde{x}$, a dual feasible point for the original problem. Find the corresponding lower bound on the optimal value of the original problem.

Запишем Лагранжиан для данной задачи:

$$L(x,\lambda) = f_0(x) + \lambda^T(Ax-b)$$

В точке $\tilde{x}$ функция $\phi(x)$ достигает своего минимального значения, следовательно, по необходимому условию экстремума:

$$\nabla\phi(\tilde{x}) = 0 \Leftrightarrow \nabla f_0(\tilde{x}) + 2\alpha A^T(A\tilde{x}-b) = 0$$

Посмотрим, в какой точке достигается минимум Лагранжиана:

$$\nabla L(x,\lambda) = \nabla f_0(x) + A^\top\lambda$$

Заметим, что при $\lambda^* = 2\alpha(A\tilde{x} - b)$, мы получаем выражение из градиента ϕ, которое равно 0. Следовательно, (так как f_0 - выпуклая по условию, а $(Ax - b)$ - линейная, то есть Лагранжиан выпуклая функция) то, равенство градиента нулю - достаточное условие экстремума и $L(x, \lambda^*)$ достигает минимума на $\tilde{x}$.
По определению двойственной функции:

$$g(\lambda) = \inf_x \left(f_0(x) + \lambda(Ax - b)\right) = \inf_x (L(x, \lambda))$$

$$g(\tilde{\lambda}) = f_0(\tilde{x}) + \alpha\|A\tilde{x} - b\|_2^2$$

Получается, для таких $x: \ Ax = b$ выполнено:

$$f_0(x) \geqslant g(\tilde{\lambda}) = f_0(\tilde{x}) + \alpha\|A\tilde{x} - b\|_2^2$$

Получили нижнюю границу на оптимальное значение прямой задачи.

Задача 4

Analytic centering. Derive a dual problem for

$$-\sum_{i=1}^m \log\left(b_i - a_i^\top x\right) \to \min_{x \in \mathbb{R}^n}$$

with domain $\left\{x \mid a_i^\top x < b_i, i = [1, m]\right\}$. First introduce new variables y_i and equality constraints $y_i = b_i - a_i^\top x$. (The solution of this problem is called the analytic center of the linear inequalities $a_i^\top x \leq b_i, i = [1, m]$. Analytic centers have geometric applications, and play an important role in barrier methods.) with domain $\left\{x \mid a_i^\top x < b_i, i = [1, m]\right\}$.

Пользуясь подсказкой сделаем замену переменных: $y_i = b_i - a_i^\top x$ или можно её записать как $y = b - Ax$, где матрица $A \in \mathbb{R}^{m \times n}$ где a_i^T строки у матрицы A
Тогда получим задачу:

$$\begin{aligned} &-\textstyle\sum_{i=1}^m \log(y_i) \to \min \\ &\text{s.t. } y = b - Ax \end{aligned}$$

Лагранжиан полученной задачи:

$$\mathrm{L}(x, y, \mu) = -\sum_{i=1}^m \log y_i + \mu^\top (y - b + Ax)$$

По определению двойственная функция:

$$g(\mu) = \inf_{x,y} \left(-\sum_{i=1}^{m} \log{(y_i)} + \mu^T (y - b - Ax) \right)$$

Слагаемое $\mu^\top Ax$ неограничено снизу по x, то есть всегда можем взять такой x что оно уйдёт в $-\infty$, следовательно, чтобы минимум был не $-\infty$, нужно $A^\top \mu = 0$

Для нахождения минимума приравниваем градиент к нулю (то что это минимум легко проверить взяв ещё раз градиент, получим сумму $\sum\limits_{i=1}^{m} \left(\frac{1}{y_i}^2\right) \geq 0$, следовательно, минимум). Минимум достигается при $y_i = \frac{1}{\mu_i}$

Так как $y_i > 0$ (стоят под логарифмом), то область определения двойственной функции:

$$A^\top \mu = 0 \text{ и } \mu \succ 0$$

а сама двойственная функция:

$$g(\mu) = \begin{cases} \sum\limits_{i=1}^{m} \log \mu_i + m - \mu^\top b, & A^\top \mu = 0,\ \mu \succ 0 \\ -\infty, \text{ иначе} \end{cases}$$

Тогда двойственная задача имеет вид:

$$\sum_{i=1}^{m} \log{(\mu_i)} + m - \mu^\top b \to \max_{\mu}$$
$$\text{s.t.} \quad A^T \mu = 0, \quad \mu \succ 0$$

10/10

Прекрасная работа!

100