Полу-определенные релаксации в оптимизации
Лекция 3: релаксации полиномиальных задач

Роланд Хильдебранд

Факультет Управления и Прикладной Математики МФТИ

весна 2020 г.

Выпуклая формулировка общей задачи оптимизации

Положительные полиномы и суммы квадратов

Релаксации типа сумм квадратов для полиномиальных задач

Моментные релаксации

Тригонометрические полиномы

Общая задача оптимизации

рассмотрим задачу

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

$X \subset \mathbb{R}^n$ — измеримое допустимое множество
f — непрерывная функция цены

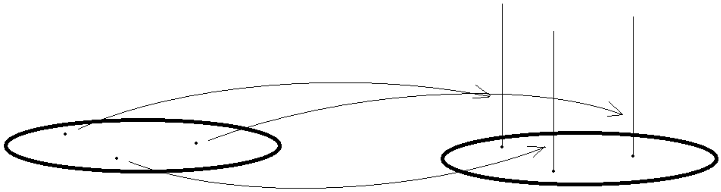
эквивалентная формулировка

$$\min_{\mu \in \mathcal{M}} \int_X f(x) d\mu(x) \;\; : \quad \int_X d\mu(x) = 1$$

$\mathcal{M}$ — конус неотрицательных мер на множестве X

формально невыпуклая задача представлена как
бесконечно-мерная коническая программа

Геометрическая интерпретация

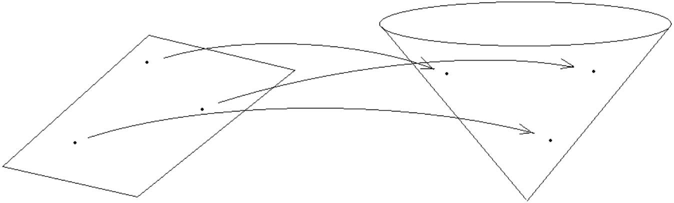


точки множества $X \subset \mathbb{R}^n$ отображаются в δ-функции в пространстве мер на X

нелинейная инъекция Γ устроена таким образом, что $f \circ \Gamma^{-1}$ можно расширить до линейной функции цены

супремум линейной цены по выпуклой оболочке δ-функций равен супремуму f по X

Релаксация задач QCQP



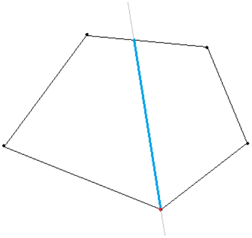
точка $x \in \mathbb{R}^n$ отображается в матрицу $\begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} \in \mathcal{S}_+^{n+1}$

нелинейное *вложение Веронезе* устроено таким образом, что

- квадратичная функция цены на $\mathbb{R}^n$ представляется линейной функцией на $\mathcal{S}^{n+1}$
- квадратичные ограничения в $\mathbb{R}^n$ представляются линейными в $\mathcal{S}^{n+1}$

Неточность релаксации

из-за наличия ограничений переход к выпуклой оболочке многообразия Веронезе приводит к релаксации



линейное сечение выпуклой оболочки больше, чем выпуклая оболочка линейного сечения

Формулировка в пространстве мер

проблема: конус неотрицательных мер бесконечно-мерный

подход: полиномиальные задачи позволяют ограничиться конечно-мерной проекцией на набор моментов мер

проблема: конус моментов неотрицательных мер в общем случае численно недоступен

подход: полу-определенные релаксации, выражающие необходимые условия

в ряде важных случаев релаксации точны

Конуса положительных полиномов

пусть $P_{d,n}$ — конус неотрицательных однородных полиномов степени d от n вещественных переменных

если d четное, то $P_{d,n}$ — регулярный выпуклый конус

внутренние элементы $P_{d,n}$ — полиномы, положительные на единичной сфере

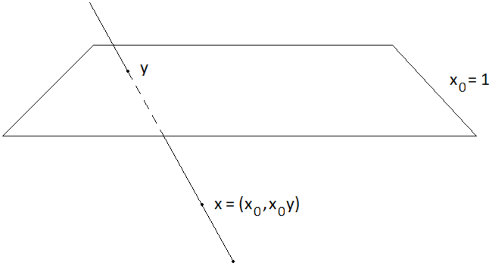
можно рассмотреть и неоднородную версию: конус полиномов степени не больше d, неотрицательных на $\mathbb{R}^n$

такой полином может быть положительным, но все равно лежать на границе конуса

пример: $p(x) \equiv 1$ равен нулю во всех бесконечно удаленных точках

Эквивалентность однородного и неоднородного конуса

конус неоднородных неотрицательных полиномов степени не больше d от n переменных изоморфен $P(d, n+1)$



неоднородному полиному $q(y)$ ставим в соответствие однородный полином $p(x) = x_0^d q(y)$

Пример: коположительный конус

симметрическая матрица $A \in \mathcal{S}^n$ называется *коположительной*, если $x^T A x \geq 0$ для всех $x \in \mathbb{R}^n_+$, или эквивалентно

$$p_A(x) = \sum_{i,j=1}^{n} A_{ij} x_i^2 x_j^2 \in P_{4,n}$$

коположительный конус $\mathcal{COP}^n$ представляем через линейное сечение конуса неотрицательных полиномов $P_{4,n}$

проверить $A \in \mathcal{COP}^n$ является ко-NP-полной задачей [Murty, Kabadi 1987]

в общем случае проверить $p \in P_{2d,n}$ является сложной проблемой

Конус сумм квадратов

Определение

Конусом сумм квадратов $\Sigma_{2d,n}$ *называется множество однородных полиномов* $p(x)$ *четной степени* $2d$ *от* n *вещественных переменных, которые представляются в виде конечной суммы* $p(x) = \sum_k q_k^2(x)$, *где* $q_k(x)$ *— однородные полиномы степени* d.

