Полу-определённое программирование
Приложения

Roland Hildebrand

LJK, Université Grenoble Alpes / CNRS

Методы оптимизации, ФУПМ МФТИ, апрель 2021 г.

Сперва небольшое отступление:

направления Нестерова-Тодда для ЛП

Основные объекты

прямая программа

$$\min_{x \in K} \langle c, x \rangle \quad : \qquad Ax = b$$

двойственная программа

$$\max_{s \in K^*} \langle b, z \rangle \quad : \qquad s = c - A^T z$$

- $M = \{(x, -F'(x)) \mid x \in K^o\}$ — граф лежандровой изометрии
- $\mathcal{A} = \{(x, s) \mid Ax = b, \quad \exists\, y : \ s = c - A^T y\}$ — аффинное подпространство ограничений
- $L = \ker\ A,\ L^\perp = Im\, A^T$ линейные подпространства допустимых смещений
- прямо-двойственный центральный путь $\mathbb{R} \cdot M \cap \mathcal{A}$

Параметр центрального пути

если $(x, s) \in \sqrt{\mu} M \cap \mathcal{A}$, то x, s — точки на прямом и двойственном центральных путях с параметром $\tau = \mu^{-1}$

для $(x, s) \in M$ имеем

$$\langle x, s \rangle = -\langle x, F'(x) \rangle = \nu$$

в силу логарифмичной однородности F
отсюда $\langle \sqrt{\mu} x, \sqrt{\mu} s \rangle = \mu\nu$

если $(x, s) \in \sqrt{\mu} M \cap \mathcal{A}$ на прямо-двойственном центральном пути, то $\mu = \frac{\langle x, s \rangle}{\nu}$

Направления поиска Нестерова-Тодда

для произвольной пары $(x, s) \in \mathcal{A}$ определим $\mu = \frac{\langle x,s \rangle}{\nu}$

имеем $(x, -\mu F'(x)) \in \sqrt{\mu}M$, $(-\mu F'_*(s), s) \in \sqrt{\mu}M$
определим невязки $\Delta_s = s + \mu F'(x)$, $\Delta_x = x + \mu F'_*(s)$
в общем случае $\Delta_x \notin L$, $\Delta_s \notin L^\perp$

поэтому проектируем на подпространства допустимых точек

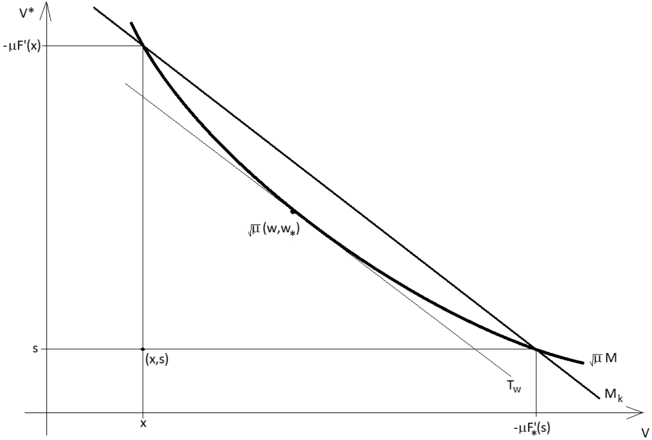
проекция ортогональна в локальной метрике $\|\cdot\|_w$ $(\|\cdot\|_{w_*})$
w — точка шкалировки: $F''(w)x = s$,
$w_* = -F'(w)$ — двойственный образ

направления *центрирующей шкалировки* (d_x, d_s) и *аффинной шкалировки* (p_x, p_s) Нестерова-Тодда определены как

$$d_x = -\Pi_{w,L}\Delta_x, \qquad d_s = -\Pi_{w_*,L^\perp}\Delta_s$$

$$p_x = -\Pi_{w,L}x, \qquad p_s = -\Pi_{w_*,L^\perp}s$$

Направления шкалировок



плоскость M_k параллельна к касательной плоскости T_w в точке $\sqrt{\mu}(w, w_*)$ и проходит через $(x, -\mu F'(x))$, $(-\mu F_*'(s), s)$

Направление Нестерова-Тодда для ЛП

прямая программа

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

двойственная программа

$$\max_{s \geq 0} \langle b, z \rangle \quad : \qquad s = c - A^T z$$

- барьер $F(x) = -\sum_{i=1}^{n} \log x_i$
- двойственный образ $s = -F'(x) = 1/x$
- текущая итерация (x, s), $Ax = b$, $s = c - A^T y$
- точка шкалировки $(w, w_*) = (\sqrt{x/s}, \sqrt{s/x})$
- гессиан $F''(w) = \text{diag}(1/w^2) = \text{diag}(s/x)$
- $\mu = \langle x, s \rangle / n$

