1 Релаксации и представимость множеств

На этой лекции мы рассмотрим разные классы задач оптимизации. Задачи некоторых классов можно преобразовать в задачи из других стандартных классов. В частности, мы рассмотрим теорию преобразований конических программ друг в друга. Приведение задачи к стандартному виду дает возможность применить сольверы для ее решения. Если задача не эквивалентна задаче стандартного вида, ее можно попытаться аппроксимировать. Мы рассмотрим некоторые случаи, в которых можно оценить ошибку, сделанную при аппроксимации.

В этой части 4 упражнения.

1.1 Задачи оптимизации стандартного вида

Задача оптимизации в самой общей форме имеет вид

$$\min_{x \in X} f(x),$$

где x — искомая переменная, X — допустимое множество, f — целевая функция. Данные задачи могут быть определены разными способами.

Функция цены f может быть задана

- аналитическим выражением
- оптимальным значением другой задачи оптимизации, параметризованной x
- черным ящиком, выдающим для каждого данного $x \in X$ значение, (суб-)градиент, гессиан

Допустимое множество X может быть задано

- равенствами $f_i(x) = 0$
- неравенствами $g_i(x) \leq 0$
- матричными неравенствами $A_i(x) \succeq 0$
- включениями $x_i \in \{0, 1\}$, $x_i \in \mathbb{Z}$, ...
- черным ящиком, выдающим $x \in X$ или нет (отделяющую плоскость)

Выражения, задающие функцию цены и допустимое множество, могут быть

- линейными
- квадратичными
- полиномиальными
- тригонометрическими
- дискретными (бинарными, целочисленными)

В зависимости от характера функции цены и ограничений можно выделять разные стандартные классы задач:

- линейные программы (linear program, LP): линейная функция цены и линейные ограничения
- линейные программы с целочисленными ограничениями (mixed-integer linear program, MILP): линейная функция цены, линейные и целочисленные ограничения
- квадратичные программы (quadratic program, QP): квадратичная функция цены, линейные ограничения

- квадратично ограниченные квадратичные программы (quadratically constrained quadratic program, QCQP): квадратичная функция цены, квадратичные ограничения
- полиномиальные задачи: полиномиальная функция цены, полиномиальные ограничения
- геометрические программы (geometric program, GP): позиномиальная функция цены, позиномиальные ограничения

Одна и та же задача оптимизации может быть сформулирована разными, эквивалентными способами.

- Функцию цены можно преобразовать в линейную, заменив

$$\min_x f(x)$$

на

$$\min_{x,t} t \quad : \quad t \geq f(x).$$

- Бинарные ограничения $x_i \in \{0,1\}$ или $x_i \in \{-1,+1\}$ можно заменить на квадратичные $x_i^2 = x_i$ или $x_i^2 = 1$, соответственно.
- Полиномиальные ограничения можно преобразовать в квадратичные заменой мономов x^α на дополнительные переменные x_α, удовлетворяющие ограничениям $x_{\alpha+\beta} = x_\alpha x_\beta$, $x_{e_i} = x_i$, где $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ — мульти-индекс, а $x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}$.

Таким образом, например, широкий класс задач сводится к QCQP, имеющей общий вид

$$\min_{x \in \mathbb{R}^n} (x^T A_0 x + 2\langle b_0, x\rangle + c_0) \quad : \quad x^T A_i x + 2\langle b_i, x\rangle + c_i \leq 0, \quad x^T A_j x + 2\langle b_j, x\rangle + c_j = 0. \tag{1}$$

Здесь A_0, A_i, A_j — вещественные симметрические матрицы, b_0, b_i, b_j — вектора, а c_0, c_i, c_j — скаляры.

Упражнение 1.1: Переписать задачу оптимизации

$$\min_{x,y,z} (x + 2yz) \quad : \quad x^3 + y^3 = 1, \ z \in \{0,1\}$$

в виде QCQP.

Сведение исходной задачи к задаче стандартного вида дает некоторое представление о сложности задачи. Например, задача класса QCQP общего вида трудно разрешима. Однако, если матрицы A_0, A_i положительно определены, а $A_j = 0$, то задача выпукла и сводится далее к конично-квадратичной задаче (second-order cone program, SOCP). Задачи класса SOCP эффективно решаются. В противном случае задачу QCQP можно аппроксимировать выпуклой задачей. Несколько вариантов такой аппроксимации мы увидим ниже.

Вообще в первом приближении задача эффективно решается, если она выпукла, т.е. функция цены f и допустимое множество X выпуклы. В ходе курса мы, однако, рассмотрим некоторые примеры, идущие вразрез с этим правилом.

Если задача формально выпукла, то ее можно переписать в виде *задачи конического программирования*

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

где $K \subset \mathbb{R}^n$ — выпуклый конус с непустой внутренностью, не содержащий прямой. Такие конуса называются *регулярными*. Помимо условия принадлежности к конусу, на x налагаются только линейные условия типа равенства. Функция цены также является линейной.

