



几何问题

Geometric Problems · Chapter 8

梁经纬 Jingwei Liang

Institute of Natural Sciences
School of Mathematical Sciences

Lecture 08: 几何问题

2026 年秋季学期

饮水思源 · 爱国荣校

本讲内容

投影、距离与分离 §§8.1–8.2
点到集合、集合间距离、support function 与分离证书



Gram 几何与椭球 §§8.3–8.4
距离/角度可实现性、极值体积椭球与仿射不变性



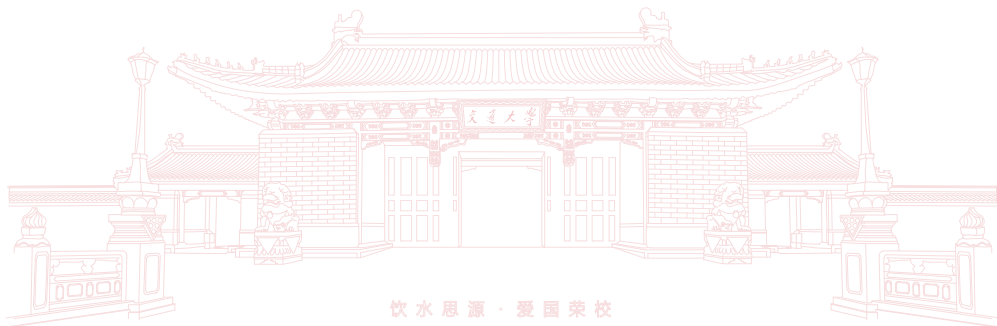
中心与分类 §§8.5–8.6
Chebyshev/analytic center、线性与非线性判别



布局与平面规划 §§8.7–8.8
facility location、path constraints、relative positioning 与 GP

01

前序知识与建模主线



饮水思源 · 爱国荣校

前序知识

Lectures 01A–01C, 02

凸集、support function、conic models
与拟合



Lectures 04–06

对偶、KKT、log-det barrier 与 interior-
point methods



Lecture 07

协方差、信息矩阵、置信椭球与实验设计

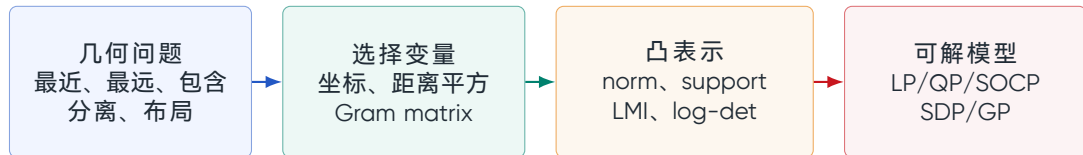
本章的统一变量

几何对象本身往往不直接作为变量，而改用

点 x , Gram matrix G ,
ellipsoid parameters (A, b) .

距离、角度、体积与位置关系由这些变量表达
成凸约束。

从几何叙述到凸模型



判断标准

固定相对位置、把距离平方写成 affine function、或把“对所有点”改写成 support function，常能消除几何表象中的非线性；离散的先后/上下选择仍保留组合性。

