Геометрия самосогласованных барьеров IV. Симметрические конуса

Roland Hildebrand

LJK, Université Grenoble Alpes / CNRS

Летняя школа "Управление, Информация, Оптимизация"
Вороново, 21 июня 2019 г.

1. Параллелизм кубической формы
   - Симметрические конуса
   - Авто-шкалированные барьеры
   - Условия параллелизма
2. Точки шкалировки как ортогональные проекции
   - Ближайшие точки
   - Reach
   - Взаимное расположение объектов

Мотивация

самыми успешными классами конических программ являются программы над *симметрическими* конусами

благодаря наличию барьеров с дополнительными свойствами: авто-шкалировка

мы покажем, что это свойство эквивалентно условию $\hat{\nabla} C = 0$, где $\hat{\nabla}$ – связность Леви-Чивита метрики, C – центро-аффинная кубическая форма

$C \sim F'''$:

- $C \equiv 0$ – $F \sim$ "полином второй степени"
- $\hat{\nabla} C \equiv 0$ – $F \sim$ "полином третьей степени"

наиболее простой класс барьеров после гиперболического

Симметрические конуса

Определение

*Самодвойственный однородный конус называется **симметрическим**.*

- однородный: группа линейных автоморфизмов конуса действует транзитивно на внутренности
- самодвойственный: линейно изоморфен двойственному

Классификация

[Винберг, 1960; Koecher, 1962] любой симметрический конус является произведением конечного числа следующих неприводимых симметрических конусов:

- конус Лоренца
  $L_n = \left\{(x_0, \dots, x_{n-1}) \,|\, x_0 \geq \sqrt{x_1^2 + \dots + x_{n-1}^2}\right\}$, $n \neq 2$
- матричный конус $S_+(n)$, $H_+(n)$, $Q_+(n)$ вещественных, комплексных или кватернионных эрмитовых положительно определенных матриц, $n \geq 3$
- конус Альберта $O_+(3)$ октонионных эрмитовых положительно определенных матриц размера 3×3

Жордановы алгебры

симметрические конуса имеют и алгебраическую структуру

Определение

Коммутативная алгебра J, удовлетворяющая условию

$$(x \bullet x) \bullet (x \bullet y) = x \bullet ((x \bullet x) \bullet y)$$

для всех $x, y \in J$ называется ***жордановой алгеброй****.*
Жордановая алгебра называется ***евклидовой*** *если из $\sum_{k=1}^{n} x_k \bullet x_k = 0$ следует $x_k = 0$ для всех $k = 1, \dots, n$.*

определены степени $x^k = x \bullet \cdots \bullet x$

Связь конусов с алгебрами

Теорема

Пусть $K \subset \mathbb{R}^n$ – симметрический конус. Тогда на $\mathbb{R}^n$ существует структура евклидовой жордановой алгебры J такой, что

$$K = \{x \bullet x \,|\, x \in J\}.$$

Для любой евклидовой жордановой алгебры J заданный выше конус квадратов является симметрическим конусом.

для внутренности конуса также есть представление

$$K^o = \exp J = \left\{ \exp x = \sum_{k=0}^{+\infty} \frac{x^k}{k!} \,|\, x \in J \right\}$$

Примеры

конусу Лоренца L_n соответствует жорданова алгебра с произведением

$$\begin{pmatrix} x_0 \\ \tilde{x} \end{pmatrix} \bullet \begin{pmatrix} y_0 \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} x_0 y_0 + \langle \tilde{x}, \tilde{y} \rangle \\ x_0 \tilde{y} + y_0 \tilde{x} \end{pmatrix}$$

она обладает единичным элементом $(1, 0, \dots, 0)^T$

матричным конусам над $\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$ соответствуют алгебры с произведением

$$A \bullet B = \frac{AB + BA}{2}$$

они обладают единичным элементом I

Детерминант

каждая жорданова алгебра имеет детерминант, полином $d : J \to \mathbb{R}$

на внутренности симметрических конусов детерминант положителен, и равен нулю на границе

для неприводимых конусов:

- L_n: $d(x) = x_0^2 - x_1^2 - \cdots - x_{n-1}^2$
- $\mathcal{S}_+^n$: $d(A) = \det A$

