HW2: выпуклая оптимизация, двойственность, условия Каруша-Куна-Таккера

Семинарист: Курузов Илья

Дедлайн: 23:59, 18.10.2021

1 Выпуклая оптимизация

1. [2] Рассмотрим задачу аппроксимации по ℓ_p норме

$$\min_{\mathbf{x}\in\mathbb{C}^n} \|A\mathbf{x}-\mathbf{b}\|_p$$

с комплексными параметрами $A \in \mathbb{C}^{m\times n}, \mathbf{b} \in \mathbb{C}^m$ и комплекснозначной переменной $\mathbf{x} \in \mathbb{C}^n$. Выразите задачу аппроксимации по ℓ_p норме для $p = 1, 2, \infty$ как QCQP или SOCP задачу с вещественными переменными и параметрами.

Напомним, что норма ℓ_p определяется как $\|\mathbf{x}\|_p = \left(\sum\limits_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}$ для $p \geq 1$ и $\|\mathbf{x}\|_\infty = \max_{i=\overline{1,n}} |x_i|$ как предельный случай для $p = \infty$.

2. [5] В данной задаче Вам предлагается привести задачу робастного квадратичного программирования к стандартной форме. Сама задача квадратичного робастного программирования формулируется следующим образом:

$$\begin{aligned}&\min_{\mathbf{x}\in\mathbb{R}^n} \sup_{P\in\mathcal{P}} \left(\mathbf{x}^\top P\mathbf{x} + \mathbf{b}^\top \mathbf{x} + c\right)\\ &\text{s.t. } A\mathbf{x} \leq \mathbf{d},\end{aligned} \tag{1}$$

где $A \in \mathbf{R}^{m\times n}, \mathbf{b} \in \mathbf{R}^n, \mathbf{d} \in \mathbf{R}^m$ - параетры задачи. Множество $\mathcal{P} \subseteq \mathbb{S}^n_+$ есть некоторое множество неотрицательно определенных матриц. Приведите задачу к стандартному виду (например, QP, QCQP, SCP, SDP) для следующих множеств $\mathcal{P}$:

   a) [1] Конечное множество матриц: $\mathcal{P} = \{P_1 \dots P_k\}$, где $P_j \in \mathbb{S}^n_+, j = \overline{1,k}$.

   b) [2] Множество, заданное как ограничение на собственное число матрицы $P - P_0$:

$$\mathcal{P} = \left\{P \in \mathbb{S}^n_+ \middle| -\gamma I \preceq P - P_0 \preceq \gamma I\right\}$$

   для некоторого $\gamma \in \mathbb{R}_+$ и $P_0 \in \mathbb{S}^n_+$ (матрица I есть единичная матрица).

   c) [2] Эллипсоид матриц:

$$\mathcal{P} = \left\{P_0 + \sum_{i=1}^k P_i u_i \middle| \forall \mathbf{u} \in \mathbb{R}^k, \|\mathbf{u}\|_2 \leq 1\right\},$$

   где $P_i \in \mathbb{S}^n_+, i = \overline{0,k}$ есть неотрицательно определенные матрицы.

Постарайтесь свести задачу к более узкому классу. Например, если Вы получили задачу в форме SDP и можете свести ее к QCQP, то сведите.

3. Выразите следующие проблемы в форме задачи геометрического программирования:

   a) [1] $\min\limits_{\mathbf{x}\in\mathbb{R}^n_+} \exp(p(\mathbf{x})) + \exp(q(\mathbf{x}))$, где p, q - позиномы.

   b) [1] $\min\limits_{\mathbf{x}\in\mathbb{R}^n_+} \frac{p(\mathbf{x})}{r(\mathbf{x}) - q(\mathbf{x})}$ при ограничении $r(\mathbf{x}) \geq q(\mathbf{x})$, где p - позином, q, r - мономы.

2 Двойственность

4. [2] Выведите двойственную задачу для задачи:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|A\mathbf{x}-\mathbf{b}\|_2^2 + \frac{\lambda}{2}\|\mathbf{x}\|_2^2,$$

введя новую переменную $\mathbf{y} \in \mathbb{R}^m$ $\mathbf{y} = A\mathbf{x} - \mathbf{b}$ и соответствующие ограничения. Параметры: $A \in \mathbb{R}^{m\times n}$, $\lambda \geq 0$. Таким образом, мы получим задачу, которая дает нижнюю оценку на решение для задачи безусловной минимизации.

