

1 Стандартные классы выпуклых задач

1.1 Линейное программирование (LP)

Стандартная форма LP:

$$\min_x c^T x \quad \text{s. t.} \quad Ax \preceq b,$$

где $c, x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$.

Пример 1 (Максимум из аффинных функций). Пусть $a_1, \dots, a_m \in \mathbb{R}^n$ и $b_1, \dots, b_m \in \mathbb{R}$. Рассмотрим задачу

$$\min_{x \in \mathbb{R}^n} \max_{1 \leq i \leq m} \{a_i^T x - b_i\}.$$

Эта задача является задачей *негладкой* безусловной минимизации (из-за присутствия максимума). Тем не менее, эта задача эквивалентна следующей задаче линейного программирования:

$$\min_{x, t} t \quad \text{s. t.} \quad t \geq a_i^T x - b_i, \quad 1 \leq i \leq m.$$

В стандартной форме это соответствует тому, что $\tilde{n} = n + 1$, $\tilde{m} = m$, $\tilde{x} = (x, t) \in \mathbb{R}^{n+1}$, $\tilde{c} := (0_n, 1) \in \mathbb{R}^{n+1}$, $\tilde{b} := b$ и $\tilde{A} := (A, 1_n) \in \mathbb{R}^{m \times (n+1)}$, где $A \in \mathbb{R}^{m \times n}$ — матрица со строками a_1, \dots, a_m .

1.2 Квадратичное программирование (QP)

Стандартная форма QP:

$$\min_x \left\{ \frac{1}{2} x^T A x - b^T x \right\} \quad \text{s. t.} \quad Gx \preceq h,$$

где $x, b \in \mathbb{R}^n$, $A \in \mathbb{S}_+^n$, $G \in \mathbb{R}^{m \times n}$, $h \in \mathbb{R}^m$.

Пример: линейная регрессия $\|Ax - b\|_2^2$ на положительном ортанте: $x \succeq 0$.

Пример: расстояние между полиэдрами: $\min \|x - y\|_2^2$, $Ax \preceq b$, $Cy \preceq d$.

1.3 Квадратичное программирование с квадратичными ограничениями (QCQP)

Стандартная форма QCQP:

$$\min_x \left\{ \frac{1}{2} x^T A x - b^T x \right\} \quad \text{s. t.} \quad \frac{1}{2} x^T P_i x + q_i^T x + c_i \leq 0, \quad 1 \leq i \leq m,$$

где $x, b, q_1, \dots, q_m \in \mathbb{R}^n$, $A, P_1, \dots, P_m \in \mathbb{S}_+^n$, $c_1, \dots, c_m \in \mathbb{R}$.

Пример 2 (Квадратичная функция на единичном шаре). Рассмотрим задачу

$$\min_x \left\{ \frac{1}{2} x^T A x - b^T x \right\} \quad \text{s. t.} \quad \|x\|_2 \leq 1,$$

где $x, b \in \mathbb{R}^n$ и $A \in \mathbb{S}_+^n$. Эта задача эквивалентна следующей задаче QCQP:

$$\min_x \left\{ \frac{1}{2} x^T A x - b^T x \right\} \quad \text{s. t.} \quad x^T x \leq 1.$$

В этом случае $m = 1$, $P_1 = I_n$, $q_1 = 0$ и $c_1 = -1$.

1.4 Коническое программирование второго порядка (SOCP)

Стандартная форма SOCP:

$$\min_x q^T x \quad \text{s. t.} \quad \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad 1 \leq i \leq m,$$

где $x, q, c_1, \dots, c_m \in \mathbb{R}^n$, $A_1 \in \mathbb{R}^{s_1 \times n}$, $b_1 \in \mathbb{R}^{s_1}$, \dots , $A_m \in \mathbb{R}^{s_m \times n}$, $b_m \in \mathbb{R}^{s_m}$, $d_1, \dots, d_m \in \mathbb{R}$.

Пример 3 (Задача Lasso). Пусть $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Рассмотрим задачу

$$\min_x \{ \|Ax - b\|_2^2 + \|x\|_2 \}. \quad (1)$$

Эта задача является негладкой из-за присутствия слагаемого $\|x\|_2$. Эта задача является эквивалентной следующей задаче SOCP:

$$\min_{x, t_1, t_2} \{ t_1 + t_2 \} \quad \text{s. t.} \quad \|Ax - b\|_2 \leq t_1, \quad \|x\|_2 \leq t_2.$$

(Почему эту задачу (1) нельзя было похожим образом переформулировать как QCQP?)

1.5 Полуопределённое программирование (SDP)

Стандартная форма SDP:

$$\min_x c^T x \quad \text{s. t.} \quad A_0 + x_1 A_1 + \dots + x_n A_n \succeq 0, \quad Cx = d,$$

где $x, c \in \mathbb{R}^n$, $A_0, \dots, A_n \in \mathbb{S}^n$, $C \in \mathbb{R}^{m \times n}$, $d \in \mathbb{R}^m$.

