Домашнее задание 4.

1. Условия оптимальности

1. Формулировка задачи. Положим $\widetilde{W} = W - \sum w_i$ и будем обозначать $E_i(s_i)$ — функции энергии из условия. Нетрудно заметить, что все эти функции выпуклые, т.к. их графиками являются левая и правая ветви парабол, склеенные в нуле. Получаем следующую задачу оптимизации:

$$\min \sum E_i(s_i)$$
$$\text{s.t.}$$
$$\sum s_i = \widetilde{W}$$

Условий типа неравенство нет, следовательно из условий ККТ остаются только два:

$$\sum s_i = \widetilde{W}$$
$$\operatorname{grad} \sum \big(E_i(s_i)\big) + \lambda \sum s_i - \lambda \widetilde{W} = 0$$

Условие на градиент распадается на n уравнений

$$\frac{dE_i(s_i)}{ds_i} = -\lambda. \tag{1}$$

При этом,

$$\frac{dE_i(s_i)}{ds_i} = \begin{cases} k_i^{\text{ext}}(s_i - N_i), & \text{if } s_i \geqslant N_i, \\ k_i^{\text{comp}}(s_i - N_i), & \text{if } s_i < N_i. \end{cases}$$

Разберем два случая:
1) $-\lambda \geqslant 0 \iff \lambda \leqslant 0$. Тогда уравнение 1 принимает вид:

$$k_i^{\text{ext}}(s_i - N_i) = -\lambda.$$

Получаем $s_i = -\frac{\lambda}{k_i^{\text{ext}}} + N_i$. Отсюда

$$\widetilde{W} = \sum s_i = -\lambda \sum \frac{1}{k_i^{\text{ext}}} + \sum N_i.$$

2) $-\lambda < 0 \iff \lambda > 0$. Тогда уравнение 1 принимает вид:

$$k_i^{\text{comp}}(s_i - N_i) = -\lambda.$$

Получаем $s_i = -\frac{\lambda}{k_i^{\text{comp}}} + N_i$. Отсюда

$$\widetilde{W} = \sum s_i = -\lambda \sum \frac{1}{k_i^{\text{comp}}} + \sum N_i.$$

Если $\widetilde{W} - \sum n_i \geqslant 0$ то попадаем в случай 1. Тогда

$$\lambda = \frac{\sum N_i - \widetilde{W}}{\sum \frac{1}{k_i^{\text{ext}}}}.$$

Отсюда

$$s_i = \frac{\widetilde{W} - \sum N_j}{k_i^{\text{ext}} \sum \frac{1}{k_j^{\text{ext}}}} + N_i.$$

Иначе, попали в случай 2. Аналогично

$$s_i = \frac{\widetilde{W} - \sum N_j}{k_i^{\mathrm{comp}} \sum \frac{1}{k_j^{\mathrm{comp}}}} + N_i.$$

3. Запишем условия ККТ:

(1) $\sum_{i=1}^{n} x_i = 1$,

(2) $-x_i \leqslant 0$,

(3) $\mu_i \geqslant 0$,

(4) $\mu_i x_i = 0$,

(5) $\mathrm{grad}_x \, \|x - y\|^2 + \lambda \sum x_i - \sum \mu_i x_i = 0$.

Последнее уравнение является системой из n уравнений:

$$2x_i - 2y_i + \lambda - \mu_i = 0$$

Получаем, что

$$x_i = y_i - \frac{\lambda}{2} + \frac{\mu_i}{2}.$$

Рассмотрим два случая:

(1) $y_i - \frac{\lambda}{2} \geqslant 0$. Тогда из условий 2, 3, 4, $x_i = y_i - \frac{\lambda}{2}$, $\mu_i = 0$.

(2) $y_i - \frac{\lambda}{2} < 0$. Тогда из условий 2, 3, 4, $x_i = 0$, $\mu_i = \lambda - 2y_i$.

Итак, получили, что

$$x_i = \max\left\{0, y_i - \frac{\lambda}{2}\right\} \tag{2}$$

Положим $\left(y_{(1)}, \ldots, y_{(n)}\right)$ — отсортированный по возрастанию вектор y. Положим

$$S(\lambda) = \sum_i x_i = \sum_i \max\left\{0, y_i - \frac{\lambda}{2}\right\}$$

Пусть $\frac{\lambda_0}{2} \in \left[y_{(k-1)}, y_{(k)}\right)$. Тогда

$$S(\lambda_0) = \sum_{i=k}^{n} \left(y_{(i)} - \frac{\lambda_0}{2}\right) = \sum_{i=k} y_{(i)} - (n - k + 1)\frac{\lambda_0}{2} \tag{3}$$

Видно, что каждое слагаемое > 0, с ростом λ каждое слагаемое становится меньше, и самих слагаемых становится меньше. Поэтому функция $S(\lambda)$ строго убывает.
Алгоритм решения.