Проекции $\Pi_{w,L}$, $\Pi_{w_*,L^\perp}$

пусть $W = \operatorname{diag} w$, $W_* = \operatorname{diag} w_*$, $X = \operatorname{diag} x$, $S = \operatorname{diag} s$

метрика в прямом пространстве $F''(w) = SX^{-1}$
метрика в двойственном пространстве $F_*''(w_*) = XS^{-1}$

разложение прямого пространства

$$\ker A \oplus XS^{-1}[Im\, A^T]$$

проекция на $\ker A$: $z \mapsto z - XS^{-1}A^T(AXS^{-1}A^T)^{-1}Az$

разложение двойственного пространства

$$Im\, A^T \oplus SX^{-1}[\ker A]$$

проекция на $Im\, A^T$: $z \mapsto A^T(AXS^{-1}A^T)^{-1}AXS^{-1}z$

Направления шкалировок

невязки имеют вид

$$\Delta_s = s - \mu X^{-1}\mathbf{1}, \qquad \Delta_x = x - \mu S^{-1}\mathbf{1}$$

направления имеют вид

$$\begin{aligned}
d_x &= -(I - XS^{-1}A^T(AXS^{-1}A^T)^{-1}A)(x - \mu S^{-1}\mathbf{1}) \\
&= XS^{-1}A^T(AXS^{-1}A^T)^{-1}b - x + \mu(I - XS^{-1}A^T(AXS^{-1}A^T)^{-1}A)S^{-1}\mathbf{1} \\
d_s &= -A^T(AXS^{-1}A^T)^{-1}AXS^{-1}(s - \mu X^{-1}\mathbf{1}) \\
&= -A^T(AXS^{-1}A^T)^{-1}b + \mu A^T(AXS^{-1}A^T)^{-1}AS^{-1}\mathbf{1} \\
p_x &= -(I - XS^{-1}A^T(AXS^{-1}A^T)^{-1}A)x \\
&= -x + XS^{-1}A^T(AXS^{-1}A^T)^{-1}b \\
p_s &= -A^T(AXS^{-1}A^T)^{-1}AXS^{-1}s = -A^T(AXS^{-1}A^T)^{-1}b
\end{aligned}$$

Методы с длинным шагом

методы с длинным шагом идут по направлению аффинной шкалировки, пока не достигнут границы некоторой «большой» окрестности центрального пути

возможна корректировка: после «длинного» шага следует один или несколько корректирующих шагов по центрирующему направлению, пока итерации не приблизятся в достаточной мере к центральному пути

возможна комбинация аффинного и центрирующего направлений

«большая» окрестность: произведение xs (для симметрических конусов) имеет ограниченное число обусловленности

- LP: $\frac{\max_i x_i s_i}{\min_i x_i s_i} \leq \gamma$
- SDP: $\frac{\lambda_{\max} XS}{\lambda_{\min} XS} \leq \gamma$

- $||x||_2^2 \leq t,\ x \in \mathbb{R}^n \Leftrightarrow \left(x, \frac{t-1}{2}, \frac{t+1}{2}\right) \in L^{n+2}$
- $\frac{||x||_2^2}{s} \leq t,\ s \geq 0,\ x \in \mathbb{R}^n \Leftrightarrow \left(x, \frac{t-s}{2}, \frac{t+s}{2}\right) \in L^{n+2}$
- $ts \geq 1,\ t, s > 0 \Leftrightarrow \left(1, \frac{t-s}{2}, \frac{t+s}{2}\right) \in L^3$
- $x^T A x + b^T x + c \leq t$ при $A \succeq 0 \Leftrightarrow$ $\left(A^{1/2}x, \frac{t-b^Tx-c-1}{2}, \frac{t-b^Tx-c+1}{2}\right) \in L^{n+2}$
- $|t| \leq \sqrt{x_1 x_2},\ x_1, x_2 \geq 0 \Leftrightarrow \left(t, \frac{x_1-x_2}{2}, \frac{x_1+x_2}{2}\right) \in L^3$
- $t \leq \sqrt{x_1 x_2},\ x_1, x_2 \geq 0 \Leftrightarrow t \leq s,\ s \geq 0,\ \left(s, \frac{x_1-x_2}{2}, \frac{x_1+x_2}{2}\right) \in L^3$

итерацией последнего пункта получаем также представление множества

$$\left\{(x_1, \dots, x_{2^k}, t) \in \mathbb{R}_+^{2^k+1} \mid \prod_{i=1}^{2^k} x_i \geq t^{2^k}\right\}$$