имеем $\Sigma_{2d,n} \subset P_{2d,n}$

$\Sigma_{2d,n}$ является внутренней аппроксимацией конуса $P_{2d,n}$

Полу-определенное представление $\Sigma_{2d,n}$

пусть $\mathbf{x}$ — вектор мономов $x^\alpha := \prod_{k=1}^n x_k^{\alpha_k}$ степени d

$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ — мульти-индекс, удовлетворяющий $|\alpha| := \sum_{k=1}^n \alpha_k = d$

$\mathbf{x}$ имеет $N = \binom{n+d-1}{d}$ элементов

Лемма

Однородный полином p степени $2d$ от n переменных $x_1, \dots, x_n$ является элементом конуса $\Sigma_{2d,n}$ тогда и только тогда, когда существует неотрицательно определенная матрица $A \in \mathcal{S}_+^N$ такая, что $p(x) = \mathbf{x}^T A \mathbf{x}$.

Построение представления в виде суммы квадратов

пусть $A \in \mathcal{S}_+^N$ и $p(x) = \mathbf{x}^T A \mathbf{x}$

тогда $A = B^T B$, $B \in \mathbb{R}^{k \times N}$
пусть b_j, $j = 1, \ldots, k$ — строки фактора B

имеем

$$p(x) = \mathbf{x}^T B^T B \mathbf{x} = \sum_{j=1}^{k} \langle b_j, \mathbf{x} \rangle^2 = \sum_{j=1}^{k} q_j^2(x)$$

и p представлено в виде суммы квадратов k полиномов $q_j(x) = \langle b_j, \mathbf{x} \rangle$ степени d

отсюда следует включение $p \in \Sigma_{2d,n}$

Построение $A \in \mathcal{S}_+^N$

пусть $p \in \Sigma_{2d,n}$

тогда $p(x) = \sum_{j=1}^k q_j^2(x)$
$q_j(x) = \langle b_j, \mathbf{x} \rangle, \ b_j \in \mathbb{R}^N$

составим матрицу $B \in \mathbb{R}^{k \times N}$ из строк b_j
имеем

$$p(x) = \sum_{j=1}^{k} \langle b_j, \mathbf{x} \rangle^2 = \mathbf{x}^T B^T B \mathbf{x} = \mathbf{x}^T A \mathbf{x}$$

матрица A не обязательно единственная

Полу-определенное представление $\Sigma_{2d,n}$

представим конус сумм квадратов $\Sigma_{2d,n}$ как проекцию матричного конуса $\mathcal{S}_+^N$

определим проекцию Π, сопоставляющую матрице $A \in \mathcal{S}^N$ полином $p(x)$ с коэффициентами

$$c_\alpha = \sum_{\beta,\gamma:\beta+\gamma=\alpha} A_{\beta\gamma}$$

т.е. $\Pi : A \mapsto \mathbf{x}^T A \mathbf{x}$

тогда $\Sigma_{2d,n} = \Pi[\mathcal{S}_+^N]$

Релаксации полу-определенного представления

порядок N матрицы A очень быстро увеличивается с ростом n и d

полу-определенное условие $A \succeq 0$ можно далее аппроксимировать достаточными условиями [Ahmadi, Majumbar 2014]

- DDSOS (diagonally dominant sums of squares): $A_{ii} \geq \sum_{j \neq i} |A_{ij}|$ для всех $i = 1, \ldots, N$
- SDDSOS (scaled diagonally dominant sums of squares): A из двойственного конуса к конусу матриц B, для которых $(\frac{B_{ii}+B_{jj}}{2}, \frac{B_{ii}-B_{jj}}{2}, B_{ij}) \in L_3$ для всех $i, j = 1, \ldots, N$

первая релаксация линейная, вторая конично-квадратичная

Точность релаксации

Теорема (Hilbert, 1888)

Пусть $n, d \in \mathbb{N}_+$, *d четное. Равенство* $\Sigma_{d,n} = P_{d,n}$ *справедливо тогда и только тогда, когда либо* $\min(d, n) \leq 2$, *либо* $(d, n) = (4, 3)$.