(1) Сортируем y, получаем $(y_{(1)}, \ldots, y_{(n)})$ — $O(n \log n)$.

(2) Считаем $S(2y_{(i)})$ до первого i такого, что $S(2y_{(i-1)}) \leqslant 1 < S(2y_{(i)})$ — $O(n)$.

(3) Вычисляем λ из уравнения 3 — $O(n)$.

(4) Вычисляем x_i из уравнений 2 — $O(n)$.

Итого, $O(n \log n)$.

2. Двойственные задачи

2. Перепишем задачу в эквивалентном виде:

$$\min t,$$
$$\text{s.t.}$$
$$\max_i (a_i^T x + b_i) \leqslant t.$$

Условие $\max_i (a_i^T x + b_i) \leqslant t$ эквивалентно системе неравенств $a_i^T x + b_i \leqslant t$.
Итак, получили задачу:

$$\min t,$$
$$\text{s.t.}$$
$$a_i^T x + b_i \leqslant t.$$

Двойственная задача

$$L(x, t, \mu) = t + \sum_i \mu_i (a_i^T x + b_i - t)$$

Рассмотрим $\operatorname{grad}_x L(x, t, \mu)$. Он равен $\sum_i \mu_i a_i^T$. Далее, $\frac{\delta L(x,t,\mu)}{\delta t} = 1 - \sum_i \mu_i$. Положим A — матрица, столбцами которой являются вектор a_i.Получаем условия на dom$g$:

$$A\mu = 0,$$
$$\sum_i \mu_i = 1.$$

Действительно, если хотя бы одно из этих условий не выполняется, можно подобрать последовательность x_k, t_k, на которой достигается $\inf = -\infty$. При этих условиях

$$L(x, t, \mu) = \sum_i \mu_i b_i = \inf_{x,t} L(x, t, \mu) = L(\mu).$$

Двойственная задача:

$$\max_\mu \sum_i \mu_i b_i,$$
$$\text{s.t.}$$
$$A\mu = 0,$$
$$\sum_i \mu_i = 1.$$

3. Положим $a_i = (a_i^1, \ldots, a_i^n)$, $A^i = a_i a_i^T$. Заметим, что $a_i^T X a_i = \sum\limits_{j=1}^n a_i^j \sum\limits_{k=0}^n X_{jk} a_i^k = \sum\limits_{jk} A^i_{jk} X_{jk}$. Применяя тождество: $(AX)_{ij} = \sum\limits_k X_{kj} A_{ik}$, получаем, что $\sum\limits_j \sum\limits_k A^i_{jk} X_{jk} = \sum\limits_j \left(A^i X^T\right)_{jj} = \operatorname{tr}(A^i X^T) = \operatorname{tr}(A^i X)$
Поэтому, неравенства из условия переписываются в виде

$$\operatorname{tr}(A^i X) - 1 \leqslant 0.$$

Двойственная задача: Запишем лагранжиан.

$$L(X, \mu) = -\log \det X + \sum_i \mu_i \operatorname{tr}(A^i X) - \sum_i \mu_i.$$

Из семинаров известно, что $\nabla \log \det X = X^{-1}$, а из явной формулы следа произведения матриц сразу же следует, что $\nabla \operatorname{tr}\left(A^i X\right) = A^i$. Получаем, что

$$\nabla_X L(X, \mu) = -X^{-1} + \sum_i \mu_i A^i.$$

Приравнивая к 0 получаем, что минимум достигается в $X = \left(\sum\limits_i \mu_i A^i\right)^{-1}$. По условию X — симметричная положительно определенная, поэтому условие на $\operatorname{dom} g$ выглядит как $\left(\sum\limits_i \mu_i A^i\right)^{-1} \in S_{++}$. Подставим полученный X в лагранжиан и, с учетом линейности следа матрицы, получим двойственную функцию:

$$g(\mu) = -\log\det\left(\left(\sum_i \mu_i A^i\right)^{-1}\right) + \operatorname{tr}\left(\left(\sum_i \mu_i A^i\right), \left(\sum_i \mu_i A^i\right)^{-1}\right) - \sum_i \mu_i,$$

Итоговая двойственная задача:

$$\min_\mu \log\det\left(\sum_i \mu_i A^i\right) - \sum_i \mu_i + n.$$

$$\text{s.t.}$$

$$\sum_i \mu_i A^i \in S^n_{++}.$$