Линейное программирование
Метод внутренней точки

Роланд Хильдебранд

Методы оптимизации, ФУПМ МФТИ, май 2022 г.

План

- допустимый метод внутренней точки
- поиск допустимой точки
- недопустимые методы
- приложение: оптимальное распределение ресурсов
- приложение: восстановление разреженного сигнала
- приложение: задача о максимальном потоке

Условия оптимальности

рассмотрим прямо-двойственную пару задач ЛП

$$\min_{x \geq 0} \langle c, x \rangle : \qquad Ax = b$$

$$\max_{s \geq 0, y} \langle b, y \rangle : \qquad s + A^T y = c$$

условия оптимальности для прямо-двойственной пары (x, s, y):

- $Ax = b$
- $s + A^T y = c$
- $x \geq 0$
- $s \geq 0$
- $x_i s_i = 0$ для всех i

Идея метода

сохраняем справедливость условий допустимости

приближаем условие комплементарности условием

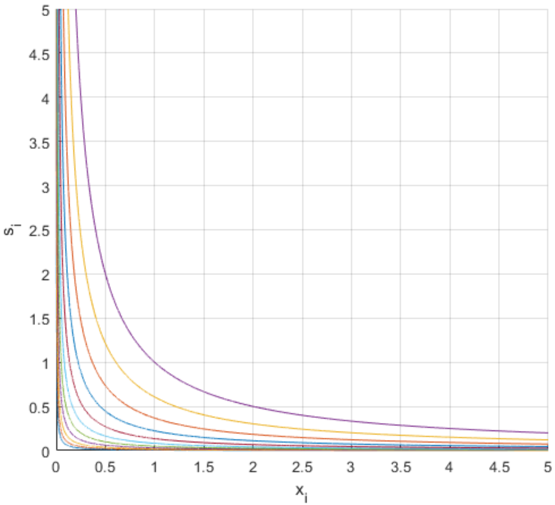
$$x \bullet s \approx \mu \cdot 1, \qquad \mu \to 0$$

здесь $\bullet$ — по-элементное умножение
пытаемся стремить произведения $x_i s_i$ к нулю с одной и той же скоростью

система уравнений

$$Ax = b, \quad s + A^T y = c, \quad x \bullet s = \mu \cdot 1$$

решается приближённо методом Ньютона



аппроксимации допустимого множества условия комплементарности $x_i s_i = 0$ гиперболами

Лемма

Допустим, что прямая и двойственная ЛП строго допустимы. Тогда для любого $z \in \mathbb{R}^n_{++}$ существует единственная строго допустимая пара точек (x, s) такая, что $z = x \bullet s$.

с другой стороны, для любой такой пары точек имеем $x \bullet s > 0$

это означает, что квадратичное отображение

$$(x, s) \mapsto x \bullet s$$

есть биекция внутренности прямого произведения $X_P \times X_D$ допустимых множеств прямой и двойственной ЛП на открытый ортант

Доказательство

пусть дано $z > 0$

рассмотрим задачу

$$\min_{x \geq 0} \left(\langle c, x \rangle - \sum_{i=1}^{n} z_i \log x_i \right) : \quad Ax = b$$

функция цены строго выпукла, допустимое множество то же, что в прямой ЛП,
на границе функция равна $+\infty$

если задача неограничена снизу, то существует рецессивное направление $\delta \geq 0$, $A\delta = 0$, на котором $\langle c, \delta \rangle \leq 0$

для допустимого $s = c - A^T y$ получаем $\langle s, \delta \rangle = \langle c, \delta \rangle \leq 0$,
противоречие со строгой допустимостью двойственной задачи

значит сущестует единственное строго допустимое решение

Доказательство

Лагранжиан имеет вид

$$\mathcal{L} = \langle c, x \rangle - \sum_{i=1}^{n} z_i \log x_i - \langle y, Ax - b \rangle$$