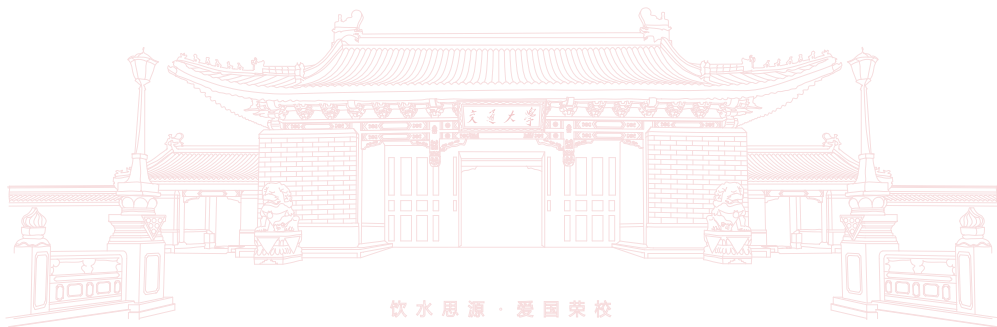
本章常见凸结构

几何对象	数学表示	典型模型	
最近点/集合间距离	$\inf \ x - y\ $	QP, SOCP	
角度与距离可实现性	$G = A^\top A \succeq 0$	SDP	
椭球体积	$\log \det A^{-1}$ 或 $\log \det B$	log-det program	同一几
最深点/中心	ball containment、barrier	LP, SDP	
分类边界	affine in parameters	LP, QP, SDP	
位置与平面规划	graph edges、relative order	LP, SOCP, GP	

何目标可能有不同参数化；凸性取决于变量与表示，而不只取决于图形外观。

02

投影与分离



距离与投影

闭集 $C \subseteq \mathbb{R}^n$ 上, 点 x_0 到 C 的距离为

$$\text{dist}(x_0, C) = \inf_{x \in C} \|x_0 - x\|.$$

任一达到最小值的 $z \in C$ 称为 projection:

$$P_C(x_0) \in \arg \min_{x \in C} \|x - x_0\|.$$

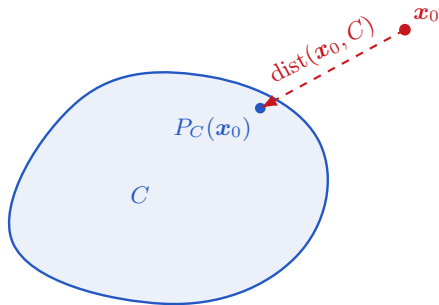
存在性

有限维空间中, C closed 即保证最短距离达到。

唯一性

C closed convex 且 norm strictly convex 时 projection 唯一; Euclidean norm 是主要情形。

投影的几何图像



Euclidean optimality

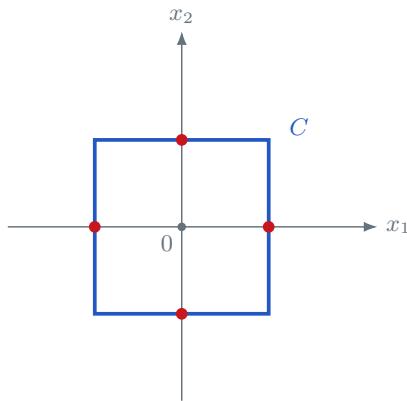
若 C closed convex, 则 $p = P_C(x_0)$ 当且仅当

$$(x_0 - p)^\top (x - p) \leq 0, \quad \forall x \in C.$$

残差是 C 在 p 处的 normal vector。

教材例 8.1: 单位方形边界

令 $C = \{x \in \mathbb{R}^2 \mid \|x\|_\infty = 1\}$, $x_0 = 0$ 。虽然 C 不是凸集, 但 projection 可以直接比较。



norm	projections
ℓ_1	$(\pm 1, 0), (0, \pm 1)$
ℓ_2	$(\pm 1, 0), (0, \pm 1)$
ℓ_∞	整个边界 C

三种距离均为 1; norm 的单位球决定“最

近”的方向结构。

教材例 8.2: 低秩投影

设

$$\mathcal{C} = \{X \in \mathbb{R}^{m \times n} \mid \text{rank } X \leq k\}, \quad X_0 = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top, \quad \sigma_1 \geq \cdots \geq \sigma_r > 0.$$

截断 SVD

$$P_{\mathcal{C}}(X_0) = \sum_{i=1}^{\min\{k, r\}} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$$

是 spectral norm 下的 projection; Frobenius norm 下也成立。



该集合非凸; 显式 projection 来自奇异值结构, 而非一般凸投影理论。

凸集上的投影问题

若

$$C = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, f_i(\mathbf{x}) \leq 0, i = 1, \dots, m\},$$

则 projection 由凸问题给出:

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & \|\mathbf{x} - \mathbf{x}_0\| \\ \text{subject to} & f_i(\mathbf{x}) \leq 0, \quad A\mathbf{x} = \mathbf{b}. \end{array}$$