для произведений:

$$d(x_1, \dots, x_m) = \prod_{i=1}^{m} d_i(x_i)$$

Авто-шкалированные барьеры

Определение

Пусть $K \subset \mathbb{R}^n$ – регулярный выпуклый конус, K^* двойственный к нему, F – самосогласованный барьер на K с параметром ν, F_* – двойственный к нему барьер на K^*. Тогда F называется *авто-шкалированным* *если для всех $x, w \in K^o$ справедливо*

$$s = F''(w)x \in \text{int}\, K^*, \qquad F_*(s) = F(x) - 2F(w) - \nu.$$

Конус K, допускающий авто-шкалированный барьер, называется *авто-шкалированным.*

Nesterov, Todd: для любой пары $(x, s) \in (K \times K^*)^o$ существует единственная точка шкалировки $w \in K^o$ такая, что

$$F''(w)x = s$$

Классификация

Hauser, Güler, Lim, Schmieta 1998 – 2002:

- авто-шкалированный конус $\Leftrightarrow$ симметрический конус
- авто-шкалированные барьеры на произведениях конусов являются взвешенными суммами авто-шкалированных барьеров на неприводимых факторах
- авто-шкалированные барьеры на неприводимых конусах являются логарифмами детерминантов

Обозначения

$\frac{\partial F}{\partial x^\alpha} = F_{,\alpha}$, $\frac{\partial^2 F}{\partial x^\alpha \partial x^\beta} = F_{,\alpha\beta}$ и т.д.

$F^{,\alpha\beta}$ = обратная гессиана $F_{,\alpha\beta}$

правило Эйнштейна суммирования по повторяющимся индексам

$$F^{,\alpha\beta} F_{,\beta\gamma} := \sum_{\beta=1}^{n} F^{,\alpha\beta} F_{,\beta\gamma} = \delta^\alpha_\gamma$$

Связь с жордановыми алгебрами

Теорема (Güler, Schmieta, Koecher)

Пусть F – авто-шкалированный барьер. Тогда величины $K^a_{bc} = F^{,ad}F_{,bcd}$ являются коэффициентами жордановой алгебры,

$$(u \bullet v)^a = K^a_{bc} u^b v^c.$$

из коэффициентов $K^a_{bc} = F^{,ad}F_{,bcd}$ восстановить F'', F''' однозначно невозможно

разные авто-шкалированные барьеры могут приводить к одной и той же алгебре

Метризованные жордановы алгебры

к алгебре нужно добавить метрику, чтобы можно было восстановить барьер

$g_{ab}K^b_{cd} = F_{,ab}F^{,be}F_{,cde} = F_{,acd}$ симметрично

Определение

Метризованной жордановой алгеброй называется пара (τ, J)*, где* J *– жорданова алгебра,* τ *– билинейная симметрическая форма на* J*, удовлетворяющие условию*

$$\tau(a \bullet b, c) = \tau(a, b \bullet c)$$

для всех $a, b, c \in J$*.*

Барьеры и метризованные алгебры

Теорема

Пусть K – симметрический конус и J соответствующая евклидова жорданова алгебра. Тогда любой авто-шкалированный барьер на K может быть записан в виде

$$F(x) = \tau(e, \log x),$$

где τ положительно определенная форма, метризующая J. С другой стороны, для любого такого τ функция $F(x)$ пропорциональна авто-шкалированному барьеру на K.

здесь e – единичный элемент алгебры, а $\log$ – обратная от экспоненты

Параллелизм F'''

Когда $K^a_{bc} = F^{,ad}F_{,bcd}$ являются коэффициентами жордановой алгебры?