в остальных случаях конус $P_{d,n}$ не является полу-определенно представимым [Scheiderer 2018]

Случай $n = 2$

вместо конуса $P_{2d,2}$ рассмотрим изоморфный конус неоднородных полиномов от одной переменной x

нам нужно представить неотрицательный полином как сумму квадратов
пусть $p(x) = c\prod_{k=1}^{m}(x - x_k)$, где m — степень, а x_k — корни полинома
из неотрицательности следует $c = (c')^2 > 0$, m — четное

если x_k — вещественный корень, то он должен иметь четную кратность
если $x_k = a_k + ib_k$ — комплексный корень, то к нему имеется сопряженный корень $x_{k'} = a_k - ib_k$
имеем $(x - x_k)(x - x_{k'}) = (x - a_k)^2 + b_k^2$

таким образом $p(x)$ — сумма квадратов

Политоп Ньютона

Определение

Пусть $p(x) = \sum_\alpha c_\alpha x^\alpha$ *— полином от* n *переменных. Политопом Ньютона полинома* p *называют выпуклую оболочку множества мульти-индексов* $\{\alpha \in \mathbb{N}^n \mid c_\alpha \neq 0\}$.

Лемма

Пусть $p(x) = \sum_\alpha c_\alpha x^\alpha$ *— полином от* n *переменных, представимый в виде суммы квадратов* $\sum_j q_j(x)^2$, *и пусть* P *— его политоп Ньютона. Тогда*

- *экстремальные точки политопа* P *имеют только четные компоненты,*
- *коэффициенты* c_α, *соответствующие экстремальным точкам* α *политопа* P, *строго положительны,*
- *мульти-индексы* β, *соответствующие ненулевым коэффициентам полиномов* q_j, *являются элементами политопа* $\frac{1}{2}P$.

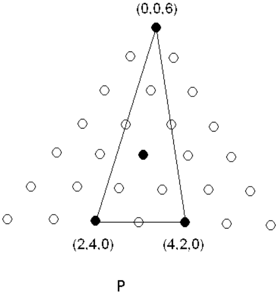
Изменение базы

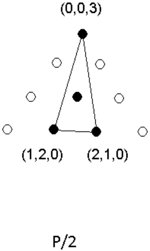
полином Моцкина $p_M(x,y,z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2$
неотрицательный вследствие неравенства между алгебраическим и геометрическим средним
не представим в виде суммы квадратов кубических полиномов

однако, p_M является суммой квадратов других функций:

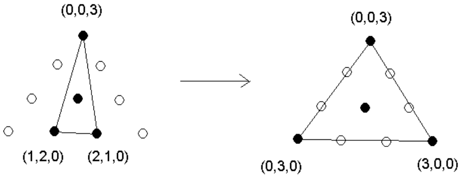
$$p_M(x,y,z) = (u^2+v^2+z^2)\cdot \begin{pmatrix} u^2 \\ v^2 \\ z^2 \end{pmatrix}^T \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} u^2 \\ v^2 \\ z^2 \end{pmatrix},$$

где $u = x^{2/3}y^{1/3}$, $v = x^{1/3}y^{2/3}$
J — неотрицательно определенная матрица ранга 2





политоп Ньютона для полинома
$p_M(x, y, z) = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2$



переход к полиному $p(u, v, z) = u^6 + v^6 + z^6 - 3u^2v^2z^2$ расширяет базу мономов, из которых составляются факторы q_j в сумме квадратов другое представление получается, если разложим произведение $(x^2 + y^2 + z^2) \cdot p_M(x, y, z)$ в сумму квадратов полиномов 4-й степени

тогда и произведение $(x^2 + y^2 + z^2)^2 \cdot p_M(x, y, z)$ является суммой квадратов полиномов, и p_M представляемо в виде суммы квадратов от рациональных функций

17-ая проблема Гильберта: Любой неотрицательный полином может быть представлен в виде суммы квадратов *рациональных* функций.

Утверждение было доказано Э. Артином в 20-х гг.

Усиление релаксации

условие $p \in \Sigma(2d, n)$ является достаточным для неотрицательности p

аппроксимацию можно усилить, потребовав

$$\left(\sum_{i=1}^{n} x_i^2\right)^r \cdot p \in \Sigma(2(d+r), n)$$

эти условия также записываются в виде линейных матричных неравенств на коэффициенты полинома p, но с ростом r их сложность быстро растет

Матричные полиномы

матричный полином $A : \mathbb{R} \to \mathcal{S}^n$ от одной переменной можно рассматривать как обычный полином $p(t, x) = x^T A(t) x$ от $n + 1$ переменных

неотрицательность $p(t, x)$ равносильна условию $A(t) \succeq 0$

Теорема

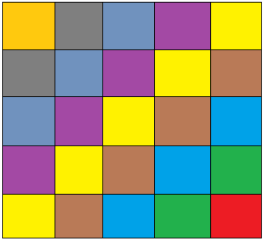
Пусть $A_0, \ldots, A_{2m} \in \mathcal{S}^n$ — симметрические матрицы такие, что матричный полином $A(t) = \sum_{j=0}^{2m} A_j t^j$ степени $2m$ положительный, т.е., $x^T A(t) x \geq 0$ для всех $t \in \mathbb{R}$ и $x \in \mathbb{R}^n$. Тогда найдутся матрицы $B_0, \ldots, B_m \in \mathbb{R}^{n \times k}$ такие, что $A(t) = B(t) B(t)^T$, где $B(t) = \sum_{j=0}^{m} B_j t^j$ — матричный полином степени m.

