Методы оптимизации
Лекция 12: Линейное программирование

Александр Катруца

Физтех-школа прикладной математики и информатики
Московский физико-технический институт

30 ноября 2020 г.

На прошлой лекции

- Метод проекции градиента

На прошлой лекции

- Метод проекции градиента
- Проксимальный метод

На прошлой лекции

- Метод проекции градиента
- Проксимальный метод
- Проксимальный градиентный метод

Постановка задачи: напоминание

- Дано: $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m\times n}$ и $\mathbf{b} \in \mathbb{R}^m$
- Стандартная форма

$$
\begin{aligned}
&\min_{\mathbf{x}} \langle \mathbf{c}, \mathbf{x} \rangle \\
\text{s.t. } &\mathbf{A}\mathbf{x} = \mathbf{b} \\
&\mathbf{x} \geq 0
\end{aligned}
$$

- Преобразование задач
  - $\mathbf{A}\mathbf{x} \leq \mathbf{b} \to \mathbf{A}\mathbf{x} + \mathbf{y} = \mathbf{b}, \; \mathbf{y} \geq 0$
  - Свободная переменная $\mathbf{x} \to \mathbf{x} = \mathbf{y} - \mathbf{z}, \; \mathbf{y} \geq 0, \; \mathbf{z} \geq 0$
  - Замена знака достигается за счёт умножения на -1
- Минимизация максимума линейных функций сводится к линейному программирования

Ключевые элементы допустимого множества

Определение

Множество P вида $P = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$, где $\mathbf{A} \in \mathbb{R}^{m \times n}$ и $m < n$, называется многогранником (polyhedron).

Определение

Точка $\mathbf{y} \in P$ называется крайней точкой многоугольника, если не существует двух других точек из P, между которыми она лежит.

Определение

Точка $\mathbf{z} \in P$ называется вершиной многоугольника, если найдётся такой вектор $\mathbf{c} \in \mathbb{R}^n$, что $\langle \mathbf{c}, \mathbf{z} \rangle < \langle \mathbf{c}, \mathbf{x} \rangle$ для всех других точек $\mathbf{x} \in P$ и $\mathbf{x} \neq \mathbf{z}$.

От геометрии к алгебре

Пусть многогранник задан в виде
$\{\mathbf{x} \mid \langle \mathbf{a}_i, \mathbf{x} \rangle = b_i \ \langle \mathbf{a}_j, \mathbf{x} \rangle \leq b_j, \ \langle \mathbf{a}_k, \mathbf{x} \rangle \geq b_k\}$.

Теорема

Пусть $\mathbf{x} \in \mathbb{R}^n$ и $I = \{i \mid \langle \mathbf{a}_i, \mathbf{x} \rangle = b_i\}$ — индексы активных ограничений. Тогда следующие утверждения эквивалентны

1. Найдётся n линейно независимых векторов в множестве $\{\mathbf{a}_i \mid i \in I\}$
2. Они образуют базис в $\mathbb{R}^n$
3. Система $\langle \mathbf{a}_i, \mathbf{x} \rangle = b_i, \ i \in I$ имеет единственное решение

Доказательство

- $1 \Leftrightarrow 2$ — очевидно из линейной алгебры
- $2 \Rightarrow 3$: если два решения $\mathbf{x}_1$ и $\mathbf{x}_2$, то $\mathbf{d} = \mathbf{x}_1 - \mathbf{x}_2$ ортогонален всем $\mathbf{a}_i$, противоречие
- $3 \Rightarrow 2$: если не базис, то возьмём $\mathbf{d}$ ортогональный подпространству для $\mathbf{a}_i$, тогда из любого решения $\mathbf{x}^*$ получим другое решение $\mathbf{x}^* + \mathbf{d}$

Ещё одно определение

Базисное решение

Пусть P задан ограничениями равенствами и неравенствами и $\mathbf{x} \in \mathbb{R}^n$. Тогда

- $\mathbf{x}$ базисное решение, если все ограничения равенства активны и среди всех активных ограничений n линейно независимых
- $\mathbf{x}$ базисное **допустимое** решение, если оно базисное и удовлетворяет всем ограничениям

Эквивалентность определений

Теорема

Пусть P многогранник и пусть $\mathbf{x} \in P$. Тогда следующие факты об $\mathbf{x}$ эквивалентны

1. $\mathbf{x}$ — вершина
2. $\mathbf{x}$ — крайняя точка
3. $\mathbf{x}$ — базовое допустимое решение

Многогранник в стандартной форме

Уточним результаты для $P = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0\}$, где строки матрицы $\mathbf{A}$ линейно независимы.