Сложность задачи конического программирования зависит от конуса K. Основным классом конусов, для которых эти задачи эффективно решаются, является класс *симметрических* конусов.

Определение 1. *Регулярный конус $K \subset \mathbb{R}^n$ называется* симметрическим, *если выполняются следующие условия:*

а) для любых двух внутренних точек $x, y \in K^o$ существует линейный автоморфизм A конуса K, который переводит x в y (однородность), и

б) при подходящем выборе базиса в $\mathbb{R}^n$ и соответствующего двойственного базиса в $\mathbb{R}_n = (\mathbb{R}^n)^*$ двойственный конус

$$K^* = \{p \in \mathbb{R}_n \mid \langle x, p\rangle \geq 0 \ \forall \ x \in K\}$$

совпадает с K.

Симметрические конуса полностью классифицированы. В оптимизации важную роль играют следующие семейства симметрических конусов:

- положительный ортант $\mathbb{R}^n_+$
- конус Лоренца $L_n = \{(x_0, \tilde{x}) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid x_0 \geq ||\tilde{x}||\}$ и произведения вида $L_{n_1} \times \cdots \times L_{n_k}$
- вещественный матричный конус $\mathcal{S}^n_+ = \{A \in \mathcal{S}^n \mid A \succeq 0\}$
- эрмитовый матричный конус $\mathcal{H}^n_+ = \{A \in \mathcal{H}^n \mid A \succeq 0\}$

Здесь $\mathcal{S}^n$ и $\mathcal{H}^n$ — пространства вещественных симметричных и комплексных эрмитовых матриц размера $n \times n$, соответственно. Отметим, что последнее рассматривается как векторное пространство над $\mathbb{R}$.

Конические программы над конусами этих семейств эквивалентны задачам линейного программирования (LP), конично-квадратичного программирования (SOCP), полу-определенного программирования (semi-definite program, SDP) с вещественными и комплексными коэффициентами, соответственно.

Преобразование исходной задачи оптимизации к конической программе над конусом из одного из вышеперечисленных семейств гарантирует ее эффективное решение при условии, что размерность задачи при этом не сильно возрастет. Если не удается произвести эквивалентное преобразование, то можно попытаться аппроксимировать задачу конической программой из одного из вышеперечисленных стандартных классов.

Рассмотрим задачу QCQP (1). Введем дополнительную матричную переменную X размера $(n+1) \times (n+1)$ и матрицы коэффициентов

$$C = \begin{pmatrix} c_0 & b_0^T \\ b_0 & A_0 \end{pmatrix}, \quad F_i = \begin{pmatrix} c_i & b_i^T \\ b_i & A_i \end{pmatrix}, \quad G_j = \begin{pmatrix} c_j & b_j^T \\ b_j & A_j \end{pmatrix}.$$

Тогда задачу можно переписать в виде

$$\min_{x,X} \langle C, X\rangle \ : \quad \langle F_i, X\rangle \leq 0, \quad \langle G_j, X\rangle = 0, \quad X = \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix}.$$

В этой формулировке функция цены и все ограничения линейны, кроме последнего. Однако, последнее ограничение можно переписать в виде

$$X_{00} = 1, \quad X_{0k} = x_k, \ k = 1, \dots, n, \quad X \succeq 0, \quad \mathrm{rk}\, X = 1.$$

Все ограничения линейны или полу-определенны, кроме последнего. Это ограничение на ранг невыпукло и делает задачу сложной.

Стандартная полу-определенная релаксация задачи QCQP состоит в опущении ограничения на ранг. Так как допустимое множество увеличивается, оптимальное значение релаксации ограничивает оптимальное значение исходной задачи снизу. Если среди решений (x, X) полу-определенной релаксации найдется такое, что $\mathrm{rk}\, X = 1$, то релаксация точна и компонента x является решением исходной задачи. Этот случай имеет место, если в исходной задачи имеется только одно ограничение. Для доказательства нам понадобится следующий результат [3].

Лемма 1. *Пусть A, B — произвольные симметрические матрицы размера $n \times n$. Тогда их* численный образ *(numerical range), задающийся множеством*

$$W(A,B) = \{(x^T Ax, x^T Bx) \in \mathbb{R}^2 \mid x \in \mathbb{R}^n\}$$

является выпуклым.

Следствие 1. *Если в задаче QCQP имеется только одно ограничение, то ее полу-определенная релаксация точна.*

Доказательство: Пусть ограничение имеет вид $\langle F, X\rangle \le 0$ или $\langle F, X\rangle = 0$.