эквивалентно, $p(t, x) = (B(t)^T x)^T (B(t)^T x)$ является обычной суммой квадратов от полиномов меньшей степени

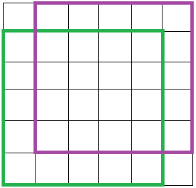
Структура блочно-ханкелевых матриц

Лемма

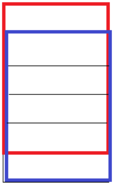
Пусть $H \in \mathcal{S}_+^{n(m+1)}$ *— неотрицательно определенная блочно-ханкелевая матрица с блоками размера* n. *Тогда* H *представляется в виде суммы неотрицательно определенных блочно-ханкелевых матриц ранга 1.*

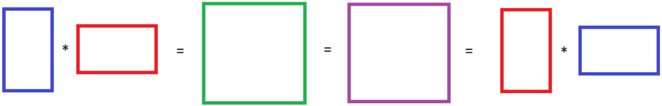


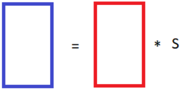
$H = FF^T =$



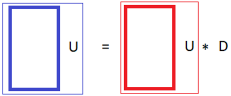
$F =$

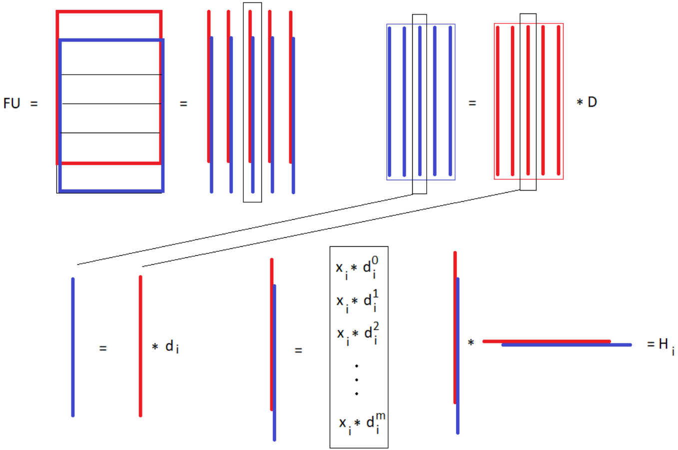






$S = S^T = UDU^T$




FU =
=
=
* D
=
* d_i
=
$x_i * d_i^0$
$x_i * d_i^1$
$x_i * d_i^2$
$x_i * d_i^m$
*
= H_i

Двойственность между полиномами и матрицами

Лемма

Пусть $A(t) = \sum_{j=0}^{2m} A_j t^j$ *— положительный матричный полином, а* $H \in \mathcal{S}_+^{n(m+1)}$ *— блочно-ханкелевая матрица с блоками* $H_0, \dots, H_{2m} \in \mathcal{S}^n$*. Тогда* $\sum_{i=0}^{2m} \langle A_i, H_i \rangle \geq 0$*. В частности, существует такая матрица* $\mathbf{A} \in \mathcal{S}_+^{n(m+1)}$*, состоящая из* $n \times n$ *блоков* A_{ij}*,* $i, j = 0, \dots, m$*, что* $A_k = \sum_{i+j=k} A_{ij}$*.*

если rk $H > 1$, то H можно разложить в сумму матриц ранга 1
если rk $H = 1$, то $H_i = \lambda^i x x^T$,
$\sum_{i=0}^{2m} \langle A_i, H_i \rangle = x^T (\sum_{i=0}^{2m} A_i \lambda^i) x \geq 0$

второе утверждение — следствие двойственности

Пример: коположительный конус

$A \in \mathcal{COP}^n$ эквивалентно

$$p_A(x) = \sum_{i,j=1}^{n} A_{ij} x_i^2 x_j^2 \in P_{4,n}$$

аппроксимация суммой квадратов дает

$$\mathcal{K}_0 = \{A \in \mathcal{S}^n \mid p_A \in \Sigma_{4,n}\}$$

Явный вид $\mathcal{K}_0$

приведем построение для $n = 3$

сформируем вектор $\mathbf{x} = (x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3, x_2x_3)^T$

$A \in \mathcal{K}_0$ эквивалентно существованию $\mathbf{A} \in \mathcal{S}_+^6$ такой, что

$$p_A(x) = \mathbf{x}^T\mathbf{A}\mathbf{x}$$

	x_1^2	x_2^2	x_3^2	x_1x_2	x_1x_3	x_2x_3
x_1^2	x_1^4	$x_1^2x_2^2$	$x_1^2x_3^2$	$x_1^3x_2$	$x_1^3x_3$	$x_1^2x_2x_3$
x_2^2	$x_1^2x_2^2$	x_2^4	$x_2^2x_3^2$	$x_1x_2^3$	$x_1x_2^2x_3$	$x_2^3x_3$
x_3^2	$x_1^2x_3^2$	$x_2^2x_3^3$	x_3^4	$x_1x_2x_3^2$	$x_1x_3^3$	$x_2x_3^3$
x_1x_2	$x_1^3x_2$	$x_1x_2^3$	$x_1x_2x_3^2$	$x_1^2x_2^2$	$x_1^2x_2x_3$	$x_1x_2^2x_3$
x_1x_3	$x_1^3x_3$	$x_1x_2^2x_3$	$x_1x_3^3$	$x_1^2x_2x_3$	$x_1^2x_3^2$	$x_1x_2x_3^2$
x_2x_3	$x_1^2x_2x_3$	$x_2^3x_3$	$x_2x_3^3$	$x_1x_2^2x_3$	$x_1x_2x_3^2$	$x_2^2x_3^2$