Полу-определённые ограничения

- $\lambda_{\max}(X) \leq t \Leftrightarrow tI - X \succeq 0$
- $||X||_\infty \leq t \Leftrightarrow -tI \preceq X \preceq tI$ для симметрических X
- $\sum_{j=1}^k \lambda_j \leq t$, где $\lambda_1, \ldots, \lambda_n$ — собственные значения $X \Leftrightarrow$ $t \geq ks + \mathrm{tr}\, Z$, $Z \succeq 0$, $Z + sI \succeq X$
- $A \succeq BC^\dagger B^T$, $C \succeq 0 \Leftrightarrow \begin{pmatrix} A & B \\ B^T & C \end{pmatrix} \succeq 0$
- $||A||_\infty \leq t \Leftrightarrow \begin{pmatrix} tI & A \\ A^T & tI \end{pmatrix} \succeq 0$ для матриц A общего вида
- $(AXB)(AXB)^T + CXD + (CXD)^T + E \preceq Y \Leftrightarrow$ $\begin{pmatrix} I & (AXB)^T \\ AXB & Y - E - CXD - (CXD)^T \end{pmatrix} \succeq 0$ (здесь X, Y — переменные, а $A, \ldots, E$ — параметры задачи)

Полу-определённые ограничения

для данной матрицы $X \succeq 0$ множество $\{\eta \in \mathbb{R}^n_+ \mid \eta \leq \lambda(X)\}$ полу-определённо представимо в виде

$$\begin{pmatrix} X & \Delta \\ \Delta^T & \operatorname{diag} \eta \end{pmatrix} \succeq 0, \ \operatorname{diag} \Delta = \eta, \ \Delta_{ij} = 0 \ \forall \ i < j$$

S-лемма: $x^T Ax \geq 0$ для всех x таких, что $x^T Bx \geq 0$ и существует x_0 такое, что $x_0^T Bx_0 > 0 \Leftrightarrow A - \lambda B \succeq 0$ и $\lambda \geq 0$

здесь матрица A и скаляр λ являются переменными, а матрица B — параметром задачи

Максимальный вписанный эллипсоид

пусть $P = \{x \mid Ax \leq b\}$ — политоп с непустой внутренностью

требуется найти вписанный в P эллипсоид

$$E = \{x = Cu + c \mid \|u\| \leq 1\}$$

максимального объёма

задача решается полу-определённой программой

$$\max\ t: \quad t^{2^k} \leq \prod_{i=1}^{n} \lambda_i(C), \ \ C \succeq 0, \ \ \|Ca_i\| \leq b_i - \langle a_i, c \rangle \ \forall\ i$$

где $n \leq 2^k$, а a_i — строки матрицы A

для того, чтобы выразить ограничение $t^{2^k} \leq \prod_{i=1}^{n} \lambda_i$, нужно воспользоваться

- представлением подграфика геометрического среднего неотрицательных величин
- представлением подграфика спектра неотрицательно определённой матрицы

Минимальный описанный эллипсоид

пусть $P = \text{conv}\{x_1, \dots, x_m\}$ — политоп с непустой внутренностью

требуется найти описанный вокруг P эллипсоид

$$E = \{x \mid (x - D^{-1}d)^T D(x - D^{-1}d) \leq 1\}$$

минимального объёма

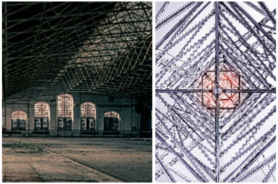
задача решается полу-определённой программой

$$\max t : \ t^{2^k} \leq \prod_{i=1}^{n} \lambda_i(D), \ D \succeq 0, \ \begin{pmatrix} s & d^T \\ d & D \end{pmatrix} \succeq 0,$$

$$x_i^T D x_i - 2x_i^T d + s \leq 1 \ \forall \ i$$

где $n \leq 2^k$

Оптимизация топологии фермы



стержневая система в строительной механике, состоящая из n узлов и соединяющих стержней

при приложении силы к некоторому подмножеству узлов конструкция деформируется и запасает энергию

возникают компенсирующие усилия растяжения–сжатия, ферма приходит в новое состояние равновесия

задача: для данной нагрузки минимизировать энергию, или максимизировать жёсткость
переменные: массы стержней при ограниченной суммарной массе

Оптимизация топологии фермы

пусть стержень k массы m_k с модулем Юнга c_k соединяет узлы i, j на позициях $v_i, v_j \in \mathbb{R}^3$

смещения $x_i, x_j \in \mathbb{R}^3$ узлов i, j вызывают силы

$$f_{ik} = -f_{jk} = -c_k m_k \left\langle \frac{v_i - v_j}{\|v_i - v_j\|^2}, x_i - x_j \right\rangle \frac{v_i - v_j}{\|v_i - v_j\|^2}$$

определим вектор $b_k = \{b_{ik}\}_{i=1,\dots,n} \in \mathbb{R}^{3n}$, где

$b_{ik} = -b_{jk} = \sqrt{c_k} \frac{v_i - v_j}{\|v_i - v_j\|^2}$, остальные $b_{lk} = 0$