Пусть $X \in \mathcal{S}_+^{n+1}$ неотрицательно определена. Тогда найдутся векторы $r_1, \dots, r_k \in \mathbb{R}^{n+1}$ такие, что $X = \frac{1}{k}\sum_{j=1}^k r_j r_j^T$. Поэтому

$$(\langle C, X\rangle, \langle F, X\rangle) = \frac{1}{k}\sum_{j=1}^k (r_j^T C r_j, r_j^T F r_j).$$

Но пары $(r_j^T C r_j, r_j^T F r_j)$ являются элементами численного образа $W(C, F)$ для всех $j = 1, \dots, k$. Так как этот численный образ выпуклый, пара $(\langle C, X\rangle, \langle F, X\rangle)$ также является его элементом. Но тогда существует матрица $\tilde{X} = rr^T \in \mathcal{S}_+^{n+1}$ ранга 1 такая, что $(\langle C, X\rangle, \langle F, X\rangle) = (\langle C, \tilde{X}\rangle, \langle F, \tilde{X}\rangle)$.

В частности, если X — решение релаксации, то найдется другое ее решение, имеющее ранг 1. Поэтому релаксация точна. □

1.2 Преобразования конических программ

В этом разделе мы рассмотрим, каким образом можно переписать коническую программу над неким конусом $K \subset V_n$ через другую коническую программу над конусом $K' \subset V_m$, где V_n, V_m — вещественные векторные пространства соответствующих размерностей.

Рассмотрим коническую программу

$$\min_{x \in K} \langle c, x\rangle \quad : \quad x \in A, \tag{2}$$

где $A \subset V_n$ — аффинное подпространство.

Пусть теперь конус K представляется в виде сечения другого конуса K', т.е. существует линейная инъекция $L : V_n \to V_m$ такая, что $L[K] = K' \cap L[V_n]$. Тогда исходная коническая программа (2) эквивалентна программе

$$\min_{y \in K'} \langle c', y\rangle \quad : \quad y \in L[A],$$

где $c' \in V_m^*$ — любой линейный функционал, удовлетворяющий соотношению $\langle c', L(x)\rangle = \langle c, x\rangle$ для всех $x \in V_n$, или эквивалентно $L^\dagger(c') = c$. Здесь $L^\dagger : V_m^* \to V_n^*$ — сопряженное к L линейное отображение, определенное соотношением

$$\langle L^\dagger(z), y\rangle = \langle z, L(y)\rangle \qquad \forall\ y \in V_n,\ z \in V_m^*.$$

Часто выражения, задающие конус K, более или менее явным образом указывают на расширение K', имеющее более простую структуру, чем K. Например, любой полиэдральный конус является сечением ортанта. Линейное сечение матричного конуса называется *спектраэдральным конусом*. Заметим также, что произвольное аффинное сечение матричного конуса называется *спектраэдром*.

Упражнение 1.2: Представить полиэдральный конус $K = \{x \mid Ax \le 0\}$ в виде сечения ортанта.

Менее очевидным является представление $K \in V_n$ линейной проекцией другого конуса $K' \in V_m$. Такое представление задается линейной сюръекцией $\Pi : V_m \to V_n$ такой, что $\Pi[K'] = K$. Пусть задана такая сюръекция, также называющаяся *поднятием* (lifting). Тогда исходная коническая программа (2) эквивалентна программе

$$\min_{y \in K'} \langle \Pi^\dagger(c), y\rangle \ : y \in \Pi^{-1}[A].$$

Часто поднятие можно реализовать вводом дополнительных переменных. В этом случае $V_m = V_n \times V$, а Π определяется как проекция $(x, y) \mapsto x$. Коническая программа над K' принимает вид

$$\min_{(x,y)\in K'} \langle (c, 0), (x, y)\rangle \quad : \quad (x, y) \in A \times V.$$

Две приведенные выше конструкции двойственны друг к другу. А именно, конус K тогда и только тогда является сечением конуса K', когда двойственный конус K^* является проекцией двойственного конуса $(K')^*$.

Упражнение 1.3: Показать, что если сечение $K \subset V_n$ конуса $K' \subset V_m$ задается инъекцией $L : V_n \to V_m$, то сюръекция $\Pi : V_m^* \to V_n^*$, задающая K^* как проекцию $(K')^*$, есть ни что иное, как сопряженное отображение $L^\dagger$.

В общем случае можно комбинировать оба подхода.

Определение 2. *Конус $K \subset V_n$ называется* представимым *через конус $K' \subset V_m$, если существуют векторное пространство V, конус $\tilde{K} \subset V$, линейная инъекция $L : V \to V_m$ и линейная сюръекция $\Pi : V \to V_n$ такие, что $\Pi[\tilde{K}] = K$, $L[\tilde{K}] = L[V] \cap K'$.*

Иными словами, включение $x \in K$ характеризуется существованием элемента $y \in \tilde{K}$ такого, что $\Pi(y) = x$, $L(y) \in K'$. Конус $\tilde{K}$ изоморфен сечению конуса K', а конус K является его проекцией. Коническая программа (2) перепишется в виде

$$\min_{z\in K'} \langle c', z\rangle \quad : \quad z \in L[\Pi^{-1}[A]].$$

Здесь c' — любой линейный функционал, удовлетворяющий соотношению $\langle c', L(y)\rangle = \langle \Pi^\dagger(c), y\rangle$ для всех $y \in V$, или эквивалентно $L^\dagger(c') = \Pi^\dagger(c)$.

