Домашнее задание № 1

Александр Катруца

Дедлайн: 5 октября 2020 г., 02:00

В задачах 1 — 5 необходимо формализовать задачу, записав её оптимизационную постановку, и объяснить, почему оптимизационная задача выглядит именно так. Решать выписанные задачи не нужно! Не забудьте ввести все необходимые обозначения перед тем, как записывать задачу. Также в случае отсутствия необходимой информации, Вы можете вводить новые обозначения.

1. Пункт назначения (1 pts)

Пусть известны приблизительные расстояния между текущим положением путешественника и m объектами, координаты которых заданы в некоторой системе координат и известны точно. Поставьте оптимизационную задачу для определения координат путешественника[1].

2. Посели меня, если сможешь (1.5 pts)

Пусть n студентов надо расселить в m комнатах, $m < n$, в каждой комнате могут одновременно жить максимум 3 студента. Предпочтение студента i жить со студентом j задано и равно p_{ij}, чем выше p_{ij}, тем более охотно студент i будет жить со студентом j. Также известно, что удовлетворённость студента i от проживания в комнате k равна b_{ik} и чем она выше, тем более охотно студент будет там проживать. Очевидно, что каждый студент может быть поселён только в одну комнату. Необходимо определить кого в какую комнату необходимо поселить? Рассмотрите случай, когда известны не все значения p_{ij}. Как изменится постановка задачи в этом случае?

3. Задача про магазины и склады (2 pts)

Пусть на i-ом складе сети М располагается a_i единиц товара, который необходимо развезти по m магазинам сети. Каждому j-ому магазину необходимо b_j единиц товара. Перевозка единицы товара из i-го склада в j-ый магазин стоит c_{ij} и занимает t_{ij} единиц времени. Сколько единиц товара с каждого склада нужно вывезти, чтобы обеспечить товаром все магазины сети? Подумайте как в этой задаче определить решение. Будет ли оно единственным и почему?

[1] Аналогичным образом работает GPS

4. Задача про определение положения базовой станции (1 pts)

Даны координаты пользователей, которые должны покрываться новой базовой станцией. Необходимо определить, где её расположить и как настроить так, чтобы все пользователи могли пользоваться мобильной связью. Поставьте задачу в предположении, что охват базовой станции представляет из себя эллипс. Параметры эллипса задаются в процессе настройки. Также учтите, что для работы станции требуется определённое количество энергии.

5. Поиск правильной метрики (3 pts)

- (1 pts) Каким образом можно параметризовать метрику в $\mathbb{R}^n$?
- (2 pts) Пусть дана выборка объекта двух классов. Необходимо найти такую метрику, чтобы объекты одного класса были как можно ближе, а объекты разных классов были как можно дальше.

6. Напоминание фактов из линейной алгебры (7 pts)

- (1 pts) Что такое норма вектора? Что такое эквивалентность норм? Покажите эквивалентность ℓ_1, ℓ_2 и ℓ_∞ норм.
- (2 pts) Что такое норма матрицы (линейного конечномерного оператора) порождённая векторной нормой? Выведите выражения для L_1, L_2, L_∞ нормы матрицы? Что такое Фробениусова норма матрицы и какой аналог среди векторных норм она имеет?
- (3.5 pts) Что такое разреженная матрица? Какие существуют форматы хранения таких матриц в памяти, опишите их структуру и способ доступа к элементу матрицы? Какие преимущества даёт работа с разреженными матрицами при вычислениях? В Jupyter Notebook файле сравните время умножение плотной и разреженной квадратной матрицы на вектор, а также покажите как время умножения зависит от размерности матрицы и заполненности разреженной матрицы. В результате должен получится график в осях время vs. размерность, на котором будут линии для плотной матрицы и для разреженных матриц с разными показателями разреженности (отношением числа ненулей к общему числу элементов). Для реализации используйте готовые объекты из библиотек NumPy и SciPy. График должен быть хорошо читаем: подписи к осям и легенда читаемы, линии не должны сливаться, масштаб (линейный или логарифмический) оси ординат должен быть выбран подходящим для получившихся данных образом. График необходимо сохранить в pdf-файл и вставить в решение, также приложите Jupyter Notebook файл для его воспроизведения
- (0.5 pts) Докажите ассоциативность матричного умножения