Теорема

Вектор $\mathbf{x}$ базисное решение тогда и только тогда, когда $\mathbf{A}\mathbf{x} = \mathbf{b}$ и найдутся индексы $B(1), \ldots, B(m)$ такие что

- столбцы $\mathbf{A}_{B(1)}, \ldots, \mathbf{A}_{B(m)}$ линейно независимы
- Если $i \notin \{B(1), \ldots, B(m)\}$ то $x_i = 0$.

Вырожденное базовое решение

Определение

Пусть $P = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0\}$ и $\mathbf{x}$ базовое решение. Тогда оно вырождено, если больше $n - m$ его элементов нули.

Существование крайней точки

Теорема

Пусть многоугольник задан в виде $\{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}\}$. Тогда следующие утверждения эквивалентны

- у P есть хотя бы одна крайняя точка
- P не содержит прямой
- найдётся n линейно независимых векторов $\mathbf{a}_1, \ldots, \mathbf{a}_m$.

Следствие

Любой ограниченный многоугольник и любой многоугольник в стандартной форме имеют крайнюю точку.

Оптимальность крайней точки

Теорема

Если многоугольник имеет хотя бы одну крайнюю точку, а задача линейного программирования имеет решение, тогда это решение в крайней точки.

Доказательство

- Множество решений $Q = \{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0, \langle \mathbf{c}, \mathbf{x} \rangle = c^*\}$ — многогранник
- $Q \subset P$
- P имеет крайнюю точку, значит и в Q она есть
- Пусть $\mathbf{x}^*$ — крайняя точка в Q, тогда она крайняя для P
- Если это не так, то найдутся точки $\mathbf{y}, \mathbf{z} \in P$ такие что $\mathbf{x}^* = \alpha\mathbf{y} + (1-\alpha)\mathbf{z}$, $\alpha \in [0, 1]$
- $c^* = \langle \mathbf{c}, \mathbf{x}^* \rangle = \alpha\langle \mathbf{c}, \mathbf{y} \rangle + (1-\alpha)\langle \mathbf{c}, \mathbf{z} \rangle$, также $\langle \mathbf{c}, \mathbf{y} \rangle \geq c^*$ и $\langle \mathbf{c}, \mathbf{z} \rangle \geq c^*$
- $\langle \mathbf{c}, \mathbf{z} \rangle = \langle \mathbf{c}, \mathbf{y} \rangle = c^*$ и $\mathbf{z}, \mathbf{y} \in Q$
- Противоречие с тем, что $\mathbf{x}^*$ крайняя точка в Q

История исследования задачи

- Разработка методики применения линейного программирования в экономике (Л.В. Канторович, 1930-ые гг.)
- Симплекс-метод (Дж. Данциг, 1949 г.)
- Доказана полиномиальность задачи линейного программирования (Л. Хачиян, 1979)
- Первый практически полезный полиномиальный алгоритм (Н. Кармаркар, 1984)

Симплекс-метод: идея

- Найти некоторую базовую допустимую точку
- Перейти в **сопряжённую** угловую точку так, чтобы целевая функция уменьшилась
- Проверить, есть ли сопряжённые точки, в которых значение целевой функции меньше

Симплекс-метод: формулировка

Дана крайняя точка $\mathbf{x}$, матрица базиса $\mathbf{B}$ и множество индексов $\mathcal{B}$.