условие оптимальности

$$c - \Delta_z x^{-1} - A^T y = 0$$

точка $s = \Delta_z x^{-1}$ строго допустима, и $z = x \bullet s$

с другой стороны, если $s = \Delta_z x^{-1}$ допустима, то x удовлетворяет условию оптимальности и совпадает с решением

Центральный путь

решение — прообраз точки $z = 0$ при биекции

$$(x, s) \mapsto z = x \bullet s$$

метод находит предел прообраза луча

$$z = \{\mu \cdot 1 \mid \mu \in \mathbb{R}_{++}\}$$

эта кривая в пространстве (x, s) называется *центральным путём*

центральный путь параметризован параметром $\mu > 0$

для того, чтобы найти точку (x^*_μ, s^*_μ) на центральном пути, нужно решить систему

$$Ax = b, \quad s + A^T y = c, \quad x_i s_i = \mu \; \forall \; i$$

пусть текущая точка (x, s, y) строго допустима, тогда смещения удовлетворяют

$$A\delta_x = 0, \quad \delta_s + A^T \delta_y = 0, \quad \Delta_x \delta_s + \Delta_s \delta_x + \delta_x \bullet \delta_s = \mu \cdot 1 - x \bullet s$$

линеаризацией получаем систему

$$\begin{pmatrix} 0 & A & 0 \\ A^T & 0 & I \\ 0 & \Delta_s & \Delta_x \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_x \\ \delta_s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \mu \cdot 1 - x \bullet s \end{pmatrix}$$

Направление Ньютона

умножаем последнюю строку на $A\Delta_s^{-1}$

$$\begin{pmatrix} 0 & A & 0 \\ A^T & 0 & I \\ 0 & A & A\Delta_{x/s} \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_x \\ \delta_s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ A(\mu \cdot s^{-1} - x) \end{pmatrix}$$

устраняем δ_x

$$\begin{pmatrix} A^T & I \\ 0 & A\Delta_{x/s} \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_s \end{pmatrix} = \begin{pmatrix} 0 \\ A(\mu \cdot s^{-1} - x) \end{pmatrix}$$

умножаем первую строку на $A\Delta_{x/s}$

$$\begin{pmatrix} A\Delta_{x/s}A^T & A\Delta_{x/s} \\ 0 & A\Delta_{x/s} \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_s \end{pmatrix} = \begin{pmatrix} 0 \\ A(\mu \cdot s^{-1} - x) \end{pmatrix}$$

устраняем δ_s

$$A\Delta_{x/s}A^T\delta_y = A(x - \mu \cdot s^{-1})$$

Направление Ньютона

система имеет размерность m, равную числу условий равенства

если $n - m \ll m$, имеет смысл перейти к двойственной задаче

матрица системы положительно определена, можно решать через факторизацию Холецкого

получаем решение

$$\delta_y = (A\Delta_{x/s}A^T)^{-1}A(x - \mu \cdot s^{-1})$$
$$\delta_s = -A^T\delta_y,$$
$$\delta_x = \mu s^{-1} - x - \Delta_{x/s}\delta_s$$

но $(x + \delta_x, s + \delta_s)$ может не удовлетворять условию положительности

Широкая окрестность

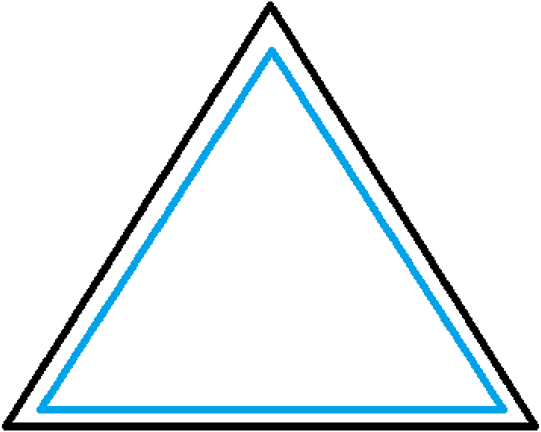
мы не хотим слишком далеко уходить от центрального пути