С точностью до линейного изоморфизма конус K можно восстановить из K', если заданы два подпространства $V' = L[V] \subset V_m$ и $V_0 = L[\ker \Pi] \subset V'$. Тогда пространство V_n можно отождествить с фактор-пространством V'/V_0, а конус K с множеством $\{[x] \in V'/V_0 \mid x \in K' \cap V'\}$.

Определение 3. *Конус K, представимый через матричный конус $\mathcal{S}_+^n$, называется* полу-определенно представимым *(semi-definite representable, SDr).*

Все симметрические конуса являются спектраэдральными.

- ортант изоморфен множеству диагональных неотрицательно определенных матриц
- конус Лоренца L_n изоморфен множеству неотрицательно определенных матриц вида

$$\begin{pmatrix} x_0 + x_1 & x_2 & \dots & x_{n-1} \\ x_2 & x_0 - x_1 & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ x_{n-1} & 0 & 0 & x_0 - x_1 \end{pmatrix}$$

- эрмитовый матричный конус, состоящий из матриц вида $S + iA$, где S — симметричная, а A — кососимметричная вещественная матрица, изоморфен множеству неотрицательно определенных матриц вида $\begin{pmatrix} S & A \\ -A & S \end{pmatrix}$

Справедлив следующий, более общий результат [2].

Теорема 1.1. *Любой однородный конус является спектраэдральным.*

Если в исходной конической задаче конус является полу-определенно представимым, причем размерность представляющего матричного конуса не слишком большая, то задача сводится к эффективно решаемой полу-определенной программе. Естественно возникает вопрос, каков наименьший размер матричного конуса, или в более общей постановке, конуса из некоторого семейства, который представляет данный конус. Для семейства $\mathbb{R}_+^n$ ответ на этот вопрос дает теорема Яннакакиса.

1.3 Теорема Яннакакиса и ее обобщения

Ответ на вопрос, представим ли данный конус $K \subset V_n$ через ортант $V_m = \mathbb{R}_+^m$ для некоторого m, имеет очевидный ответ. Такое представление существует тогда и только тогда, когда K полиэдральный. Однако установить, какова наименьшая размерность m, которая для этого потребуется, оказывается нетривиальной задачей.

Этот вопрос был изучен Михалисом Яннакакисом в работе [9] в связи с неоднократным выходом в свет "доказательств"соотношения $P = NP$ посредством представления экспоненциально сложных политопов через полиномиально сложные.

Пусть $K \subset V_n$ — регулярный полиэдральный конус, а $K^* \subset V_n^*$ — двойственный к нему конус. Пусть K имеет r экстремальных лучей, а K^* — s экстремальных лучей. Построим матрицы A, B размера $n \times r$ и $n \times s$, соответственно, такие что столбцы A и B генерируют все экстремальные лучи конусов K, K^*, соответственно. Матрица $M = B^T A$ размера $s \times r$ называется *матрицой невязок* (slack matrix). Очевидно элементы этой матрицы неотрицательны. Заметим, что матрица невязок не зависит от выбора системы координат в пространстве V_n, а изменение нумерации экстремальных лучей конусов K, K^* приводит к пермутации ее столбцов и строк.

Определение 4. Неотрицательным рангом *матрицы M размера $s \times r$ с неотрицательными элементами называется минимальное число m такое, что существуют неотрицательные матрицы F, G размера $m \times r$ и $m \times s$, соответственно, такие что $M = G^T F$.*

Вычисление неотрицательного ранга произвольной неотрицательной матрицы является NP-сложной задачей. Понятно, что неотрицательный ранг матрицы невязок зависит только от самого конуса K. Мы имеем следующую связь с представлениями конуса K через ортант.

Лемма 2. *Пусть $M = G^T F$ — неотрицательная факторизация матрицы невязок конуса K, в которой матрицы F, G имеют размер $m \times r$ и $m \times s$, соответственно. Тогда K имеет следующее представление через ортант $\mathbb{R}_+^m$:*

$$K = \{(BB^T)^{-1} BG^T y \mid y \in \mathbb{R}_+^m \cap Im\, F\}.$$

Доказательство: Заметим сначала, что B имеет полный ранг и $r \geq n$, поскольку K и вместе с ним и K^* — регулярный конус. Поэтому BB^T обратимо.

Пусть a_j — j-ый столбец матрицы A, y_j — j-ый столбец матрицы F, а e_j — j-ый базисный вектор в $\mathbb{R}^r$. Заметим, что $y_j \in \mathbb{R}_+^m \cap Im\, F$, поскольку F неотрицательна. Тогда имеем

$$(BB^T)^{-1} BG^T y_j = (BB^T)^{-1} BG^T F e_j = (BB^T)^{-1} BM e_j = (BB^T)^{-1} BB^T A e_j = A e_j = a_j.$$

Из этого следует, что все генераторы экстремальных лучей K, а значит и любые элементы K, имеют требуемое представление.