Ответ: тогда и только тогда, когда $\hat{\nabla}F''' = 0$, где $\hat{\nabla}$ – связность Леви-Чивита метрики F''

Необходимость условия

уравнение $\hat{\nabla} F''' = 0$ записывается в виде

$$F_{,\alpha\beta\gamma\delta} - \Gamma^{\rho}_{\alpha\delta} F_{,\rho\beta\gamma} - \Gamma^{\rho}_{\beta\delta} F_{,\alpha\rho\gamma} - \Gamma^{\rho}_{\gamma\delta} F_{,\alpha\beta\rho} = 0$$

это эквивалентно квазилинейному уравнению в частных производных 4-го порядка

$$F_{,\alpha\beta\gamma\delta} = \frac{1}{2} F^{,\rho\sigma} \left(F_{,\alpha\beta\rho} F_{,\gamma\delta\sigma} + F_{,\alpha\gamma\rho} F_{,\beta\delta\sigma} + F_{,\alpha\delta\rho} F_{,\beta\gamma\sigma} \right)$$

$\Gamma^{a}_{bc} = \frac{1}{2} F^{,ad} F_{,bcd}$ – символы Христоффеля связности $\hat{\nabla}$

Условие интегрируемости

продифференцируем уравнение по x^η и заменим возникающие производные 4-го порядка на функции от 3-х, используя само уравнение

$$
\begin{aligned}
F_{,\alpha\beta\gamma\delta\eta} = \frac{1}{4} F^{,\rho\sigma} F^{,\mu\nu} \big(& F_{,\beta\eta\nu} F_{,\alpha\rho\mu} F_{,\gamma\delta\sigma} + F_{,\alpha\eta\mu} F_{,\rho\beta\nu} F_{,\gamma\delta\sigma} \\
+ \quad & F_{,\gamma\eta\nu} F_{,\alpha\rho\mu} F_{,\beta\delta\sigma} + F_{,\alpha\eta\mu} F_{,\rho\gamma\nu} F_{,\beta\delta\sigma} + F_{,\beta\eta\nu} F_{,\gamma\rho\mu} F_{,\alpha\delta\sigma} \\
+ \quad & F_{,\gamma\eta\mu} F_{,\rho\beta\nu} F_{,\alpha\delta\sigma} + F_{,\beta\eta\nu} F_{,\delta\rho\mu} F_{,\alpha\gamma\sigma} + F_{,\delta\eta\mu} F_{,\rho\beta\nu} F_{,\alpha\gamma\sigma} \\
+ \quad & F_{,\delta\eta\nu} F_{,\alpha\rho\mu} F_{,\beta\gamma\sigma} + F_{,\alpha\eta\mu} F_{,\rho\delta\nu} F_{,\beta\gamma\sigma} + F_{,\delta\eta\nu} F_{,\gamma\rho\mu} F_{,\alpha\beta\sigma} \\
+ \quad & F_{,\gamma\eta\mu} F_{,\rho\delta\nu} F_{,\alpha\beta\sigma} \big)
\end{aligned}
$$

антикоммутируя δ, η, получаем условие интегрируемости

Условие интегрируемости

$$F^{,\rho\sigma}F^{,\mu\nu}(F_{,\beta\eta\nu}F_{,\delta\rho\mu}F_{,\alpha\gamma\sigma}+F_{,\alpha\eta\mu}F_{,\rho\delta\nu}F_{,\beta\gamma\sigma}+F_{,\gamma\eta\mu}F_{,\rho\delta\nu}F_{,\alpha\beta\sigma}$$
$$-F_{,\beta\delta\nu}F_{,\eta\rho\mu}F_{,\alpha\gamma\sigma}-F_{,\alpha\delta\mu}F_{,\rho\eta\nu}F_{,\beta\gamma\sigma}-F_{,\gamma\delta\mu}F_{,\rho\eta\nu}F_{,\alpha\beta\sigma})=0.$$

умножаем на $(F'')^{-1}$ и получаем

$$\begin{aligned}&\Gamma^{\eta}_{\alpha\mu}\Gamma^{\mu}_{\delta\rho}\Gamma^{\rho}_{\beta\gamma}+\Gamma^{\eta}_{\beta\mu}\Gamma^{\mu}_{\delta\rho}\Gamma^{\rho}_{\alpha\gamma}+\Gamma^{\eta}_{\gamma\mu}\Gamma^{\mu}_{\delta\rho}\Gamma^{\rho}_{\alpha\beta}\\ -\quad&\Gamma^{\mu}_{\alpha\delta}\Gamma^{\eta}_{\rho\mu}\Gamma^{\rho}_{\beta\gamma}-\Gamma^{\mu}_{\beta\delta}\Gamma^{\eta}_{\rho\mu}\Gamma^{\rho}_{\alpha\gamma}-\Gamma^{\mu}_{\gamma\delta}\Gamma^{\eta}_{\rho\mu}\Gamma^{\rho}_{\alpha\beta}=0\end{aligned}$$