Пример 4 (Минимизация спектральной нормы). Рассмотрим задачу

$$\min_x \|A(x) := A_0 + x_1 A_1 + \dots + x_n A_n\|_2,$$

где $x \in \mathbb{R}^n$, $A_0, \dots, A_n \in \mathbb{S}^n$. Представим эту задачу в виде SDP. Для этого сначала перейдем эквивалентной формулировке через надграфик:

$$\min_{x, t} t \quad \text{s. t.} \quad t \geq \|A(x)\|_2.$$

Заметим, что

$$\|A(x)\|_2 \leq t \quad \Leftrightarrow \quad \lambda_{\max}(A(x)A(x)^T) \leq t^2 \quad \Leftrightarrow \quad A(x)A(x)^T \leq t^2 I_n.$$

По лемме о дополнении Шура, получаем

$$A(x)A(x)^T \leq t^2 I_n \quad \Leftrightarrow \quad \begin{pmatrix} tI_n & A(x)^T \\ A(x) & tI_n \end{pmatrix} \succeq 0.$$

Таким образом, исходная задача эквивалентна следующей SDP:

$$\min_{x, t} t \quad \text{s. t.} \quad \begin{pmatrix} tI_n & A(x)^T \\ A(x) & tI_n \end{pmatrix} \succeq 0.$$

1.6 Обсуждение

Можно показать, что $LP \subset QR \subset QCQP \subset SOCP \subset SDP$. Таким образом, SDP являются наиболее общей постановкой выпуклых задач.

Замечание 1. Полезно заметить, что целевую функцию всегда можно перенести в ограничения и получить эквивалентную задачу с линейной целевой функцией.

2 Двойственные задачи

Ещё одно важное понятие, возникающее в условной оптимизации — двойственная задача.

Мы предложим универсальный способ: как для любой условной задачи оптимизации $\min_{x \in Q} f(x)$ записать новую *двойственную* задачу $\max_{y \in \Omega} g(y)$, при этом, двойственная задача будет оценивать прямую задачу снизу:

$$g(y) \leq f(x) \quad \text{для всех } y \in \Omega, x \in Q.$$

Поскольку неравенство справедливо для всех допустимых x и y , значит:

$$\max_{y \in \Omega} g(y) \leq \min_{x \in Q} f(x).$$

Прежде чем перейти к конкретному (одному из возможных) способов построения двойственной задачи, обсудим, зачем это может быть полезно.

1. *Построение оценки снизу на решение прямой задачи.* Решить исходную задачу может быть очень сложно. Но если у нас есть двойственная к ней, то мы можем взять произвольный $y \in \Omega$ и подставить его в $g(y)$ — получим некоторую оценку снизу:

$$g(y) \leq \min_{x \in Q} f(x).$$

2. *Проверка на допустимость задачи и ограниченность решения.* Из неравенства $\max_{y \in \Omega} g(y) \leq \min_{x \in Q} f(x)$ следует: если $\min_{x \in Q} f(x) = -\infty$, значит $\Omega = \emptyset$ и наоборот.
3. *Двойственную задачу бывает решить легче, чем прямую.* При этом, если выполнена *сильная двойственность*: $g(y^*) = f(x^*)$ то мы ничего не теряем.
4. *Получение оценки сверху на невязку по функции:* $f(x) - f^* \leq f(x) - g(y)$ для произвольного $y \in \Omega$.

Опишем возможный способ построения двойственной задачи для случая, когда Q задаётся функциональными ограничениями.

Заметим, что в этом случае, исходная условная задача $\min_{x \in Q} f(x)$ эквивалента следующей безусловной задаче:

$$\min_{x \in Q} \max_{\lambda \in \mathbb{R}_+^m, \mu \in \mathbb{R}^k} \mathcal{L}(x; \lambda, \mu).$$

Поменяем минимум и максимум местами — получим оценку снизу (это универсальное правило, которое всегда можно использовать):

$$\max_{\lambda \in \mathbb{R}_+^m, \mu \in \mathbb{R}^k} \min_{x \in Q} \mathcal{L}(x; \lambda, \mu) \leq \min_{x \in Q} \max_{\lambda \in \mathbb{R}_+^m, \mu \in \mathbb{R}^k} \mathcal{L}(x; \lambda, \mu).$$

Обозначим внутреннюю задачу оптимизации в оценке снизу за $g(\lambda, \mu)$:

$$g(\lambda, \mu) := \min_{x \in Q} \mathcal{L}(x; \lambda, \mu).$$

Получили *двойственную задачу*:

$$\max_{\lambda \in \mathbb{R}_+^m, \mu \in \mathbb{R}^k} g(\lambda, \mu).$$

Замечание 2. Заметим, что двойственная функция $g(\lambda, \mu)$ всегда будет вогнутой, как максимум аффинных функций.

По построению всегда выполнена *слабая двойственность*: $g^* \leq f^*$.

В случае, если для исходной выпуклой задачи выполнено *условие Слейтера*, справедлива *сильная двойственность*: $g^* = f^*$.

Замечание 3. Если выполнена сильная двойственность ($g^* = f^*$), то решения прямой задачи x^* и двойственной (λ^*, μ^*) тоже связаны:

$$x^* \in \underset{x \in \mathbb{R}^n}{\text{Argmin}} \mathcal{L}(x, \lambda^*, \mu^*),$$

— решения прямой задачи принадлежат множеству минимумов лагранжиана по прямым переменным, при фиксированных оптимальных двойственных.