Пусть теперь y — произвольный вектор из множества $\mathbb{R}_+^m \cap Im\, F$. Покажем, что $x = (BB^T)^{-1} BG^T y$ является элементом K. Для этого достаточно показать, что $B^T x \geq 0$, т.е., x неотрицателен на всех элементах двойственного конуса K^*. Так как $y \in Im\, F$, существует вектор $z \in \mathbb{R}^r$ такой, что $y = Fz$. Из этого следует, что

$$B^T x = B^T (BB^T)^{-1} BG^T y = B^T (BB^T)^{-1} BG^T F z = B^T (BB^T)^{-1} BB^T A z = B^T A z = G^T F z = G^T y \geq 0$$

в силу неотрицательности матрицы G. Этого и требовалось доказать. □

С другой стороны, любое представление конуса K через ортант определяет неотрицательную факторизацию матрицы невязок. Для доказательства этого факта нам понадобится следующая вспомогательная лемма.

Лемма 3. *Пусть $K \subset V_m$ — выпуклый полиэдральный конус, а $L : V \subset V_m$ — линейная инъекция. Пусть $w \in V^*$ — линейный функционал такой, что образ $L[H]$ открытого полу-пространства $H = \{x \in L \mid \langle w, x \rangle < 0\}$ имеет пустое пересечение с K. Тогда существует элемент $g \in K^*$ такой, что $L^\dagger(g) = w$.*

Доказательство: Так как L инъективно, сопряженное отображение $L^\dagger$ будет сюръективным. Поэтому найдется элемент $\tilde{g} \in V_m^*$ такой, что $L^\dagger(\tilde{g}) = w$. Пусть далее $l_1, \dots, l_k \in V_m^*$ генерируют ортогональное подпространство $(L[V])^\perp \subset V_m^*$, т.е. ядро отображения $L^\dagger$, а $b_1,\dots,b_N \in V_m^*$ генерируют двойственный конус K^*, т.е. $K^* = \{y \in V_m \mid \langle b_i, y\rangle \geq 0 \ \forall\ i = 1,\dots,N\}$. Тогда система линейных равенств и неравенств

$$\langle b_i, y\rangle \geq 0 \ \forall\ i = 1,\dots,N, \quad \langle l_j, y\rangle = 0,\ \forall\ j = 1,\dots,k, \quad \langle \tilde g, y\rangle < 0$$

на вектор $y \in V_m$ несовместима. По теореме об альтернативе существуют $\lambda_1,\dots,\lambda_k \in \mathbb{R}$ и $\mu_1,\dots,\mu_N \geq 0$ такие, что $\tilde g - \sum_{i=1}^N \lambda_i b_i + \sum_{j=1}^k \mu_j l_j = 0$. Положим $g = \tilde g + \sum_{j=1}^k \mu_j l_j$. В силу $L^\dagger(l_j) = 0$ имеем $L^\dagger(g) = w$, а в силу $g = \sum_{i=1}^N \lambda_i b_i$ имеем $g \in K^*$, что и требовалось доказать. □

Лемма 4. *Пусть существует представление конуса K через ортант $\mathbb{R}_+^m$. Тогда неотрицательный ранг матрицы невязок не превосходит m.*

Доказательство: Пусть отображения $L : V \to \mathbb{R}^m$, $\Pi : V \to V_n$ задают представление K через $\mathbb{R}_+^m$, т.е. существует конус $\tilde K \subset V$ такой, что $\Pi[\tilde K] = K$, $L[\tilde K] = L[V] \cap \mathbb{R}_+^m$.

Построим неотрицательную матрицу F размера $m \times r$ следующим образом. В качестве j-го столбца F определим вектор $y_j = L(z_j) \in L[V] \cap \mathbb{R}_+^m$, где z_j — любой прообраз из пересечения $\tilde K \cap \Pi^{-1}(a_j)$.

Определим векторы $w_i = \Pi^\dagger(b_i) \in V^*$. Имеем $\langle w_i, z\rangle = \langle b_i, \Pi(z)\rangle \geq 0$ для любого $z \in \Pi^{-1}[K]$, и поэтому $w_i \in (\Pi^{-1}[K])^*$. Вследствие $K = \Pi[\tilde K]$ имеем $\tilde K \subset \Pi^{-1}[K]$, $(\Pi^{-1}[K])^* \subset \tilde K^*$, и $w_i \in \tilde K^*$. Поэтому открытое полу-пространство $H = \{z \in V \mid \langle w_i, z\rangle < 0\}$ имеет пустое пересечение с $\tilde K$, а его образ $L[H]$ имеет пустое пересечение с $\mathbb{R}_+^m$. В силу леммы 3 найдется элемент $g_i \geq 0$ такой, что $L^\dagger(g_i) = w_i$. Построим матрицу G размера $m \times s$ из столбцов g_i.

По построению имеем $\langle g_i, y_j\rangle = \langle w_i, z_j\rangle = \langle b_i, a_j\rangle$ для всех индексных пар (i,j). Это дает нам искомую неотрицательную факторизацию $M = G^T F$ матрицы невязок. □

Комбинируя обе леммы, получаем следующий результат.

