Методы оптимизации. Семинар 11. От LP до SDP, примеры задач.

Корнилов Никита Максимович

Московский физико-технический институт

12 ноября 2024г

Линейное программирование

Общий вид

$$\min_{x\in\mathbb{R}^n} \langle c, x\rangle$$
$$\textbf{s.t. } Ax = b,$$
$$Gx \leq h.$$

Стандартный вид

$$\min_{x\in\mathbb{R}^n} \langle c, x\rangle$$
$$\textbf{s.t. } Ax = b,$$
$$x \geq 0.$$

Задачу в общем виде можно свести к стандартному

Дробно-линейное программирование

Example

$$\min_{x \in \mathbb{R}^n} \left[f(x) := \frac{c^\top x + d}{e^\top x + f} \right]$$

$$\text{s.t. } Ax = b,$$

$$Gx \leq h,$$

где $\{x | e^\top x + f > 0\}$.

Задача невыпуклая, однако можно свести к LP.

Замена:

$$y = \frac{x}{e^\top x + f} \in \mathbb{R}^n, \quad z = \frac{1}{e^\top x + f} \in \mathbb{R}.$$

Эквивалентная система (почему?):

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \; & c^\top y + dz \\ \textbf{s.t.} \; & Ay - bz = 0, \\ & Gy - hz \leq 0, \\ & e^\top y + fz = 1, \\ & z \geq 0. \end{aligned}$$

Сведение через двойственность

Example

$$\min_{x \in \mathbb{R}^n} f(x) := \sum_{i=1}^{r} x_{[i]}$$
$$\textbf{s.t. } Ax = b,$$
$$Gx \le h,$$

где $x_{[i]}$ - i-ая по величине координата, то есть $x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$.

Сведение через двойственность

Замена

$$
\begin{aligned}
f(x) = \max_{y \in \mathbb{R}^n} \ & y^\top x \\
\textbf{s.t.} \ & 0 \le y \le 1, \\
& 1^\top y = r.
\end{aligned}
$$

Сильная двойственность и двойственная задача (накладываем только условия $y \le 1$):

$$
\begin{aligned}
\min_{t,\nu} \ & rt + \mathbf{1}^\top \nu \\
\textbf{s.t.} \ & t\mathbf{1} + \nu \ge x, \\
& \nu \ge 0.
\end{aligned}
$$

Quadratic Constrained Quadratic Programming (QCQP)

$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \ & \frac{1}{2} x^\top A_0 x + b_0^\top x + c_0 \\
\textbf{s.t.} \ & \frac{1}{2} x^\top A_i x + b_i^\top x + c_i \leq 0, \quad i = \overline{1, m},
\end{aligned}
$$

$A_i \in \mathbb{S}^n$ – необязательно неотрицательно определенные.
В частности, неравенства

- $x_i \in \{-1, 1\} \quad x_i^2 - 1 = 0,$
- $x_i \in \{0, 1\} \quad x_i^2 - x_i = 0,$
- $x_i \in \{-M, -M+1, \ldots, N-1, N\}, \quad x_i = \sum_{j=1}^N u_j^+ - \sum_{j=1}^M u_j^-.$

Если все ограничения вида неравенства линейные, то задача называется просто Quadratic Programming.

LP лежит в QCQP!

Метод наименьших квадратов

Example

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2}\|Ax - b\|_2^2.$$

Составление портфеля

Предположим имеется

- n активов, куда можно вложить капитал
- $x \in \mathbb{R}^n, x \geq 0, \mathbf{1}^\top x = 1$ – распределение по активам
- ξ_i –случайная величина прибыли i-го актива со средним p_i и дисперсей σ_i
- Ковариация между активами задана матрицей Σ_{ij}
- Хотим среднюю прибыль не меньше α и с наименьшей дисперсией