Euclidean projection on a polyhedron:

$$\min_{A\mathbf{x} \leq \mathbf{b}} \frac{1}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2,$$

是 strongly convex QP, feasible 时解唯一。

KKT 给出

$$\mathbf{x}^* - \mathbf{x}_0 + A^\top \boldsymbol{\lambda}^* = 0,$$

因此 projection residual 是 active normals 的非负组合。

三个显式 Euclidean projections

Hyperplane

$$C = \{x \mid a^\top x = b\}:$$

$$P_C(x_0) = x_0 + \frac{b - a^\top x_0}{\|a\|_2^2} a.$$

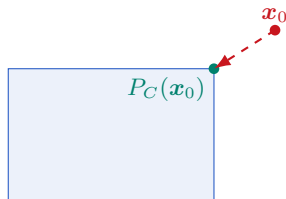
Halfspace

$C = \{x \mid a^\top x \leq b\}$: 若已 feasible, 保持 x_0 ; 否则使用上式投到边界。

Rectangle

$$C = \{x \mid l \leq x \leq u\}:$$

$$[P_C(x_0)]_j = \min\{u_j, \max\{l_j, x_{0j}\}\}.$$



Proper cone 上的投影

对 proper cone K , 令 $x_+ = P_K(x_0)$ 。KKT 给出 Moreau decomposition:

$$x_0 = x_+ - x_-, \quad x_+ \succeq_K 0, \quad x_- \succeq_{K^*} 0, \quad x_+^\top x_- = 0.$$

$$K = \mathbb{R}_+^n$$

$$[P_K(x_0)]_i = \max\{x_{0i}, 0\}.$$

负分量逐项截断为零。

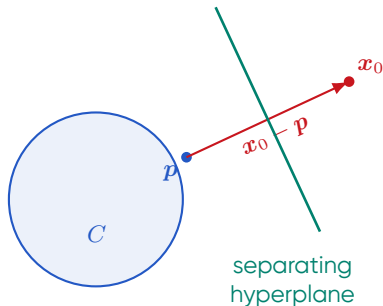
$$K = \mathbb{S}_+^n$$

若 $X_0 = \sum_i \lambda_i q_i q_i^\top$, 则

$$P_K(X_0) = \sum_i \max\{\lambda_i, 0\} q_i q_i^\top.$$

projection 同时给出 primal cone component 与 orthogonal dual-cone certificate。

投影导出分离超平面



令 $p = P_C(x_0)$, $x_0 \notin C$ 。中垂超平面

$$(p - x_0)^\top \left(x - \frac{x_0 + p}{2} \right) = 0$$

严格分离 x_0 与 C 。该直接构造依赖

Euclidean geometry; 一般 norm 可通过 duality 构造分离方向。

Boyd & Vandenberghe, *Convex Optimization*, §8.1.2, Figure 8.1

任意 norm 下的分离证书

将 $y = x_0 - x$ 引入 projection problem, dual 为

$$\begin{aligned} & \underset{\lambda, \mu, \nu}{\text{maximize}} && \mu^\top x_0 + \inf_x \left(\sum_i \lambda_i f_i(x) + \nu^\top (Ax - b) - \mu^\top x \right) \\ & \text{subject to} && \lambda \geq 0, \quad \|\mu\|_* \leq 1. \end{aligned}$$

若某个 dual feasible point 的目标值为正, 则对所有 $x \in C$,

$$\mu^\top x_0 > \mu^\top x;$$

μ 因而定义严格分离超平面。若 projection problem 满足 Slater, positive distance 与这种 certificate 等价。

Indicator 与 support function

定义

$$I_C(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in C, \\ +\infty, & \mathbf{x} \notin C, \end{cases} \quad S_C(\mathbf{z}) = \sup_{\mathbf{x} \in C} \mathbf{z}^\top \mathbf{x}.$$