Теорема 1.2. *Пусть K — регулярный полиэдральный конус. Тогда наименьшее число m такое, что существует представление K через ортант $\mathbb{R}_+^m$, равно неотрицательному рангу матрицы невязок конуса K.* □

Упражнение 1.4: Построить представление конуса над правильным шестигранником через ортант $\mathbb{R}_+^5$.

Общий случай был исследован в работе [6]. Вместо матрицы невязок регулярному конусу K общего вида соответствует *оператор невязок* (slack operator). Пусть $\mathrm{ext}(K), \mathrm{ext}(K^*)$ — множества генераторов экстремальных лучей конусов K, K^*, соответственно. Тогда оператор невязок конуса K задается отображением $M_K : \mathrm{ext}(K^*) \times \mathrm{ext}(K) \to \mathbb{R}_+$, $M_K : (y,x) \mapsto \langle y, x\rangle$.

Определение 5. *Пусть X, Y — произвольные множества, а $M : Y \times X \to \mathbb{R}_+$ — неотрицательная функция на их произведении. Пусть K — произвольный регулярный конус. Функция M* факторизуется *через K если существуют функции $F : X \to K$, $G : Y \to K^*$ такие, что для всех $x \in X$, $y \in Y$ имеет место соотношение $M(y,x) = \langle G(y), F(x)\rangle$.*

Заметим, что разложение неотрицательной матрицы $M = G^T F$ на неотрицательные факторы F, G размера $m \times r$, $m \times s$ является факторизацией M через ортант $\mathbb{R}_+^m$.

Имеем следующее обобщение леммы 2.

Лемма 5. *Пусть оператор невязок M_K регулярного конуса $K \subset V_n$ факторизуется через регулярный конус $K' \subset V_m$. Тогда K имеет представление через конус K'.*

Доказательство: Рассмотрим линейное пространство

$$V = \{(x,y) \in V_n \times \operatorname{span} Im\, F \mid \langle w, x\rangle = \langle G(w), y\rangle \ \forall\ w \in \mathrm{ext}(K^*)\},$$

его проекцию $\Pi : (x,y) \mapsto x$ на V_n, а также отображение $L : (x,y) \mapsto y$. Покажем, что этим задается представление K через K'. Обозначим $\tilde K = \{(x,y) \in V \mid y \in K'\}$.

Для любого $x \in \mathrm{ext}(K)$ имеем $(x, F(x)) \in V$ по определению M_K. Так как K — регулярный конус, то Π сюръективно. Более того, вследствие $F(x) \in K'$ имеем $(x, F(x)) \in \tilde K$, $x \in \Pi[\tilde K]$, и тем самым $K \subset \Pi[\tilde K]$.

Далее, из $(x, 0) \in V$ следует, что $\langle w, x\rangle = 0$ для всех $w \in \mathrm{ext}(K^*)$, что вследствие регулярности K может иметь место только тогда, когда $x = 0$. Поэтому L инъективно.

Осталось показать, что $\Pi[\tilde{K}] \subset K$. Пусть $y \in K'$, $(x, y) \in \tilde{K}$, $w \in \mathrm{ext}(K^*)$. Тогда $\langle w, x\rangle = \langle G(w), y\rangle \geq 0$, и x неотрицателен на всех экстремальных лучах двойственного конуса K^*. Но тогда $x \in K$, что и требовалось доказать. $\square$

Лемму 4 также можно обобщить, но вследствие отсутствия аналога леммы 3 необходимо потребовать дополнительное условие.

Лемма 6. *Пусть существует представление регулярного конуса $K \subset V_n$ через регулярный конус $K' \subset V_m$ такое, что подпространство $L[V] \subset V_m$ имеет непустое пересечение с внутренностью конуса K'. Тогда оператор невязок M_K конуса K факторизуется через конус K'.*

Доказательство: Отображение F строится как в доказательстве леммы 4, т.е. $a \in \mathrm{ext}(K)$ отображается в $L(z)$, где $z \in \tilde{K} \cap \Pi^{-1}(a)$.

Построим отображение G. Пусть $b \in \mathrm{ext}(K^*)$, $w = \Pi^\dagger(b)$. Пусть функционал $g \in V_m^*$ отделяет конус K' от множества $L[H]$, где $H = \{z \in V \mid \langle w, z\rangle < 0\}$. По построению образ $L^\dagger(g)$ будет пропорционален функционалу w, однако надо исключить случай $L^\dagger(g) = 0$. Это гарантируется дополнительным условием. Если $L^\dagger(g) = 0$, то функционал g равен нулю на всех точках подпространства $L[V]$ и таким образом и на внутренних точках K', что невозможно вследствие регулярности K'. Итак, g можно нормализовать так, что $L^\dagger(g) = w$. Положим $G(b) = g$.

Тогда имеем $\langle G(b), F(a)\rangle = \langle g, L(z)\rangle = \langle w, z\rangle = \langle \Pi^\dagger(b), z\rangle = \langle b, a\rangle$. Таким образом, F, G задают искомую факторизацию.