Составление портфеля

Предположим имеется

- n активов, куда можно вложить капитал
- $x \in \mathbb{R}^n, x \geq 0, \mathbf{1}^\top x = 1$ – распределение по активам
- ξ_i –случайная величина прибыли i-го актива со средним p_i и дисперсей σ_i
- Ковариация между активами задана матрицей Σ_{ij}
- Хотим среднюю прибыль не меньше α и с наименьшей дисперсией

$$\min_{x \in \mathbb{R}^n} x^\top \Sigma x$$
$$\text{s.t. } x \geq 0, \mathbf{1}^\top x = 1$$
$$\sum_{i=1}^n x_i p_i \geq \alpha.$$

Приближение сферой

Пусть $\{x_i\}_{i=1}^k$ – зашумленные нормальным шумом $\mathcal{N}(0, I_n\varepsilon^2)$ точки с евклидовой сферы $S_r^n(x_c) := \{x \in \mathbb{R}^n, \quad \|x - x_c\|_2 = r\}$. То есть

$$x_i = \hat{x}_i + u_i, \quad \hat{x}_i \in S_r^n(x_c), \quad u_i \sim \mathcal{N}(0, I_n\varepsilon^2).$$

Приближение сферой

Пусть $\{x_i\}_{i=1}^k$ – зашумленные нормальным шумом $\mathcal{N}(0, I_n\varepsilon^2)$ точки с евклидовой сферы $S_r^n(x_c) := \{x \in \mathbb{R}^n, \quad \|x - x_c\|_2 = r\}$. То есть

$$x_i = \hat{x}_i + u_i, \quad \hat{x}_i \in S_r^n(x_c), \quad u_i \sim \mathcal{N}(0, I_n\varepsilon^2).$$

Задача оптимизации

$$\min_{x_c \in \mathbb{R}^n, r \in \mathbb{R}_+} \sum_{i=1}^k (\|x_i - x_c\|_2^2 - r^2)^2.$$

В общем случае, функция $(x^2 - y^2)^2$ невыпуклая.

Эквивалентная задача

Перепишем задачу:

$$f(x) = \sum_{i=1}^{k} (\|x_i - x_c\|_2^2 - r^2)^2 = \sum_{i=1}^{k} (\|x_i\|_2^2 - 2x_i^\top x_c + \|x_c\|_2^2 - r^2)^2.$$

Замена:

$$(x_c, r) \to (x_c, t = \|x_c\|_2^2 - r^2)$$

Получившаяся задача

$$\min_{x_c \in \mathbb{R}^n, t \in \mathbb{R}} \sum_{i=1}^{k} (\|x_i\|_2^2 - 2x_i^\top x_c + t)^2.$$

Множество решений шире, но в точке оптимума $\|x_c\|^2 \geq t$ (почему?).

Second-order Conic Programming

$$
\begin{aligned}
\min_{x \in \mathbb{R}^n} \ & \langle c, x \rangle \\
\textbf{s.t. } & Ax = b \\
& \|G_i x - h_i\|_2 \leq e_i^\top x + f_i, \quad i = \overline{1, M}.
\end{aligned}
$$

Последнее условие означает, что пара $(G_i x - h_i, e_i^\top x + f_i)$ лежит в конусе $K_2 = \{(y, t) | \quad \|y\|_2 \leq t, t \geq 0\}$.

Выпуклые задачи QCQP лежат в SOCP!

Сведение через эпиграф

Example

$$\min_{x\in\mathbb{R}^n}\sum_{i=1}^{m}\|A_ix-b_i\|_2,$$

$$\min_{x\in\mathbb{R}^n}\max_{i=\overline{1,m}}\|A_ix-b_i\|_2.$$

Example

$$\min_{x\in\mathbb{R}^n} f(x) := \sup_{(A,b)\in\mathcal{A}} \|Ax-b\|_2,$$

где $\mathcal{A} = \{(A,b)| \quad \|(A-A_0, b-b_0)\|_F \leq \rho\}$.