投影可写成 $\underset{\mathbf{x}}{\text{minimize}} \|\mathbf{x} - \mathbf{x}_0\| + I_C(\mathbf{x})$; Fenchel dual 化简为

$$\begin{aligned} & \underset{\mathbf{z}}{\text{maximize}} && \mathbf{z}^\top \mathbf{x}_0 - S_C(\mathbf{z}) \\ & \text{subject to} && \|\mathbf{z}\|_* \leq 1. \end{aligned}$$

若最优值为正, $\mathbf{z}^\top \mathbf{x}_0 > S_C(\mathbf{z})$, 即 $\mathbf{x}_0$ 位于 C 的 support hyperplane 之外。

应用例 1: 概率向量校正

分类器给出的 raw score 为 $z = (0.8, -0.1, 0.6)^\top$; 需要投影到 probability simplex

$$\Delta_3 = \{p \geq 0 \mid \mathbf{1}^\top p = 1\}.$$

Euclidean projection 具有 threshold 形式

$$p_i^* = \max\{z_i - \tau, 0\}, \quad \sum_i \max\{z_i - \tau, 0\} = 1.$$

本例 $\tau = 0.2$, 所以

$$P_{\Delta_3}(z) = (0.6, 0, 0.4)^\top.$$

解释

归一化 $z/(\mathbf{1}^\top z)$ 会保留负概率; simplex projection 同时满足非负性、总和约束, 并使 Euclidean 改动最小。

自编示例: 将未校准分类得分投影到 probability simplex

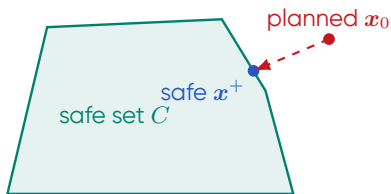
应用例 2: 机器人安全位置

移动机器人计划位置 x_0 , 但可行空间由走廊、禁区的线性近似给出:

$$C = \{x \in \mathbb{R}^2 \mid Ax \leq b\}.$$

控制器每次求解

$$x^+ = \arg \min_{Ax \leq b} \frac{1}{2} \|x - x_0\|_2^2.$$

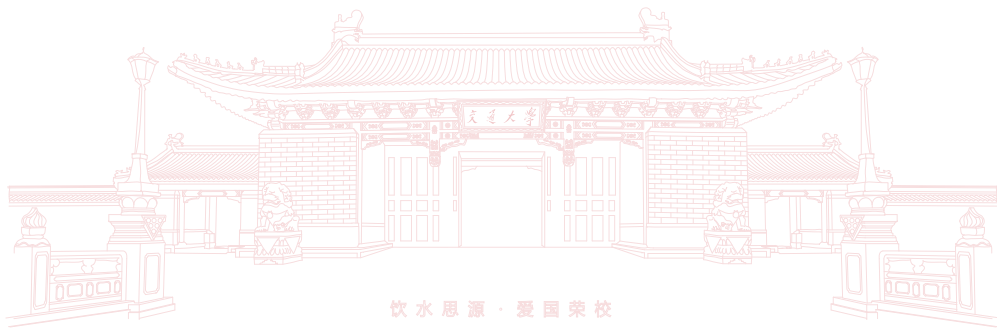


active constraints 的 multipliers 给出把计划点推回安全集的合力方向, 也量化每个边界对修正量的贡献。

自编示例: model-predictive safety filter via Euclidean projection

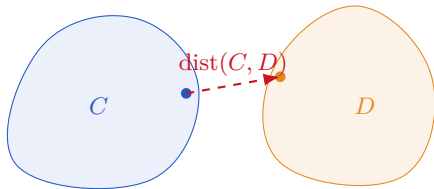
03

集合间距离



两个集合的距离

$$\text{dist}(C, D) = \inf\{\|x - y\| \mid x \in C, y \in D\} = \text{dist}(0, D - C).$$



零距离的含义

若 $C \cap D \neq \emptyset$, 距离为零。反向结论需要 infimum attained; 例如 sets closed 且 其中一个 bounded 时成立。