5. [3] Рассмотрим задачу бинарного линейного программирования:

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} \mathbf{c}^\top \mathbf{x} \\ &\text{s.t. } A\mathbf{x} \leq \mathbf{b} \\ &x_i(1-x_i) = 0, i = \overline{1,n} \end{aligned}$$

с непустым допустимым множеством.

a) [1] Постройте двойственную задачу. Данная двойственная задача будет давать оценку снизу на оптимальное значение исходной дискретной задачи. Такой метод релаксации называется релаксацией Лагранжиана.

b) [1] Упростите полученную в предыдущем пункте задачу, аналитически решив ее по одному из блоков переменных.

c) [1] Докажите, что оценка, которую дает двойственная задача, совпадает с решением релаксации этой задачи:

$$\begin{aligned} &\mathbf{c}^\top \mathbf{x} \to \min_{x\in\mathbb{R}^n} \\ &\text{s.t. } A\mathbf{x} \leq \mathbf{b} \\ &0 \leq x_i \leq 1, i = \overline{1,n}. \end{aligned}$$

Hint: Один из возможных способов сделать это - вывести двойственную задачу к этой релаксации и показать эквивалентность этой двойственной задачи и полученной в пункте b.

6. [2] Рассмотрим следующую модификации задачи Optimal Experiment Design:

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^p} \operatorname{tr}\left(\sum_{i=1}^p x_i \mathbf{v}_i \mathbf{v}_i^\top\right)^{-1} \\ &\text{s.t. } \mathbf{x} \geq \mathbf{0}, \quad \mathbf{1}^\top \mathbf{x} = 1. \end{aligned}$$

Данная задача называется A-optimal Design. Область определения функции $\{\mathbf{x} \in \mathbb{R}^p | \sum_{i=1}^n x_i \mathbf{v}_i \mathbf{v}_i^\top \in \mathbb{S}^n_{++}\}$. Параметрами данной задачи являются p вектором $\mathbf{v}_1 \dots \mathbf{v}_p \in \mathbb{R}^n$. Введите новую переменную $X \in \mathbb{S}^n$ и ограничение $X = \sum_{i=1}^n x_i \mathbf{v}_i \mathbf{v}_i^\top$. Выведите для новой проблемы двойственную задачу. Упростите её, насколько сможете.

Hint: Можно пользоваться без доказательства, что $\nabla_X \operatorname{tr} X^{-1} = -X^{-2}$ для $X \in \mathbb{S}^n$.

3 Условия ККТ

7. a) [2] Пусть параметры $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n_{++}$ имеют положительные компоненты, при этом компоненты первого вектора отсортированы в порядке убывания $a_n \geq a_k \geq \dots a_1 > 0$, а компоненты второго вектора определены, как $b_k = \frac{1}{a_k}$. Выведите условия ККТ для задачи

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^n} -\log(\mathbf{a}^\top \mathbf{x}) - \log(\mathbf{b}^\top \mathbf{x}) \\ &\text{s.t. } \mathbf{x} \geq 0, \quad \mathbf{1}^\top \mathbf{x} = 1 \end{aligned}$$

и покажите, что вектор $\mathbf{x} = \left(\frac{1}{2}, 0, 0 \dots, 0, 0, \frac{1}{2}\right)^\top$ является решением этой задачи.

b) [1] Пусть $A \in \mathbb{S}^n_{++}$. Примените результат первой части задачи для вектора $\mathbf{a} = (\lambda_n \dots \lambda_1)^\top$, собственные значения в котором расположены в порядке убывания, чтобы доказать неравенство Канторовича:

$$2\left(\mathbf{u}^\top A \mathbf{u}\right)^{\frac{1}{2}}\left(\mathbf{u}^\top A^{-1} \mathbf{u}\right)^{\frac{1}{2}} \leq \sqrt{\frac{\lambda_n}{\lambda_1}} + \sqrt{\frac{\lambda_1}{\lambda_n}},$$

где вектор $\mathbf{u} \in \mathbb{R}^n$, $\|\mathbf{u}\|_2 = 1$.

Hint: Если $A\mathbf{v} = \lambda \mathbf{v}$ и A обратимая матрица, то $A^{-1}\mathbf{v} = \frac{1}{\lambda}\mathbf{v}$.