7. Диаграмма Вороного (3.5 pts)

Пусть $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^n$. Рассмотрим множество точек, которые ближе к точке $\mathbf{x}_0$, чем к точкам $\mathbf{x}_1, \dots, \mathbf{x}_k$:

$$V = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\|_2 \leq \|\mathbf{x} - \mathbf{x}_i\|_2,\ i = 1, \dots, k\}$$

Множество V называется областью Вороного для точки $\mathbf{x}_0$.

- (1 pts) Докажите, что множество V является многоугольником. Представьте его в виде $\{\mathbf{x} \in\ \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq\ \mathbf{b}\}$. Обозначение $\mathbf{x} \leq \mathbf{y}$ означает, что $x_i \leq y_i,\ i = 1, \dots, n$.

- (1 pts) Обратно, покажите как по данному многоугольнику восстановить точки $\mathbf{x}_0, \dots, \mathbf{x}_k$, для которых он определяет область Вороного.

- (1 pts) Сделайте обобщение областей Вороного для $p > 1$ точек и покажите, как по данному разбиению пространства на многоугольники (возможно открытые) восстановить множество точек, для которого это разбиение является разбиением Вороного.

- (0.5 pts) Выберите область науки и технологии, которая Вам интересна и опишите, как в ней можно использовать разбиение Вороного. Для вдохновения можете посмотреть статью в Википедии[2]

8. Выпуклый, аффинный, конический (2 pts)

Определите являются ли следующие множества выпуклыми, аффинными или конусами и объясните почему?

1. (0.2 pts) $\{\mathbf{x} \in \mathbb{R}^n \mid \alpha \leq \mathbf{a}^\top \mathbf{x} \leq \beta\}$
2. (0.2 pts) $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_1^\top \mathbf{x} \leq b_1,\ \mathbf{a}_2^\top \mathbf{x} \leq b_2\}$
3. (0.4 pts) $\{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_0\|_2 \leq \|\mathbf{x} - \mathbf{y}\|_2,\ \forall \mathbf{y} \in S \subseteq \mathbb{R}^n\}$
4. (0.4 pts) Множество точек, расстояние от которых до $\mathbf{a}$ не превышает доли $\theta \in [0, 1]$ от расстояния до точки $\mathbf{b}$: $\{\mathbf{x} \mid \|\mathbf{x} - \mathbf{a}\|_2 \leq \theta \|\mathbf{x} - \mathbf{b}\|_2\},\ \mathbf{a} \neq \mathbf{b}$

Также проверьте, верны ли следующие утверждения:

1. (0.4 pts) Пересечение любого (конечного или бесконечного) числа выпуклых конусов является конусом
2. (0.4 pts) Пересечение любого (конечного или бесконечного) числа аффинных множеств является аффинным множеством

9. Сумма Минковского (1 pts)

1. (0.5 pts) Найдите множество $M = M_1 + M_2$, где $M_1 = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$, а $M_2 = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 4\}$
2. (0.5 pts) Приведите пример множеств A и B таких что хотя бы одно из них невыпукло, но их сумма $A + B$ выпукла.

[2] https://en.wikipedia.org/wiki/Voronoi_diagram#Applications

10. Сопряги это (2 pts)

Найдите сопряжённые множества к следующим множествам:

1. (0.1 pts) $\{0\}$
2. (0.2 pts) $\{\mathbf{A}\mathbf{x} \mid x_i \geq 0,\ \mathbf{A} \in \mathbb{R}^{m \times n},\ \mathbf{x} \in \mathbb{R}^n\}$
3. (0.2 pts) $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{x}_0 + \alpha \mathbf{s},\ \alpha \in \mathbb{R},\ \mathbf{s} \in \mathbb{R}^n\}$
4. (0.5 pts) $\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_1 \leq 1\}$
5. (1 pts) Найти $\mathcal{G}^*, \mathcal{G}^{**}, \mathcal{G}^{***}$ для $\mathcal{G} = \{\mathbf{x} \in \mathbb{R}^2 \mid |x_1 - 1|^{1/5} + |x_2 - 1|^{1/5} \leq 1\}$