凸集间最近点

若

$$C = \{\mathbf{x} \mid f_i(\mathbf{x}) \leq 0\}, \quad D = \{\mathbf{y} \mid g_j(\mathbf{y}) \leq 0\},$$

则

$$\begin{aligned} & \underset{\mathbf{x}, \mathbf{y}}{\text{minimize}} && \|\mathbf{x} - \mathbf{y}\| \\ & \text{subject to} && f_i(\mathbf{x}) \leq 0, \quad g_j(\mathbf{y}) \leq 0 \end{aligned} \tag{8.3}$$

是 convex optimization problem。对 polyhedra $A_1\mathbf{x} \leq \mathbf{b}_1$ 、 $A_2\mathbf{y} \leq \mathbf{b}_2$ 和 Euclidean norm，平方目标得到 QP:

$$\min \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 \quad \text{s.t.} \quad A_1\mathbf{x} \leq \mathbf{b}_1, \quad A_2\mathbf{y} \leq \mathbf{b}_2.$$

两组 active face normals 在最优点沿连接线达到平衡。

距离问题的 dual separation

引入 $w = x - y$ 后, 对偶方向 z 满足 $\|z\|_* \leq 1$ 。一般形式为

$$\begin{aligned} & \underset{\lambda, \mu, z}{\text{maximize}} && \inf_x \left(\sum_i \lambda_i f_i(x) + z^\top x \right) + \inf_y \left(\sum_j \mu_j g_j(y) - z^\top y \right) \\ & \text{subject to} && \lambda \geq 0, \quad \mu \geq 0, \quad \|z\|_* \leq 1. \end{aligned}$$

正的 dual value 蕴含

$$z^\top x > z^\top y, \quad \forall x \in C, \forall y \in D,$$

故 z 定义 strict separating hyperplane。strong duality 时, positive distance 与分离 certificate 相互对应。

Polyhedra 的显式分离证书

对 $C = \{x \mid A_1 x \leq b_1\}$ 、 $D = \{y \mid A_2 y \leq b_2\}$, dual 为

$$\begin{aligned} & \underset{\lambda, \mu, z}{\text{maximize}} && -b_1^\top \lambda - b_2^\top \mu \\ & \text{subject to} && A_1^\top \lambda + z = 0, \quad A_2^\top \mu - z = 0, \\ & && \|z\|_* \leq 1, \quad \lambda \geq 0, \mu \geq 0. \end{aligned}$$

对任意 $x \in C$ 、 $y \in D$,

$$z^\top x - z^\top y \geq -b_1^\top \lambda - b_2^\top \mu.$$

右端若为正, 不仅证明不相交, 还给出 normalized separation margin 的下界。

Support functions 的紧凑表达

集合间距离的 dual 可写成

$$\begin{array}{ll} \underset{z}{\text{maximize}} & -S_C(-z) - S_D(z) \\ \text{subject to} & \|z\|_* \leq 1. \end{array}$$

若 dual objective 为正, 则

$$S_D(z) < -S_C(-z) \iff \sup_{y \in D} z^\top y < \inf_{x \in C} z^\top x.$$

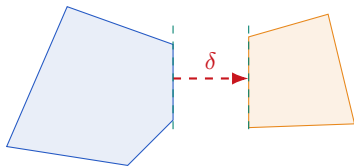
几何意义

$S_D(z)$ 是 D 沿 z 的最远投影, $-S_C(-z)$ 是 C 沿 z 的最近投影; 两者间隙正是该方向的 separation margin。

应用例 3: 机械臂碰撞裕量

把两个刚体在当前姿态附近近似为 polyhedra $C(q)$ 、 D ；最近点对给出 clearance

$$\delta(q) = \min_{\mathbf{x} \in C(q), \mathbf{y} \in D} \|\mathbf{x} - \mathbf{y}\|_2.$$