В случае полу-определенных представлений особенности матричных конусов позволяют получить полный аналог теоремы 1.2. А именно, если конус K представим через матричный конус $\mathcal{S}_+^m$, но подпространство $L[V]$ не содержит положительно определенных матриц, то $K' = \mathcal{S}_+^m$ можно заменить на минимальный фасад, содержащий пересечение $L[V] \cap \mathcal{S}_+^m$, а объемлющее пространство V_m — на линейную оболочку этого фасада. Но этот фасад изоморфен матричному конусу $\mathcal{S}_+^{m'}$ для некоторого $m' < m$, а его линейная оболочка — пространству симметрических матриц $\mathcal{S}^{m'}$ размера $m' \times m'$. Поэтому K представим через $\mathcal{S}_+^{m'}$, но теперь дополнительное условие в формулировке леммы 6 уже выполняется. Для формулировки аналога теоремы 1.2 нам понадобится следующее обобщение неотрицательного ранга.

Определение 6. *Пусть X, Y — произвольные множества, а $M : Y \times X \to \mathbb{R}_+$ — неотрицательная функция на их произведении.* Полу-определенным рангом *функции M называется минимальное число m такое, что существуют отображения $F : X \to \mathcal{S}_+^m$, $G : Y \to \mathcal{S}_+^m$ такие, что $M(y, x) = \langle G(y), F(x)\rangle$ для всех $x \in X$, $y \in Y$.*

В случае, когда M — матрица, свойства полу-определенного ранга были изучены в [4].

Мы получаем следующий результат.

Теорема 1.3. *Пусть K — регулярный конус. Тогда наименьшее число m такое, что существует представление K через матричный конус $\mathcal{S}_+^m$, равно полу-определенному рангу оператора невязок конуса K.* $\square$

В частности, конус K не является полу-определенно представимым, если его оператор невязок не допускает факторизации через матричный конус какой бы то ни было размерности.

Так как нетривиальные фасады конуса Лоренца изоморфны лучу $\mathbb{R}_+ \simeq L_1$, для представления конуса K через конуса Лоренца верен аналогичный результат.

1.4 Аппроксимация полу-определенно представимыми конусами

Представлением одного конуса через другой можно порою сильно понизить сложность задачи. Если представление неизвестно или не сущетвует, то можно попытаться сделать аппроксимацию.

Например, конус над правильным многогранником с N гранями можно представить через ортант размерности $O(\log N)$. Конусом над таким многогранником можно аппроксимировать конус Лоренца L_3 с точностью $\epsilon = O(N^{-1})$, а конус Лоренца L_n можно представить через прямое произведение L_3^{n-2}. В итоге конус Лоренца L_n можно с точностью ϵ аппроксимировать полиэдральным конусом, обладающем представлением через ортант $\mathbb{R}^m$ с $m \sim O(n \log \epsilon)$ [1].

Иногда можно гарантировать некоторую точность определенной релаксации, но нет семейства релаксаций со все возрастающей точностью. Рассмотрим два примера.

Рассмотрим задачу максимизации однородной квадратической формы A на вершинах гиперкуба $[-1,1]^n$. Эта задача записывается в виде квадратичной программы с квадратичными ограничениями

$$\max_{x \in \mathbb{R}^n} x^T A x \quad : \quad x_i^2 = 1, \ \forall\, i = 1, \dots, n.$$

Стандартная полу-определенная релаксация этой задачи получается из эквивалентной формулировки

$$\max_{X \in \mathcal{S}_+^n} \langle A, X \rangle \quad : \quad \operatorname{diag} X = \mathbf{1}, \quad \operatorname{rk} X = 1$$

опущением условия на ранг.

Матрицы $X \succeq 0$ ранга 1 с единичной диагональю являются вершинами политопа максимального разреза (MaxCut polytope), обозначаемого через $\mathcal{MC}$. Этот политоп имеет 2^{n-1} вершин. Множество допустимых матриц в полу-определенной релаксации, т.е., неотрицательно опреледенных матриц с единичной диагональю, содержит политоп $\mathcal{MC}$ и является его внешней аппроксимацией. Это множество обозначается через $\mathcal{SR}$. Кроме того, введем еще *тригонометрическую аппроксимацию* [7]

$$\mathcal{TA} = \{\frac{2}{\pi} \arcsin X \mid X \in \mathcal{SR}\},$$

где функция $t \mapsto \frac{2}{\pi} \arcsin t$ применяется к X по-элементно. Следующий результат показывает, что $\mathcal{TA}$ является внутренней аппроксимацией политопа $\mathcal{MC}$ [8].

Лемма 7. *Пусть $Z \in \mathcal{TA}$, $X = \sin(\frac{\pi}{2} Z) \in \mathcal{SR}$. Пусть $\xi \sim \mathcal{N}(\mathbf{0}, X)$ — нормально распределенный случайный вектор, и пусть $x = \operatorname{sgn} \xi$. Тогда $Z = \mathbb{E}_\xi x x^T$.*

Лемма доказывается прямым счетом.