коэффициенты **A**, стоящие вне блочно-диагональной под-матрицы $\mathrm{diag}(B, c_{12}, c_{13}, c_{23})$, можно положить равными нулю

Явный вид $\mathcal{K}_0$

условие $A \in \mathcal{K}_0$ эквивалентно существованию матрицы $B \in \mathcal{S}^3_+$ и скаляров $c_{12}, c_{13}, c_{23} \geq 0$, что

$$\sum_{i,j=1}^{n} A_{ij} x_i^2 x_j^2 = \sum_{i,j=1}^{n} B_{ij} x_i^2 x_j^2 + \sum_{i<j} c_{ij} x_i^2 x_j^2$$

сравнивая коэффициенты, получаем

$$\operatorname{diag} A = \operatorname{diag} B, A_{ij} = B_{ij} + c_{ij} \forall\ i < j$$

отсюда получаем $\mathcal{K}_0 = \mathcal{S}^3_+ + \mathcal{N}^3$, где $\mathcal{N}^3$ — конус поэлементно неотрицательных симметрических 3×3 матриц с нулевой диагональю

условие на диагональ можно опустить

Явный вид $\mathcal{K}_0$

в общем случае

$$\mathcal{K}_0 = \mathcal{S}_+^n + \mathcal{N}^n,$$

где $\mathcal{N}^n$ — конус поэлементно неотрицательных симметрических $n \times n$ матриц

Теорема (Diananda 1962)

Равенство $\mathcal{COP}^n = \mathcal{S}_+^n + \mathcal{N}^n$ имеет место тогда и только тогда, когда $n \leq 4$.

Усиление релаксации

определим

$$\mathcal{K}_r = \left\{ A \in \mathcal{S}^n \mid \left(\sum_{j=1}^{n} x_j^2 \right)^r \cdot p_A(x) \in \Sigma_{4+2r,n} \right\}.$$

иерархия релаксаций Паррило $\mathcal{K}_0 \subset \mathcal{K}_1 \ldots$

Теорема (Parrilo 2000)

Пусть $A \in \operatorname{int} \mathcal{COP}^n$. *Тогда существует* $r \geq 0$ *такое, что* $A \in \mathcal{K}_{r'}$ *для всех* $r' \geq r$.

Класс задач

Определение

Множество $K \subset \mathbb{R}^n$ называется базовым полу-алгебраическим если оно задается

$$K = \{x \mid f_i(x) = 0,\ g_j(x) \leq 0\}$$

для некоторых полинимов $f_i, g_j : \mathbb{R}^n \to \mathbb{R}$.

рассмотрим проблему минимизации полинома на базовом полу-алгебраическом множестве K:

$$\min_{x \in K} f_0(x)$$

данные задачи — полиномы f_i, g_j

Коническая формулировка

введем конус $P_{d,K}$, состоящий из полиномов степени, не превосходящей d, которые неотрицательны на базовом полу-алгебраическом множестве K

$P_{d,K}$ — конечно-мерный, замкнутый и выпуклый

задача запишется в виде

$$\max \tau : \quad f_0(x) - \tau \in P_{d,K},$$

где d — не меньше степени f_0

это коническая программа в *пространстве полиномов*

в качестве неизвестных переменных выступают коэффициенты полинома и переменная τ

Релаксация задачи

конус $P_{d,K}$ не поддается эффективному описанию

введем конус $\Sigma_{d,K}$, состоящий из всех полиномов $p(x)$ степени, не превосходящей d, вида

$$p(x) = \sigma_0(x) + \sum_i p_i(x) f_i(x) - \sum_j \sigma_j(x) g_j(x),$$

где $p_i(x)$ — произвольные полиномы, а $\sigma_0(x), \sigma_j(x)$ — суммы квадратов полиномов

конус $\Sigma_{d,K}$ является внутренней аппроксимацией $P_{d,K}$

аппроксимируем исходную полиномиальную задачу оптимизации конической программой

$$\max \tau : \quad f_0(x) - \tau \in \Sigma_{d,K}.$$

Полу-определенная представимость

коэффициенты

$$p(x) = \sigma_0(x) + \sum_i p_i(x) f_i(x) - \sum_j \sigma_j(x) g_j(x)$$

линейны по σ_0, p_i, σ_j, а конуса, по которым пробегают σ_0, σ_j, полу-определенно представимы

поэтому условие $p \in \Sigma_{d,K}$ полу-определенно представимо, оно эквивалентно конечному набору полу-определенных и линейных ограничений

в итоге релаксация — полу-определенная коническая задача

Усиление релаксации

релаксация может быть усилена, если вместо конуса $\Sigma_{d,K}$ использовать конус $\Sigma^r_{d,K}$ полиномов p, удовлетворяющих условию

$$p(x) \cdot \left(\sum_{i=1}^{n} x_i^2 \right)^r = \sigma_0(x) + \sum_i p_i(x) f_i(x) - \sum_j \sigma_j(x) g_j(x)$$