определим окрестность

$$N_\gamma = \{(x, s) \in X_P \times X_D \mid \min_i (x_i s_i) \geq \gamma \psi(x, s)\}$$

где $\psi(x, s) = \frac{\langle x,s \rangle}{n}$ — среднее значение произведений $x_i s_i$
типичное значение $\gamma = 10^{-3}$

N_γ является прообразом симплициального конуса в пространстве $z = x \bullet s$, который получается смещением фасадов ортанта внутрь



сечение ортанта
сечение образа N_γ

Длина шага

шаг имеет вид

$$(x, s, y) \leftarrow (x', s', y') = (x + \alpha\delta_x, s + \alpha\delta_s, y + \alpha\delta_y)$$

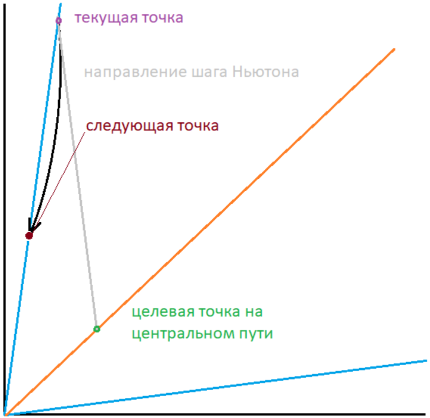
где α — максимальное значение, при котором $(x', s', y') \in N_\gamma$

$z = x' \bullet s'$ — квадратична по α,

$$\psi(x', s') = \frac{\langle x + \alpha\delta_x, s + \alpha\delta_s \rangle}{n} = \psi(x, s) + \frac{\alpha(\langle x, \delta_s \rangle + \langle \delta_x, s \rangle)}{n}$$

линейна

нужно решить n квадратичных уравнений



образ в пространстве переменных z является параболой, которая снова приходит на границу окрестности N_γ

Другие параметры алгоритма

обновление μ

- длина центрального пути пропорциональна $-\log \mu$
- малое μ означает далёкую целевую точку — линейная аппроксимация менее точная
- медленное понижение μ означает небольшой прогресс

можно полагать $\mu = \theta \cdot \psi(x, s)$, $\theta \in (0, 1)$

на центральном пути имеем

$$\psi(x^*(\mu), s^*(\mu)) = \mu$$

с другой стороны

$$\langle c, x \rangle - \langle b, y \rangle = \langle s, x \rangle = n \cdot \psi(x, s)$$

двойственный зазор стремится к нулю

обычно критерий останова — когда двойственный зазор достигнет $\sqrt{\epsilon} \approx 10^{-8}$

далее алгоритм теряет устойчивость

Фаза 1

бывает ситуация, когда строго допустимой точки не имеется

для построения прямой допустимой точки можно решать вспомогательную задачу

$$\min_{x \geq 0, \tau} \tau : \qquad Ax = \tau A x_0 + (1-\tau) b$$

где $x_0 > 0$ — произвольная точка

тогда пара $(x, \tau) = (x_0, 1)$ строго допустима для вспомогательной задачи

любая строго допустимая точка $(x, \tau) = (\hat{x}, 0)$ даёт строго допустимую точку $\hat{x}$ для прямой ЛП

если оптимальное значение $\tau^* > 0$, то задача недопустима для построения двойственной допустимой точки можно решать вспомогательную задачу

$$\min_{s \geq 0, y, \tau} \tau : \qquad s + A^T y = \tau s_0 + (1 - \tau) c$$

где $s_0 > 0$ — произвольная точка

тогда пара $(s, y, \tau) = (s_0, 0, 1)$ строго допустима для вспомогательной задачи