Так как экстремальные точки политопа $\mathcal{MC}$ содержаться в $\mathcal{TA}$, то из включения $\mathcal{TA} \subset \mathcal{MC}$ также следует равенство $\operatorname{conv} \mathcal{TA} = \mathcal{MC}$.

Теорема 1.4 ($\frac{\pi}{2}$-теорема Нестерова)**.** *Пусть $A \succeq 0$. Тогда имеет место оценка*

$$\frac{2}{\pi} \max_{X \in \mathcal{SR}} \langle A, X \rangle \leq \max_{X \in \mathcal{TA}} \langle A, X \rangle = \max_{X \in \mathcal{MC}} \langle A, X \rangle \leq \max_{X \in \mathcal{SR}} \langle A, X \rangle.$$

Доказательство: Равенство имеет место, поскольку линейная форма на компактном множестве принимает те же экстремальные значения, что и на его выпуклой оболочке. Правое неравенство следует из включения $\mathcal{MC} \subset \mathcal{SR}$. Для доказательства левого неравенства рассмотрим матрицу $X \in \mathcal{SR}$, в которой линейная форма A принимает максимум, и положим $Z = \frac{2}{\pi} \arcsin X \in \mathcal{TA}$. Тогда

$$\langle A, Z \rangle = \frac{2}{\pi} \left\langle A, \sum_{k=0}^{\infty} \frac{(2k-1)!!}{(2k)!!} \frac{X^{2k+1}}{2k+1} \right\rangle \geq \frac{2}{\pi} \langle A, X \rangle,$$

поскольку $\langle A, X^k \rangle \geq 0$ для всех $k \in \mathbb{N}$ вследствие неотрицательной определенности A и X^k. $\square$

Задача нахождения максимального разреза в графе с неотрицательными весами сводится к программе

$$\frac{1}{4} \max_{X \in \mathcal{MC}} \langle L, X \rangle = \frac{1}{4} \max_{X \in \mathcal{MC}} \langle W, \mathbf{1} - X \rangle,$$

где W — неотрицательная по-элементно матрица весов, а $L = \operatorname{diag}(W\mathbf{1}) - W$ — ее лапласиан. Теорема Геманса-Виллиамсона дает следующую оценку ошибки полу-опреледенной релаксации [5].

Теорема 1.5. *Пусть W — неотрицательная по-элементно матрица. Тогда*

$$\gamma \max_{X\in\mathcal{SR}} \langle W, \mathbf{1} - X\rangle \le \max_{X\in\mathcal{TA}} \langle W, \mathbf{1} - X\rangle = \max_{X\in\mathcal{MC}} \langle W, \mathbf{1} - X\rangle \le \max_{X\in\mathcal{SR}} \langle W, \mathbf{1} - X\rangle,$$

где $\gamma = \min_{t\in[-1,1)} \frac{1-\frac{2}{\pi}\arcsin t}{1-t} \approx 0.878$.

Доказательство: Правое неравенство и равенство доказываются как и в предыдущем случае. Для доказательства левого неравенства рассмотрим матрицу $X \in \mathcal{SR}$, в которой выражение $\langle W, \mathbf{1}-X\rangle$ принимает максимум, и положим $Z = \frac{2}{\pi}\arcsin X \in \mathcal{TA}$. Тогда

$$\langle W, \mathbf{1} - Z\rangle = \langle W, \mathbf{1} - \frac{2}{\pi}\arcsin X\rangle \ge \gamma \langle W, \mathbf{1} - X\rangle,$$

поскольку $W \ge 0$. □

Список литературы

[1] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on Modern Convex Optimization - Analysis, Algorithms, and Engineering Applications*, volume 2 of *MPS/SIAM Series on Optimization*. SIAM, 2001.

[2] Chek Beng Chua. Relating homogeneous cones and positive definite cones via T-algebras. *SIAM J. Optim.*, 14(2):500–506, 2003.

[3] Lloyd L. Dines. On the mapping of quadratic forms. *Bull. Amer. Math. Soc.*, 47(6):494–498, 1941.

[4] Hamza Fawzi, Joao Gouveia, Pablo A. Parrilo, Richard Z. Robinson, and Rekha R. Thomas. Positive semidefinite rank. *Math. Program.*, 153:133–177, 2015.

[5] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, Nov 1995.

[6] Joao Gouveia, Pablo A. Parrilo, and Rekha R. Thomas. Lifts of convex sets and cone factorizations. *Math. Oper. Res.*, 38(2):248–264, 2013.

[7] Bernd Hirschfeld. *Approximative Lösungen des Max-Cut Problems mit semi-definiten Programmen*. PhD thesis, Universität Düsseldorf, 2004.

[8] Yurii Nesterov. Semidefinite relaxation and nonconvex quadratic optimization. *Optim. Methods Softw.*, 9(1–3):141–160, 1998.

[9] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *J. Comput. Syst. Sci.*, 43:441–466, 1991.