тогда $f_k = -m_k b_k b_k^T x$ — сила, генерируемая стержнем k

прилагаемая внешняя сила задаётся

$$f = -\sum_k f_k = \sum_k m_k b_k b_k^T x = Ax$$

минимизируемая энергия деформации на подпространстве $V \subset \mathbb{R}^{3n}$ допустимых смещений задаётся

$$c_f(m) = \langle f, x \rangle = \sup_{u \in V} \left(2\langle f, u \rangle - u^T A u \right).$$

компоненты силы f известны в узлах $l \in L$, в которых прилагается нагрузка g_l, и неизвестны в узлах $j \in J$, зафиксированных в опорах
в остальных узлах $i \in I$ они равны нулю

таким образом получаем полу-определённую программу

$$\min_{t,m_k,(f_j)_{j\in J}} t: \quad \Pi \begin{pmatrix} t & -f^T \\ -f & \sum_k m_k b_k b_k^T \end{pmatrix} \Pi^T \succeq 0,$$

$$\sum_k m_k = m, \; f_l = g_l, \; l \in L, \; f_i = 0, \; i \in I, \; \Pi = \mathrm{diag}(1, \Pi_V)$$

здесь Π_V — проектор на подпространство V

Поиск функции Ляпунова

рассмотрим линейную динамическую систему

$$\dot{x}(t) = A(t)x(t)$$

про которую известно, что матрица системы $A(t)$ в любой момент времени принадлежит некоторому множеству $\mathcal{U}$

квадратичная функция $L(x) = x^T Xx$, $X \succ 0$, является функцией Ляпунова системы, если

$$\frac{d}{dt}L = x^T(A^T X + XA)x \leq -sL \quad \forall\ t$$

для некоторого $s > 0$

в этом случае система устойчива и любая её траектория стремится к началу координат

это условие выполняется, если

$$A^T X + XA \preceq -sX \quad \forall\ A \in \mathcal{U}$$

условие эквивалентно условиям

$$X \succ 0, \quad A^T X + XA \preceq Z \prec 0 \quad \forall\, A \in \mathcal{U}$$

обе матрицы X, Z можно ограничить кратными единичной матрице

пусть $\mathcal{U} = \mathrm{conv}\{A_1, \dots, A_m\}$ — политоп

тогда поиск функции Ляпунова сводится к полу-определённой программе

$$\min\, s : \;\; X \succeq I, \;\; sI - A_i^T X - XA_i \succeq 0 \;\forall\, i$$

Задача из теории управления

рассмотрим линейную управляемую систему

$$\dot{x} = Ax + Bu$$

здесь x — вектор состояния системы, u — вектор управления

необходимо построить линейный закон управления $u = Kx$, стабилизирующий систему

ищем квадратичную функцию Ляпунова $L(x) = x^T Xx$, $X \succ 0$, удовлетворяющую условию

$$\frac{d}{dt}L = x^T((A+BK)^T X + X(A+BK))x \leq -sL = -sx^T Xx$$

для некоторого $s > 0$

Задача из теории управления

проблему можно решить полу-определённой программой

$$\min s: \; sI - (AY + BZ) - (AY + BZ)^T \succeq 0, \; Y \succeq I$$

функция Ляпунова и управление восстанавливаются по формуле

$$X = Y^{-1}, \; K = ZX$$

билинейность по X, K устраняется сопряжением с $X^{-1} = Y$

$$X(A + BK) = Y^{-1}(AY + BZ)Y^{-1}$$

рассмотрим линейную программу

$$\min \langle c, x \rangle : \quad Ax \leq b$$

где данные (A, b) не известны достоверно, а задаются по формуле

$$A = A_0 + \sum_{i=1}^{m} u_i A_i, \quad b = b_0 + \sum_{i=1}^{m} u_i b_i, \quad \|u\| \leq 1$$

необходимо решить *робастную версию* программы

$$\min \langle c, x \rangle : \quad A(u)x \leq b(u) \quad \forall \ \|u\| \leq 1$$

Робастное линейное программирование

рассмотрим робастную версию одного линейного неравенства $\langle a, x\rangle \leq \beta$:

$$\left\langle a_0 + \sum_{i=1}^m u_i a_i, x \right\rangle \leq \beta_0 + \sum_{i=1}^m u_i \beta_i \ \forall \ \|u\| \leq 1$$

условие перепишется в виде

$$\sum_{i=1}^m u_i(\langle a_i, x\rangle - \beta_i) \leq \beta_0 - \langle a_0, x\rangle \ \forall \ \|u\| \leq 1$$

$$\|(\langle a_i, x\rangle - \beta_i)_{i=1,\dots,m}\| \leq \beta_0 - \langle a_0, x\rangle$$

и становится конично-квадратичным ограничением

формулируя это ограничение для каждого неравенства исходной программы, получаем конично-квадратичную задачу

Спасибо за внимание