8. [2] Выпишите условия ККТ для следующей задачи:

$$\begin{aligned} &\|A\mathbf{x} - \mathbf{b}\|_2^2 \to \min_{x \in \mathbb{R}^n} \\ &\text{s.t. } G\mathbf{x} = \mathbf{h}, \end{aligned}$$

$A \in \mathbb{R}^{m \times n}, \operatorname{rank} A = n, G \in \mathbb{R}^{p \times n}, \operatorname{rank} G = p$ и найдите выражения для решения прямой задачи $\mathbf{x}^*$ и двойственной ν^* (одна из этих точек может выражаться через другую).

9. [4*] Дивергенцией Кульбака-Лейбера между двумя распределениями $\mathcal{P}_1$ и $\mathcal{P}_2$ с абсолютно непрерывными распределениями и функциями распределения p_1 и p_2 называется величина:

$$D_{KL}(\mathcal{P}_1 || \mathcal{P}_2) = \mathbb{E}_{x \sim \mathcal{P}_1}\left[\log \frac{p_1(x)}{p_2(x)}\right] = \int\limits_{x \in X} p_1(x) \log \frac{p_1(x)}{p_2(x)} dx.$$

a) [1] Покажите, что если распределения $\mathcal{P}_1 = \mathcal{N}(\mu_1, \Sigma_1)$, $\mathcal{P}_2 = \mathcal{N}(\mu_2, \Sigma_2)$ есть многомерные нормальные распределения с невырожденными матрицами корреляции $\Sigma_{1,2} \in \mathbb{S}^n_{++}$, то дивергенция KL примет вид:

$$2D_{KL}(\mu_1, \Sigma_1 || \mu_2, \Sigma_2) = \operatorname{tr}\left(\Sigma_2^{-1}\Sigma_1\right) - \log\det\left(\Sigma_2^{-1}\Sigma_1\right) + (\mu_1 - \mu_0)^\top \Sigma_2^{-1}(\mu_1 - \mu_0) - n$$

Hint: Пара трюков со следом: $\mathbf{a}^\top A \mathbf{a} = \operatorname{tr}(\mathbf{a}^\top A \mathbf{a}) = \operatorname{tr}(\mathbf{a}\mathbf{a}^\top A)$; $\mathbb{E}[\operatorname{tr}(AB)] = \operatorname{tr}\left(\mathbb{E}[AB]\right)$.

b) [2] Рассмотрим следующую задачу минимизации дивергенции Кульбака-Лейбера между нормальными распределениями при линейных ограничениях:

$$\begin{aligned} &\min_{X \in \mathbb{S}^n_{++}} 2D_{KL}(\mathbf{0}, X || \mathbf{0}, \Sigma) = \operatorname{tr}\left(\Sigma^{-1}X\right) - \log\det\left(\Sigma^{-1}X\right) - n \\ &\text{s.t. } X\mathbf{s} = \mathbf{y} \end{aligned}$$

где параметры имеют следующие размерности $A \in \mathbb{R}^{m \times n}, \Sigma \in \mathbb{S}^n_{++}, \mathbf{s}, \mathbf{y} \in \mathbb{R}^n$.Покажите, что оптимальное X дается формулой:

$$X = I + \mathbf{y}\mathbf{y}^\top - \frac{\mathbf{s}\mathbf{s}^\top}{\mathbf{s}^\top\mathbf{s}} = \left(I + \frac{\mathbf{y}\mathbf{s}^\top}{\|\mathbf{s}\|_2} - \frac{\mathbf{s}\mathbf{s}^\top}{\|\mathbf{s}\|_2}\right)\left(I + \frac{\mathbf{y}\mathbf{s}^\top}{\|\mathbf{s}\|_2^2} - \frac{\mathbf{s}\mathbf{s}^\top}{\|\mathbf{s}\|_2^2}\right)^\top$$

в случае, если $\Sigma = I$ и $\mathbf{s}^\top\mathbf{y} = 1$.

c) [1] Покажите, как свести задачу минимизации Кульбака-Лейбера с произвольной матрицей Σ и векторами $\mathbf{s}, \mathbf{y}$ такими, что $\mathbf{s}^\top\mathbf{y} > 0$ к частному случаю, описанному в предыдущем случае. Выведите решение в этом случае, используя результат предыдущего пункта.