тождество справедливо если

$$\Gamma^{\eta}_{\alpha\mu}\Gamma^{\mu}_{\delta\rho}\Gamma^{\rho}_{\beta\gamma}u^{\alpha}u^{\beta}u^{\gamma}v^{\delta}=\Gamma^{\mu}_{\alpha\delta}\Gamma^{\eta}_{\rho\mu}\Gamma^{\rho}_{\beta\gamma}u^{\alpha}u^{\beta}u^{\gamma}v^{\delta}$$

для всех касательных векторов u, v

Переход к алгебре

в произвольной точке x определим следующее умножение на касательном пространстве:

$$\Gamma(u,v)^\alpha = (u \bullet v)^\alpha = \frac{1}{2} F^{,\alpha\delta} F_{,\delta\beta\gamma} u^\beta v^\gamma = \Gamma^\alpha_{\beta\gamma} u^\beta v^\gamma$$

условие интегрируемости записывается в виде

$$\Gamma(\Gamma(\Gamma(u,u),v),u) = \Gamma(\Gamma(u,v),\Gamma(u,u))$$

умножение определяет коммутативную алгебру, удовлетворяющую тождеству Жордана

$$(u^2 \bullet v) \bullet u = (u \bullet v) \bullet u^2$$

т.е. жорданову алгебру

Гессиан метризует алгебру

гессиан F'' удовлетворяет условию

$$\begin{aligned} F''(u \bullet v, w) &= F_{,\beta\gamma}\Gamma^{\beta}_{\delta\rho}u^{\delta}v^{\rho}w^{\gamma} = \frac{1}{2}F_{,\beta\gamma}F_{,\delta\rho\sigma}F^{,\sigma\beta}u^{\delta}v^{\rho}w^{\gamma} \\ &= \frac{1}{2}F_{,\delta\rho\gamma}u^{\delta}v^{\rho}w^{\gamma} = \frac{1}{2}F_{,\beta\delta}u^{\delta}F_{,\rho\gamma\sigma}F^{,\sigma\beta}v^{\rho}w^{\gamma} \\ &= F_{,\delta\beta}u^{\delta}\Gamma^{\beta}_{\rho\gamma}v^{\rho}w^{\gamma} = F''(u, v \bullet w). \end{aligned}$$

поэтому F'' метризует алгебру

Достаточность условия

Теорема

Пусть (τ, J) *– метризованная жорданова алгебра. Тогда существует окрестность* $U \subset J$ *нуля такая, что аналитическая функция*

$$F(x) = \sum_{k=2}^{\infty} \frac{(-1)^k}{k} \tau(x, x^{k-1})$$

является решением уравнения $\hat\nabla F''' = 0$ *на* U.

Параллелизм градиента

производная F' является $\hat{\nabla}$-параллельной если

$$F_{,\alpha\beta} - \Gamma^{\gamma}_{\alpha\beta} F_{,\gamma} = F_{,\alpha\beta} - \frac{1}{2} F^{,\gamma\delta} F_{,\alpha\beta\delta} F_{,\gamma} = 0$$

это можно записать в виде

$$2F''(\cdot,\cdot) = F'''(\cdot,\cdot,(F'')^{-1}F')$$

обозначим красный вектор через $-e$

Интегрирование уравнения

положим $e^\gamma = -F_{,\delta}F^{,\gamma\delta}$

тогда получим

$$2F_{,\alpha\beta} = -F_{,\alpha\beta\delta}e^\delta$$

дифференцируем e^γ

$$\begin{aligned} e^\gamma_{,\alpha} &= -F_{,\alpha\delta}F^{,\gamma\delta} + F^{,\gamma\rho}F_{,\rho\sigma\alpha}F^{,\sigma\delta}F_{,\delta} \\ &= -F_{,\alpha\delta}F^{,\gamma\delta} - F^{,\gamma\rho}F_{,\rho\sigma\alpha}e^\sigma = -\delta^\gamma_\alpha + 2F^{,\gamma\rho}F_{,\rho\alpha} = \delta^\gamma_\alpha \end{aligned}$$