это условие слабее, но все еще достаточно для неотрицательности полинома p на K

конус $\Sigma^r_{d,K}$ также полу-определенно представим, и $\Sigma^r_{d,K} \subset P_{d,K}$

релаксации усиливаются и становятся сложнее с ростом d и r

Пример

рассмотрим полиномиальную проблему оптимизации

$$\min x + y : \quad x \geq 0,\ x^2 + y^2 = 1$$

на базовом полу-алгебраическом множестве

$$K = \{(x, y) \,|\, x \geq 0,\ x^2 + y^2 = 1\}$$

пусть $d = 3$
аппроксимируем конус $P_{3,K}$ конусом $\Sigma_{3,K}$ полиномов, представимых в виде

$$p(x, y) = \sigma_0(x, y) + l(x, y)(x^2 + y^2 - 1) + \sigma_1(x, y)x,$$

где σ_0, σ_1 — суммы квадратов линейных полиномов, а l — линейный полином

Представление $\Sigma_{3,K}$

введем вектор мономов $\mathbf{x} = (x, y, 1)^T$

тогда $\Sigma_{3,K}$ состоит из полиномов вида

$$
\begin{aligned}
p(x,y) &= \mathbf{x}^T A^0 \mathbf{x} + \mathbf{l}^T \mathbf{x} \cdot (x^2 + y^2 - 1) + (\mathbf{x}^T A^1 \mathbf{x}) \cdot x \\
&= (A^1_{11} + l_x)x^3 + (2A^1_{12} + l_y)x^2 y + (A^0_{11} + 2A^1_{13} + l_1)x^2 \\
&\quad + (A^1_{22} + l_x)xy^2 + (2A^0_{12} + 2A^1_{23})xy + (2A^0_{13} + A^1_{33} - l_x)x \\
&\quad + l_y y^3 + (A^0_{22} + l_1)y^2 + (2A^0_{23} - l_y)y + A^0_{33} - l_1,
\end{aligned}
$$

где $\mathbf{l} = (l_x, l_y, l_1)^T \in \mathbb{R}^3$ и $A^0, A^1 \in \mathcal{S}^3_+$

приравниваем p к $x + y - \tau$, получаем

$$
\max_{A^0, A^1 \in \mathcal{S}^3_+} \tau :
$$

$$
\begin{gathered}
A^1_{11} + l_x = 2A^1_{12} + l_y = A^0_{11} + 2A^1_{13} + l_1 = 0, \\
A^1_{22} + l_x = 2A^0_{12} + 2A^1_{23} = l_y = A^0_{22} + l_1 = 0, \\
2A^0_{13} + A^1_{33} - l_x = 2A^0_{23} - l_y = 1, \;\; A^0_{33} - l_1 = -\tau
\end{gathered}
$$

элиминируем часть переменных, получаем

$$\max -(A^0_{33} + A^0_{11} + 2A^1_{13}) :$$

$$\begin{pmatrix} A^0_{11} & A^0_{12} & A^0_{13} \\ A^0_{12} & A^0_{11} + 2A^1_{13} & \frac{1}{2} \\ A^0_{13} & \frac{1}{2} & A^0_{33} \end{pmatrix} \succeq 0,$$

$$\begin{pmatrix} A^1_{11} & 0 & A^1_{13} \\ 0 & A^1_{11} & -A^0_{12} \\ A^1_{13} & -A^0_{12} & 1 - A^1_{11} - 2A^0_{13} \end{pmatrix} \succeq 0$$

решение приводит к оптимальному значению -1, которое является оптимальным также для исходной задачи

Моменты меры

Определение

Пусть $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ *— мульти-индекс. Моментом* m_α *меры* μ *будем называть значение интеграла*

$$m_\alpha(\mu) = \int_{\mathbb{R}^n} x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \cdots \cdot x_n^{\alpha_n} \mu(x)\, dx = \int_{\mathbb{R}^n} x^\alpha \mu(x)\, dx.$$

m_α является линейным функционалом на пространстве мер

его область определения не обязательно совпадает со всем пространством, поскольку интеграл может расходиться

для δ-функции $\mu(x) = \delta(x - \hat{x})$ все моменты существуют и равны $m_\alpha(\mu) = \hat{x}^\alpha$

Конус моментов

фиксируем d и рассматриваем только моменты m_α, для которых степень $|\alpha| = \sum_{i=1}^{n} \alpha_i$ не превосходит d

их конечное число $N = \binom{n+d}{d}$, и они образуют N-мерный *вектор моментов* $m(\mu) = (m_\alpha(\mu))_{\alpha:|\alpha|\le d}$

Конус моментов $M_d \subset \mathbb{R}^N$ определяется как множество всех векторов, которые представляются как вектор моментов некоторой неотрицательной меры μ.