любая строго допустимая точка $(s, y, \tau) = (\hat{s}, \hat{y}, 0)$ даёт строго допустимую точку $(\hat{s}, \hat{y})$ для двойственной ЛП

если оптимальное значение $\tau^* > 0$, то задача недопустима

Недопустимая внутренняя точка

можно сразу начать основную фазу с недопустимой начальной точки

тогда в построении системы для вычисления направления надо учитывать невязки

$$\begin{pmatrix} 0 & A & 0 \\ A^T & 0 & I \\ 0 & \Delta_s & \Delta_x \end{pmatrix} \begin{pmatrix} \delta_y \\ \delta_x \\ \delta_s \end{pmatrix} = \begin{pmatrix} b - Ax \\ c - s - A^T y \\ \mu \cdot 1 - x \bullet s \end{pmatrix}$$

шаг с коэффициентом α умножает невязки на множитель $(1 - \alpha)$

в частности, полный шаг приводит к допустимым точкам, и далее они уже остаются допустимыми

длина шага выбирается при условии что невязки сокращаются как минимум пропорционально μ

Оптимальное распределение ресурсов

рассмотрим задачу планировки распределения ресурсов $k = 1, \dots, K$ для производства продуктов $l = 1, \dots, n$

цены продуктов $p_1, \dots, p_n$

для производства единицы продукта l нужно a_{kl} единиц ресурса k

в наличие имеется $r_1, \dots, r_K$ единиц ресурсов

нужно найти оптимальные количества $x_1, \dots, x_n$ продуктов, максимизирующие выручку

задача формулируется в виде ЛП

$$\min_x(-\langle p, x\rangle) : \qquad Ax \leq r, \quad x \geq 0$$

если продукты производятся по-штучно, получаем смешанно-целочисленную ЛП пусть x — разреженный сигнал

задача — восстановить x из зашумлённого линейного образа

$$y = Ax + \xi$$

где $A \in \mathbb{R}^{m \times n}$ и шум ξ ограничен δ в $\|\cdot\|_\infty$ норме

здесь $m \ll n$, т.е. система недоопределена

нужно использовать информацию о разреженности x

Восстановление разреженного сигнала

в идеале нужно решить проблему

$$\min_x \|x\|_0 : \qquad \|Ax - y\|_\infty \leq \delta$$

где $\|x\|_0 := \#\{i \mid x_i \neq 0\}$

линейная релаксация проблемы получается заменой 0-"нормы" на 1-норму:

$$\min_x \|x\|_1 : \qquad \|Ax - y\|_\infty \leq \delta$$

эта задача формулируется в виде ЛП

$$\min_{x,t} \langle 1, t\rangle : \qquad -t \leq x \leq t, \quad -\delta \leq Ax - y \leq \delta$$

Задача о максимальном потоке

пусть дан граф с n вершинами и весами W_{ij} на гранях

граф моделирует систему труб с пропускными способностями W_{ij}

требуется найти пропускную способность сети из выделенной вершины 1 в вершину n

пусть $F_{ij} = -F_{ji}$ — поток из вершины i в вершину j

тогда задача формулируется в виде ЛП

$$\max_{F=-F^T} \sum_{i=1}^{n} F_{1i} : \qquad F \leq W, \quad \sum_{i=2}^{n-1} F_{ji} = 0 \ \forall \ j$$

рассмотрим граф с весами $W_{ij} = W_{ji} \geq 0$ на гранях и с двумя выделенными вершинами $1, n$

необходимо разделить вершины на две группы S, T, в каждой из которых по одной выделенной, и при этом минимизировать суммарный вес граней, соединяющих S, T

задача эквивалентна задаче о максимальном потоке

пусть F^* — оптимальный поток, $\tilde{W} = W - F^*$ — вспомогательная матрицы ограничений

тогда S — множество вершин, *достижимых* из вершины 1 с помощью потока $\tilde{F} = -\tilde{F}^T$ с ограничением $\tilde{F} \leq \tilde{W}$