таким образом, $e(x) = x + const$

Интегрирование уравнения

выберем начало координат так, что $e(x) = x$

$$\begin{aligned} & F_{,\delta} + F_{,\gamma\delta}x^\gamma = (F_{,\gamma}x^\gamma)_{,\delta} = 0 \\ \Rightarrow\quad & F_{,\gamma}x^\gamma = const = \nu \\ \Rightarrow\quad & F(\alpha x) = \nu \log \alpha + F(x), \quad \alpha > 0 \end{aligned}$$

значит F логарифмично однородно

рассуждение обратимо если $\det F'' \neq 0$

и степень однородности ν, и положение начала координат возникают как константы интегрирования

Параллелизм кубической формы

Теорема

Пусть $M \subset \mathbb{R}^n$ – центро-аффинная гиперповерхность с положительно определенной метрикой. Следующие утверждения эквивалентны:
а) M является частью поверхности уровня авто-шкалированного барьера
б) центро-аффинная кубическая форма параллельна по отношению к связности Леви-Чивита центро-аффинной метрики на M.

Итоги

найдена геометрическая интерпретация авто-шкалированности барьера

- локальное условие в форме уравнения в частных производных 4-го порядка
- $\hat{\nabla} F''' = 0$: "кубический полином"
- наиболее простой класс после гиперболических барьеров
- алгоритмы?

Мотивация

в прямо-двойственных методах итерация генерирует пару (x, s)

$s \neq -F'(x)$ и $(x, s) \notin M$

в какой точке строить аппроксимирующую квадратичную функцию / аппроксимирующий конус?

Нестеров, Тодд: в точке шкалировки

геометрическая интуиция: в ближайшей точке

Точка шкалировки как ближайшая точка

точка шкалировки, заданная условием

$$F''(w)x = s$$

является ближайшей к (x, s) точкой на M в псевдо-римановой метрике произведения:

$$\max_{w \in K^o} \langle x - w, s + F'(w) \rangle$$

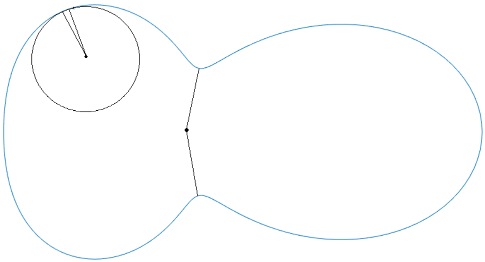
производная по w приводит к условию

$$-(s + F'(w)) + F''(w)(x - w) = 0$$

красные члены сокращаются, поскольку $F''(w)w = -F'(w)$

похожая проблема рассмотрена в [Nesterov, Todd 1997]

Ближайшие точки



препятствия для единственности ближайшей точки:

- глобальные: точки, далекие на подмногообразии, близки в объемлющем пространстве
- локальные: кривизна многообразия

Reach

Определение (Federer 1959)

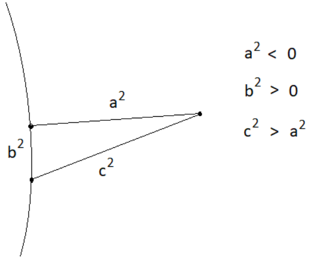
Пусть $A \subset E$ – подмножество Евклидова пространства.
Близкой к A точкой назовем точку $x \in E$ такую, что существует единственная точка $a \in A$, удовлетворяющая $||x - a|| = d(x, A)$.
Reach точки $a \in A$ – это максимум по $r \geq 0$ таких, что открытый шар $B_r^o(a)$ вокруг a состоит из близких точек.
The Reach подмножества A есть инфимум по $a \in A$ от reach(a).