для подмножества $K \subset \mathbb{R}^n$ определим конус $M_{d,K} \subset \mathbb{R}^N$, состоящий из векторов моментов неотрицательных мер μ с носителем в K

конуса моментов являются *проекциями* конуса неотрицательных мер

Релаксация конуса моментов

найдем необходимые условия на вектор мометнов неотрицательной меры μ

пусть $\mathbf{x} = (1, x_1, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_n^{[d/2]})^T$ — вектор мономов x^α, для которых степень $|\alpha|$ не превосходит целую часть от $\frac{d}{2}$

тогда все элементы одноранговой матрицы $\mathbf{x}\mathbf{x}^T$ являются мономами со степенью, не превосходящей d

элементы матрицы

$$\int_{\mathbb{R}^n} \mathbf{x}\mathbf{x}^T \mu(x)\, dx \succeq 0$$

равны неким элементам вектора моментов $m(\mu)$

отсюда получаем полу-определенное ограничение на вектор моментов

Случай с носителем в K

пусть $K = \{x \in \mathbb{R}^n \mid f_i(x) = 0,\ g_j(x) \leq 0\}$ — базовое полу-алгебраическое множество, а μ — неотрицательная мера с носителем в K

для любого полинома p интеграл

$$\int_K p(x) f_i(x) \mu(x)\, dx = 0$$

выражается через линейную комбинацию моментов

получаем линейное ограничение типа равенства на вектор моментов $m(\mu)$

если полином $p(x)$ пробегает все мономы x^β степени $|\beta| \leq d - d_i$, где $d_i = \deg f_i$, то мы получим максимальный линейно независимый набор ограничений на моменты до степени d

пусть $d_j = \deg g_j$
образуем вектор $\mathbf{x}'$ всех мономов степени, не превосходящей целую часть от $\frac{d-d_j}{2}$

каждый элемент интеграла

$$-\int_K \mathbf{x}'(\mathbf{x}')^T g_j(x)\mu(x)\,dx \succeq 0$$

представляется в виде линейной комбинации моментов степени $\leq d$

получаем полу-определенное ограничение на $m(\mu) \in M_{d,K}$

Моментная релаксация

рассмотрим задачу

$$min_{x \in K} f_0(x),$$

где $f_0 = \sum_\alpha c_\alpha x^\alpha$, deg $f_0 \leq d$, K — базовое полу-алгебраическое множество

эквивалентно получаем коническую задачу

$$\min_{\mu \geq 0 : \operatorname{supp} \mu \subset K} \int_{\mathbb{R}^n} f_0(x)\mu(x)\, dx = \sum_\alpha c_\alpha m_\alpha(\mu) :$$

$$\int_{\mathbb{R}^n} \mu(x)\, dx = m_0(\mu) = 1$$

или

$$\min_{m \in M_{d,K}} \sum_\alpha c_\alpha m_\alpha : \quad m_0 = 1$$

заменив включение $m \in M_{d,K}$ полу-определенными и линейными необходимыми условиями получаем релаксацию

Пример

рассмотрим снова проблему

$$\min x + y : \quad x \geq 0,\ x^2 + y^2 - 1 = 0$$

при $d = 3$ вектор моментов будет 10-мерным
релаксация запишется в виде

$$\min m_{10} + m_{01} : \quad \begin{pmatrix} m_{00} & m_{10} & m_{01} \\ m_{10} & m_{20} & m_{11} \\ m_{01} & m_{11} & m_{02} \end{pmatrix} \succeq 0,$$

$$m_{20} + m_{02} - m_{00} = m_{30} + m_{12} - m_{10} = m_{21} + m_{03} - m_{01} = 0,$$

$$\begin{pmatrix} m_{10} & m_{20} & m_{11} \\ m_{20} & m_{30} & m_{21} \\ m_{11} & m_{21} & m_{12} \end{pmatrix} \succeq 0,\ m_{00} = 1$$

решение также дает оптимальное значение -1

Параметризация

в тригонометрическом полиноме независимой переменной выступает либо угол $\phi \in [-\pi, \pi]$, либо комплексная величина $z = e^{i\phi}$, принимающая значения на единичном круге $\mathbb{T}$

в первом случае полином — 2π-периодическая функция на $\mathbb{R}$, во втором — конечный ряд Лорана по переменной z

тригонометрические полиномы сводятся к обычным заменой независимой переменной

$$\cos\phi = \frac{z+\bar{z}}{2} = \frac{1-t^2}{1+t^2}, \ \sin\phi = \frac{z-\bar{z}}{2i} = \frac{2t}{1+t^2}, \ \ t \in \mathbb{R} \cup \{\infty\}$$

в приложениях тригонометрические полиномы часто матрично-значные

Положительность

полином степени d имеет общий вид

$$p(\phi) = \sum_{k=-d}^{d} A_k e^{ik\phi} = \sum_{k=-d}^{d} A_k z^k,$$

где $A_{-k} = A_k^*$ для всех $k = 0, \ldots, d$,
A_k — матрицы размера $n \times n$, A^* — эрмитово сопряженная (комплексно сопряженная транспонированная) к A

полином принимает значения в пространстве $\mathcal{H}^n$ эрмитовых матриц размера n ($\mathcal{H}^1 = \mathbb{R}$)

p неотрицательный, если $p(\phi) \succeq 0$ для всех ϕ
множество неотрицательных полиномов степени $\leq d$ является выпуклым конусом