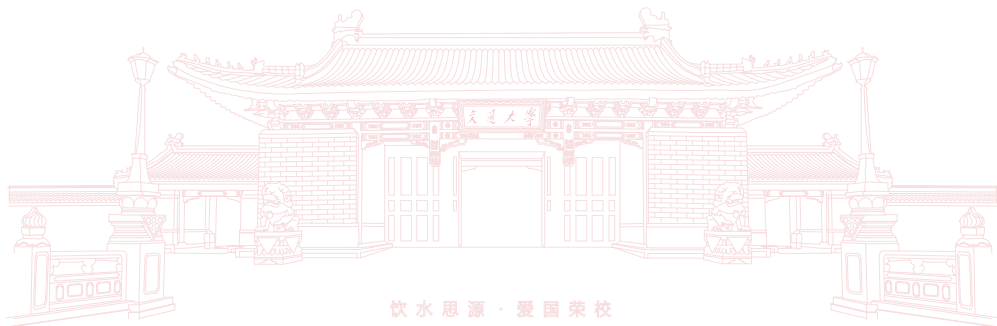


$\delta(q) \geq \delta_{\min}$ 是安全要求。最近点方向同时产生 supporting planes，可在线性化 motion planning 中形成 separating constraint。

自编示例: robot collision clearance and separating-plane certificate

04

距离、角度与 Gram 矩阵



Gram matrix 统一几何量

令 $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$, Gram matrix

$$G = A^\top A \succeq 0, \quad G_{ij} = \mathbf{a}_i^\top \mathbf{a}_j, \quad G_{ii} = \ell_i^2.$$

长度、距离、相关系数与角度全部由 G 表达:

$$d_{ij}^2 = \ell_i^2 + \ell_j^2 - 2G_{ij}, \quad \rho_{ij} = \frac{G_{ij}}{\ell_i \ell_j}, \quad \theta_{ij} = \arccos \rho_{ij}.$$

orthogonal transformation QA 不改变 G , 因此 G 保存旋转/反射不变量, 并消去坐标系自由度。

Realizability 是一个 LMI

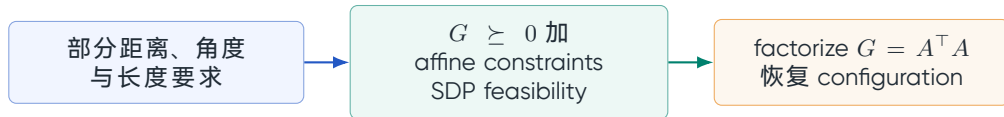
一组内积数据可由某些 vectors 实现，当且仅当

$$G \in \mathbb{S}_+^n.$$

若 $G \succeq 0$ ，可取

$$A = G^{1/2}, \quad A^\top A = G;$$

若 $G \succ 0$ 且 $LL^\top = G$ ，也可取 $A = L^\top$ 。



任意两种 realization 相差 orthogonal transformation (允许嵌入到更高维空间)。

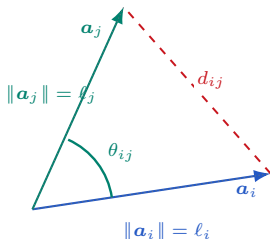
角度与距离约束变成 affine constraints

已知 ℓ_i 时, 角度区间 $\alpha \leq \theta_{ij} \leq \beta$ 等价于

$$\ell_i \ell_j \cos \alpha \geq G_{ij} \geq \ell_i \ell_j \cos \beta.$$

距离区间 $d_{\min} \leq d_{ij} \leq d_{\max}$ 等价于

$$d_{\min}^2 \leq \ell_i^2 + \ell_j^2 - 2G_{ij} \leq d_{\max}^2.$$



G_{ij} simultaneously controls the included angle and the opposite side through the cosine law.

优化含义

限制部分 distances/angles 后, 未知几何量的极值可由 SDP 求出; factorize 最优 G 即恢复 extreme configuration.