Свойства reach

- A имеет бесконечный reach тогда и только тогда, когда A замкнуто и выпукло
- гладкие компактные связные подмногообразия имеют положительный reach
- $\text{reach}(a)$ – непрерывная функция на A
- для гладких многообразий A обратная от reach ограничена снизу кривизной A
- понятия можно обобщить на подмножества римановых многообразий

Reach в псевдо-римановом пространстве $\mathcal{M}$

- расстояние до множества должно быть положительным
- в ближайшей точке должен достигаться экстремум



нормальное и касательное пространства должны быть знако-определенными

Reach в псевдо-римановых пространстве $\mathcal{M}$

Определение

Пусть $M \subset \mathcal{M}$ – положительно определенное подмногообразие максимальной размерности в псевдо-римановом пространстве $\mathcal{M}$.
Близкой к M точкой назовем точку $x \in \mathcal{M}$ такую, что существует единственная точка $z \in M$, максимизирующая $d(x, z')$ по $z' \in M$.
Reach точки $a \in M$ – это максимум по $r \geq 0$ таких, что открытый шар $B_r^o(a)$ вокруг a в нормальном подмногообразии к M в точке z состоит из близких точек.
The Reach подмногообразия M есть инфимум по $a \in M$ от $reach(a)$.

Положительность reach

Теорема

Пусть $K \subset \mathbb{R}^n$ – регулярный выпуклый конус, а F – самосогласованный барьер на K с параметром ν. Соответствующее лагранжевое подмногообразие в $\mathbb{R}^n \times \mathbb{R}_n$ имеет reach $\nu^{-1/2}$. Соответствующее лагранжевое подмногообразие в $\mathbb{R}P^{n-1} \times \mathbb{R}P_{n-1}$ имеет reach $\arccos\sqrt{\frac{\nu-1}{\nu}}$.

в частности, в окрестности соответствующей толщины вокруг подмногообразия M ближайшие точки существуют и единственны

Объекты

рассмотрим прямую и двойственную коническую программы над $K \subset \mathbb{R}^{n+1}$, $K^* \subset \mathbb{R}_{n+1}$, соответственно

пусть $P_A \subset \mathbb{R}^{n+1}$, $D_A \subset \mathbb{R}_{n+1}$ – аффинные подпространства, задаваемые линейными ограничениями,
P_L, D_L – их линейные оболочки

имеем $\dim P_A + \dim D_A = n + 1$, $\dim P_L + \dim D_L = n + 3$

пусть (x, s) – текущая прямо-двойственная итерация, а $w \in M$ – соответствующая ближайшая точка

пусть $M_L \subset \mathcal{M}$ – вполне геодезическая аппроксимация M

Нормальные координаты

перейдем в систему координат, где M_L соответствует конусу Лоренца

$$L_n = \{x \mid x_0 \geq \sqrt{x_1^2 = \ldots x_{n-1}^2}\}$$

а точке шкалировки соответствует центр конуса $e_0 = (1, 0, \ldots, 0)$

мы все еще можем производить вращения вокруг центральной оси конуса

Линейные ограничения

пространства P_L, D_L генерируются столбцами матриц

$$\begin{pmatrix} \cos\varphi & 0 & \sin\varphi\cos\xi & \mathbf{0} \\ 0 & 1 & 0 & \mathbf{0} \\ \sin\varphi & 0 & -\cos\varphi\cos\xi & \mathbf{0} \\ 0 & 0 & -\sin\xi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & I \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \begin{pmatrix} \cos\varphi & 0 & \sin\varphi\sin\xi & \mathbf{0} \\ 0 & 1 & 0 & \mathbf{0} \\ \sin\varphi & 0 & -\cos\varphi\sin\xi & \mathbf{0} \\ 0 & 0 & \cos\xi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & I \end{pmatrix}$$

$\varphi < \frac{\pi}{4}$

взаимное расположение задается двумя параметрами

Текущая итерация

текущая прямо-двойственная пара задается

$$\begin{pmatrix} \cos\varphi \\ \beta \\ \sin\varphi \\ 0 \end{pmatrix}, \begin{pmatrix} \cos\varphi \\ \beta \\ \sin\varphi \\ 0 \end{pmatrix}$$

это дает третий параметр $\beta \in (-\sqrt{\cos 2\varphi}, \sqrt{\cos 2\varphi})$

расстояние до точки шкалировки равно $\frac{\cos^2\varphi}{1+\beta^2}$

Спасибо за внимание