Полу-определенное представление

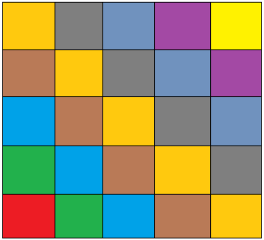
Теорема

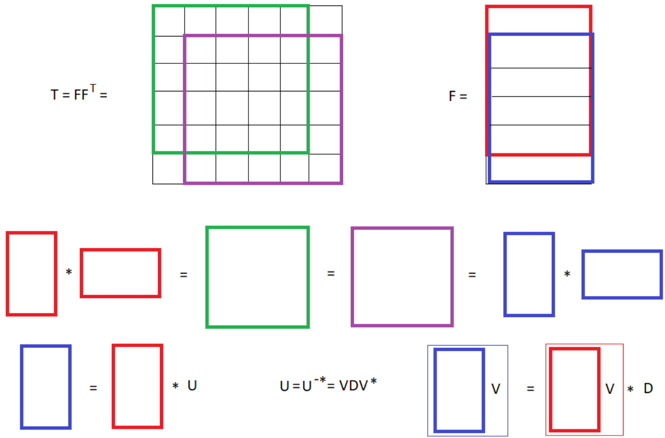
Пусть $p(\phi) = \sum_{k=-m}^{m} A_k e^{ik\phi}$ *— тригонометрический полином. Полином* p *является неотрицательным тогда и только тогда, когда существует эрмитовая неотрицательно определенная матрица* $\mathbf{A} \in \mathcal{H}_+^{n(m+1)}$ *такая, что* $A_k = \sum_{i-j=k} A_{ij}$*, где* A_{ij}*,* $i, j = 0, \ldots, m$ *— блоки матрицы* $\mathbf{A}$ *размера* n*.*

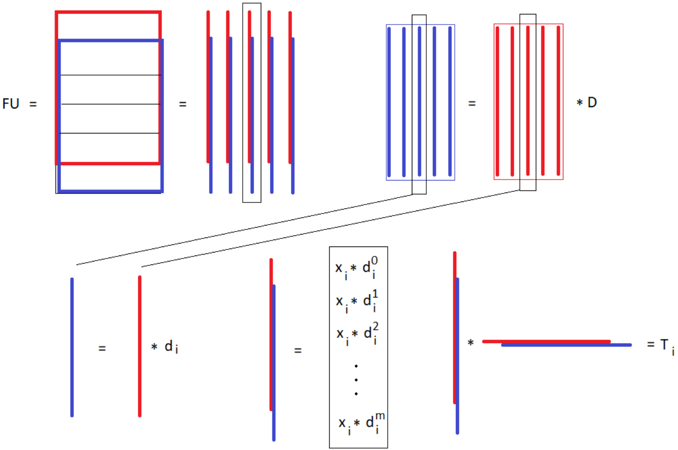
Блочно-теплицевые матрицы

Лемма

Пусть $T \in \mathcal{H}_+^{n(m+1)}$ *— неотрицательно определенная блочно-теплицевая эрмитова матрица с блоками размера* n*. Тогда* T *представляется в виде суммы неотрицательно определенных блочно-теплицевых матриц ранга 1.*




T = FF^T =
F =
*
=
=
=
*
=
* U
U = U^{-*} = VDV^*
V
=
V
* D


FU =
=
=
* D
=
* d_i
=
$x_i * d_i^0$
$x_i * d_i^1$
$x_i * d_i^2$
$x_i * d_i^m$
*
= T_i

Двойственность между полиномами и матрицами

Лемма

Пусть $p(\phi) = \sum_{j=-m}^{m} A_j e^{ij\phi}$ *— положительный тригонометрический матричный полином, а* $T \in \mathcal{H}_+^{n(m+1)}$ *— блочно-теплицевая матрица с блоками* $T_{-m}, \dots, T_m \in \mathcal{H}^n$*. Тогда* $\sum_{j=-m}^{m} \langle A_j, T_j \rangle \geq 0$*. В частности, существует такая матрица* $\mathbf{A} \in \mathcal{H}_+^{n(m+1)}$*, состоящая из* $n \times n$ *блоков* A_{ij}*,* $i, j = 0, \dots, m$*, что* $A_k = \sum_{i-j=k} A_{ij}$*.*

если $\mathrm{rk}\ T > 1$, то T можно разложить в сумму матриц ранга 1
если $\mathrm{rk}\ T = 1$, то $T_j = z^j x x^T$,
$\sum_{j=-m}^{m} \langle A_j, T_j \rangle = x^T (\sum_{j=-m}^{m} A_j z^j) x \geq 0$

второе утверждение — следствие двойственности

Представление в виде сумм квадратов

пусть p — положительный тригонометрический полином, а $\mathbf{A} \succeq 0$ — соответствующая матрица

факторизуя $\mathbf{A} = \mathbf{B}\mathbf{B}^*$ и разбивая фактор $\mathbf{B} \in \mathbb{C}^{n(m+1)\times l}$ на блоки $B_0, \ldots, B_m \in \mathbb{C}^{n\times l}$, получаем

$$p(\phi) = \left(\sum_{k=0}^{m} B_k e^{ik\phi}\right)\left(\sum_{k=0}^{m} B_k e^{ik\phi}\right)^*$$

т.е. p — матричный эрмитовый квадрат