1. Вычислить оценки замещения $\overline{c}_j = c_j - c_{\mathcal{B}}^\top \mathbf{B}^{-1}\mathbf{A}_j$ для всех $j \notin \mathcal{B}$.
   - если $\overline{c}_j \geq 0$ для всех j, то текущее значение является оптимальным и уменьшить целевую функцию нельзя
   - иначе выбрать индекс j^*, для которого $\overline{c}_{j^*} < 0$
2. Вычислить $\mathbf{u} = \mathbf{B}^{-1}\mathbf{A}_{j^*}$
   - если все компоненты $\mathbf{u}$ неположительны, то задача неограничена, оптимальное значение равно $-\infty$
   - если есть положительные компоненты, то

$$\theta^* = \min_{\{i | u_i > 0\}} \frac{x_{\mathcal{B}(i)}}{u_i}$$

3. Пусть ℓ такой индекс, что $\theta^* = \frac{x_{\mathcal{B}(\ell)}}{u_\ell}$. Новая матрица базиса $\hat{\mathbf{B}}$ – замена столбца $\mathbf{A}_{\mathcal{B}(\ell)}$ на столбец $\mathbf{A}_{j^*}$. Новая крайняя точка $\hat{\mathbf{x}}$

$$\hat{x}_{j^*} = \theta^*$$

$$\hat{x}_{\mathcal{B}(k)} = x_{\mathcal{B}(k)} - \theta^* u_k, \text{если } k \neq \ell$$

Условие оптимальности

Теорема

Пусть $\mathbf{x}$ — базовое допустимое решение, которому соответствует матрица базиса $\mathbf{B}$ и оценки замещения $\bar{\mathbf{c}}$. Тогда если $\bar{\mathbf{c}} \geq 0$, то $\mathbf{x}$ решение.

Доказательство

1. Пусть $\mathbf{y}$ допустимое решение, тогда $\mathbf{d} = \mathbf{y} - \mathbf{x}$ и $\mathbf{Ad} = 0$
2. Можно переписать в виде $\mathbf{B}\mathbf{d}_\mathbf{B} + \sum\limits_{i \in N} \mathbf{A}_i d_i = 0$
3. Откуда следует $\mathbf{d}_\mathbf{B} = -\sum\limits_{i \in N} \mathbf{B}^{-1}\mathbf{A}_i d_i$
4. Вместе с тем $\langle \mathbf{c}, \mathbf{d} \rangle = \langle \mathbf{c}_\mathbf{B}, \mathbf{d}_\mathbf{B} \rangle + \sum\limits_{i \in N} c_i d_i =$
   $\sum\limits_{i \in N} (c_i - \langle \mathbf{c}_\mathbf{B}, \mathbf{B}^{-1}\mathbf{A}_i \rangle) d_i = \sum\limits_{i \in N} \bar{c}_i d_i$
5. Так как $i \in N$, то $x_i = 0$. Также $\mathbf{y}$ допустимая точка, значит $y_i \geq 0$ и $d_i = y_i - x_i \geq 0$
6. Но $\bar{c}_i \geq 0$ по предположению, а значит $\langle \mathbf{c}, \mathbf{d} \rangle \geq 0$

Двухфазный симплекс-метод

Q: как найти начальную базовую допустимую точку?

$$\begin{aligned} &\min y_1 + \ldots + y_m \\ \text{s.t. } &\mathbf{Az} + \mathbf{y} = \mathbf{b} \\ &\mathbf{z} \geq 0;\ \mathbf{y} \geq 0 \end{aligned}$$

- Начальная точка $\mathbf{z} = 0$, $\mathbf{y} = \mathbf{b}$
- Если оптимальное значение целевой функции равно нулю, то $\mathbf{z}^*$ — начальное базовое допустимое решение
- Иначе, допустимое множество пусто

Экспоненциальная сложность симплекс-метода

Пример Klee-Minty

$$
\begin{aligned}
& \max 2^{n-1}x_1 + \ldots + 2x_{n-1} + x_n \\
\text{s.t. } & x_1 \leq 5 \\
& 4x_1 + x_2 \leq 25 \\
& \ldots \\
& 2^n x_1 + 2^{n-1}x_2 + \ldots + x_n \leq 5^n \\
& \mathbf{x} \geq 0
\end{aligned}
$$

- $\mathbf{x}_0 = 0$
- $2^n - 1$ вершин — экспоненциально много!

Как получить полиномиальный метод?

Как получить полиномиальный метод?

- Приближаться к нужной вершине изнутри множества
- Одна итерация по сложности сравнима с методом Ньютона
- Очень быстрая сходимость

Резюме

- Постановки и преобразования задач линейного программирования
- Свойства допустимого множества
- Симплекс-метод и его свойства