11. Марковские цепи (3.5 pts)

Пусть $\mathbf{P} = [p_{ij}] \in \mathbb{R}^{n \times n}$ — стохастическая матрица, то есть

$$\mathbf{P}^\top \mathbf{1} = \mathbf{1}, \qquad p_{ij} \geq 0,$$

где $\mathbf{1}$ — столбец из 1.

- (1 pts) Какое максимальное собственное значение матрицы $\mathbf{P}$ и почему?
- (1 pts) Докажите с использованием леммы Фаркаша, что существует решение следующей системы:

$$\mathbf{P}\mathbf{x} = \mathbf{x}, \qquad \mathbf{x}^\top \mathbf{1} = 1, \qquad x_i \geq 0.$$

Решение этой системы можно интерпретировать как равновесное распределение вероятностей Марковской цепи с n состояниями и матрицей переходов $\mathbf{P}$. Поэтому ответьте на следующие дополнительные вопросы:

1. (0.5 pts) Что такое Марковская цепь с конечным числом состояний?
2. (0.5 pts) Сформулируйте алгоритм PageRank для ранжирования веб-страниц.
3. (0.1 pts) Почему он называется PageRank?
4. (0.4 pts) Как алгоритм PageRank связан с марковскими цепями?

Подробно про Марковские цепи и их свойства Вам расскажут в следующем семестре в рамках курса «Случайные процессы».

11.1 Степенной метод* (2 pts)

Реализуйте алгоритм PageRank на Python[3] в Jupyter Notebook'е и приведите названия и авторов 10 самых важных статей по физике высоких энергий в соответствии с графом[4], а также названия журналов, в которых были опубликованы. Для эффективной реализации используйте разреженные матрицы. В дополнении к PDF-файлу, пришлите Jupyter Notebook файл с реализацией. Результат, приведённый в PDF, должен быть воспроизводим.

[3] Любая версия подойдёт.

[4] `https://snap.stanford.edu/data/cit-HepTh.html`

12. Альтернатива Фредгольма (1 pts)

С помощью леммы Фаркаша докажите альтернативу Фредгольма из линейной алгебры, а именно

Теорема 1 (Фредгольм) *Система линейных уравнений* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *имеет решение тогда и только тогда, когда не имеет решения следующая система:* $\mathbf{A}^\top \mathbf{y} = 0$, $\mathbf{b}^\top \mathbf{y} \neq 0$.

13. Векторное дифференцирование (5 pts)

Во всех нижеприведённых задачах найдите требуемые величины аналитически, после чего реализуйте функции для их вычисления и сравните скорость их работы с автоматическим дифференцированием в JAX. Jupyter Notebook со сравнением также приложите к решению. Рассмотрите различные размерности и сделайте вывод о том, какой подход асимптотически более быстрый.

1. (1 pts) Найдите градиент по $\mathbf{U} \in \mathbb{R}^{n\times n}$ и по $\mathbf{V} \in \mathbb{R}^{n\times n}$ функции $J(\mathbf{U}, \mathbf{V}) = \|\mathbf{U}\mathbf{V} - \mathbf{Y}\|_F^2 + \frac{\lambda}{2}(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2)$.

2. (1 pts) Найдите градиент и гессиан функции $f(\mathbf{w}) = \sum\limits_{i=1}^m \log(1 + e^{-y_i \mathbf{w}^\top \mathbf{x}_i})$, где $\mathbf{x}_i \in \mathbb{R}^n$, $y_i \in \mathbb{R}$.