Оценим сверху

$$f(x) \leq \|A_0x-b_0\|_2 + \rho\left\|\begin{pmatrix} x \\ 1 \end{pmatrix}\right\|_2.$$

Robust Linear Regression

- При $A_0x + b_0 = 0$ берём $(\Delta A, \Delta b) = \frac{\rho}{\sqrt{n}\|(x,1)\|_2}\mathbf{1}_m \cdot (x^T, 1)$.
- При $A_0x + b_0 \neq 0$ берём $(\Delta A, \Delta b) = uv^\top$, где $u = \rho\frac{A_0x - b_0}{\|A_0x - b_0\|_2}, v = \frac{(x,1)}{\|(x,1)\|_2}$.

Итого, эквивалентная задача

$$\min_{x \in \mathbb{R}^n} \|A_0x - b_0\|_2 + \rho \left\| \begin{pmatrix} x \\ 1 \end{pmatrix} \right\|_2.$$

Стандартная форма

$$
\begin{aligned}
\min_{X \in \mathbb{S}^n} \ & \langle C, X \rangle \\
\textbf{s.t.} \ & \langle A_i, X \rangle = b_i, \quad i = \overline{1, m}, \\
& X \succeq 0.
\end{aligned}
$$

Двойственная задача

$$
\begin{aligned}
\max_{x \in \mathbb{R}^m} \ & \langle b, x \rangle \\
\textbf{s.t.} \ & \sum_{i=1}^{m} A_i x_i \preceq C.
\end{aligned}
$$

SOCP лежит в SDP!

Минимум спектральной нормы

Example

Пусть $A(x) := A_0 + \sum_{i=1}^{n} A_i x_i, \quad A_i \in \mathbb{R}^{m\times n}$ и требуется найти

$$\min_{x\in\mathbb{R}^n} \|A(x)\|_2.$$

Theorem (Дополнение по Шуру)

Для блочной матрицы $M \in \mathbb{S}^{n+m}$ с обратимой $A \in \mathbb{S}^n$ верно

$$M = \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix} \succeq 0 \quad \Longleftrightarrow \quad \begin{cases} A \succ 0, \\ C - B^\top A^{-1} B \succeq 0. \end{cases}$$

Минимум спектральной нормы

$$\min_{x\in\mathbb{R}^m; t\in\mathbb{R}} \; t$$
$$\textbf{s.t. } \|A(x)\|_2^2 \le t^2.$$

Эквивалентная задача

$$\lambda_{\max}(A(x)^\top A(x)) \le t^2, t \ge 0 \iff t^2 I_n \succeq A(x)^\top A(x).$$

Дополнение по Шуру

$$\begin{pmatrix} tI_n & A(x)^\top \\ A(x) & tI_m \end{pmatrix} \succeq 0.$$

Релаксация QCQP

Замена:
Введём $X = xx^\top$ - новые переменные, забирающие всю сложность.
Релаксация к SDP:

$$X \succeq xx^\top \quad \iff \quad \begin{pmatrix} X & x \\ x^\top & 1 \end{pmatrix} \succeq 0,$$

$$\frac{1}{2}x^\top A_i x + b_i^\top x + c_i = \frac{1}{2}\langle A_i, X\rangle + b_i^\top x + c_i.$$

Если у оптимума $\mathrm{rank}(X) = 1$, то это решение исходной задачи. Если эта задача недопустима, то и исходная тоже.

$$\min_{x \in \mathcal{K}} \langle c, x \rangle$$
$$\textbf{s.t. } Ax = b,$$

где $\mathcal{K}$ - конус в гильбертовом пространстве.

- Экспоненциальный конус:

$$K_{\text{exp}} = \text{cl}\left\{(x, y, z) \in \mathbb{R}^3 \middle| x \geq y \exp\left(\frac{z}{y}\right), y > 0\right\}.$$

  Ограничения: $\log(1 + e^x) \leq t$ и $e^x < t$.
- Степенной конус $0 < \alpha < 1$:

$$\mathcal{P}_n^{\alpha, 1-\alpha} = \left\{x \in \mathbb{R}^n \middle| x_1^\alpha x_2^{1-\alpha} \geq \sqrt{\sum_{i=3}^n x_i^2}, x_1 \geq 0, x_2 \geq 0\right\}.$$

  Ограничения: $\|x\|_p \leq t$.