3. (1 pts) Найдите матрицу Якоби для функции $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(\mathbf{w})_j = \dfrac{e^{w_j}}{\sum\limits_{k=1}^n e^{w_k}}$.

4. (0.5 pts) Найдите градиенты следующих функций от матрицы $\mathbf{X}$

   (a) $f(\mathbf{X}) = \sum\limits_{i=1}^n \lambda_i(\mathbf{X})$

   (b) $f(\mathbf{X}) = \prod\limits_{i=1}^n \lambda_i(\mathbf{X})$,

   где $\lambda_i(\mathbf{X})$ — собственные значения матрицы $\mathbf{X}$.

14. Дивергенция Кульбака-Лейблера (2 pts)

Пусть даны два вероятностных распределения $p(x)$ и $q(x)$. Тогда дивергенция Кульбака-Лейблера между ними $D_{KL}(p \parallel q)$ определяется как

$$D_{KL}(p \parallel q) = \int p(x) \log \frac{p(x)}{q(x)} dx.$$

Эта величина показывает схожесть между двумя распределениями и часто используется. Подробнее Вам расскажут в курсе «Математическая статистика» в 1 семестре 4 курса. Докажите следующие свойства дивергенции Кульбака-Лейблера:

1. (0.5 pts)$D_{KL}(p \parallel q)$ выпуклая функция на множестве пар распределений (p, q)

2. (0.5 pts) $D_{KL}(p \parallel q) = 0 \Leftrightarrow p \equiv q$ почти всюду.

3. (0.5 pts) $D_{KL}(p \parallel q) \geq 0$

4. (0.5 pts) $D_{KL}(p \parallel q) \neq D_{KL}(q \parallel p)$

15. Обратное неравенство Йенсена (1 pts)

Пусть $f : \mathbb{R}^n \to \mathbb{R}$ выпуклая функция, $\lambda_1 > 0$, $\lambda_i \leq 0$, $i = 2, \ldots, n$ и $\sum\limits_{i=1}^{n} \lambda_i = 1$ и пусть $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathrm{dom}\, f$. Докажите, что

$$f(\lambda_1 \mathbf{x}_1 + \ldots + \lambda_n \mathbf{x}_n) \geq \lambda_1 f(\mathbf{x}_1) + \ldots + \lambda_n f(\mathbf{x}_n).$$

16. Кратчайший путь в графе (1 pts)

Пусть задан ориентированный взвешенный граф $G = (V, E)$. Проверьте на выпуклость/вогнутость функцию $p_{ij}(\mathbf{c})$ кратчайшего расстояния между некоторой парой вершин (i, j), зависящую от вектора весов $\mathbf{c}$ рёбер графа.

17. Логарифмический барьер для конуса второго порядка (1 pts)

Функция $f(\mathbf{x}, t) = -\log(t^2 - \mathbf{x}^\top \mathbf{x})$, определённая на множестве $E = \{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{x}\|_2 < t\}$ называется логарифмическим барьером для конуса второго порядка. Она используется при переходе от условной оптимизации с ограничением на конус второго порядка к безусловной с помощью введения штрафа за выход из этого конуса. Поэтому такие функции называются барьерными. Докажите, что функция f выпукла.

18. Вычисление субградиентов (4 pts)

1. (1.5 pts) Найдите субдифференциал функции $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} + \mathbf{b}\|_1$.

2. (1 pts) Найдите субдифференциал функции $L(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|_2^2 + \sum\limits_{i=1}^{m} \max(0, 1 - y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b_i))$.

3. (0.5 pts) Найдите субдифференциал функции $f(x) = \begin{cases} 0 & x \in [-1, 1] \\ |x| - 1 & x \in [-2, -1) \cup (1, 2] \\ \infty & x \in (-\infty, -2) \cup (2, \infty) \end{cases}$

4. (1 pts) Найдите субдифференциал функции $f(\mathbf{x}) = |x_1 - x_2| + |x_1 + x_2|$ на множестве $X = \{\mathbf{x} | \|\mathbf{x}\|_2^2 \leq 2\}$.