数值实例

若 $\ell_i = 2$ 、 $\ell_j = 3$ 、 $\theta_{ij} = 60^\circ$, 则 $G_{ij} = \ell_i \ell_j \cos \theta_{ij} = 3$, 并且 $d_{ij}^2 = 2^2 + 3^2 - 2G_{ij} = 7$ 。

条件数与 dual basis

Singular values

$\sigma_i(A)^2 = \lambda_i(G)$, 因此

$$\sigma_1(A)^2 = \lambda_{\max}(G), \quad \sigma_n(A)^2 = \lambda_{\min}(G).$$

最小化 condition number 可用 bisection 检查

$$tI \preceq G \preceq \kappa^2 tI.$$

Dual basis

若 $G \succ 0$, dual basis 的 Gram matrix 为 G^{-1} , 且

$$\|b_i\|_2^2 = e_i^\top G^{-1} e_i, \quad \sum_i \|b_i\|_2^2 = \text{tr } G^{-1}.$$

这些都是 G 的 convex functions。

体积也由 Gram matrix 决定

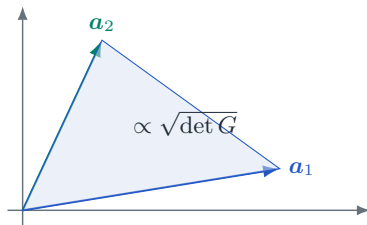
image ellipsoid $\{Au \mid \|u\|_2 \leq 1\}$ 的体积为

$$\gamma_n \sqrt{\det(A^\top A)} = \gamma_n \sqrt{\det G}.$$

simplex $\text{conv}\{0, a_1, \dots, a_n\}$ 的体积也与 $\sqrt{\det G}$ 成正比。因此

$$\underset{G}{\text{maximize}} \quad \log \det G$$

在 convex geometric constraints 下是 convex optimization。



Boyd & Vandenberghe, *Convex Optimization*, §8.3.1, ellipsoid and simplex volume

只考虑角度: correlation matrix

若 $G = \text{diag}(\boldsymbol{l})C \text{diag}(\boldsymbol{l})$, 其中 $C_{ij} = \cos \theta_{ij}$, 则

$$G \succeq 0 \iff C \succeq 0$$

对任意正长度向量 $\boldsymbol{l}$ 成立。因此 angles realizable 当且仅当

$$C \succeq 0, \quad C_{ii} = 1.$$

常见误区

逐对满足 $-1 \leq \rho_{ij} \leq 1$ 还不够; 所有 pairwise correlations 必须共同组成 PSD matrix。PSD 条件编码了不可见的全局一致性。

教材例 8.3: 相关系数边界

已知

$$0.6 \leq \rho_{12} \leq 0.9, \quad 0.8 \leq \rho_{13} \leq 0.9, \quad 0.5 \leq \rho_{24} \leq 0.7, \quad -0.8 \leq \rho_{34} \leq -0.4.$$

用两个 SDP 最小化/最大化 ρ_{14} , subject to $C \succeq 0$, $C_{ii} = 1$ 及上述 box constraints, 得到

$$-0.39 \leq \rho_{14} \leq 0.23.$$

教材给出的 extremal correlation matrices 为

$$\begin{bmatrix} 1 & .60 & .87 & -.39 \\ .60 & 1 & .33 & .50 \\ .87 & .33 & 1 & -.55 \\ -.39 & .50 & -.55 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & .71 & .80 & .23 \\ .71 & 1 & .31 & .59 \\ .80 & .31 & 1 & -.40 \\ .23 & .59 & -.40 & 1 \end{bmatrix}.$$

ρ_{14} 的有效区间明显窄于 $[-1, 1]$, 说明 PSD consistency 传播了局部相关信息。

Euclidean distance matrix

只关心 pairwise distances 时, 可平移 configuration 令 $\mathbf{A}\mathbf{1} = 0$ 。定义

$$D_{ij} = d_{ij}^2, \quad z_i = \|\mathbf{a}_i\|_2^2.$$

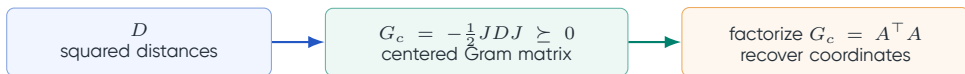
存在相应 configuration 当且仅当存在 $\mathbf{z} \succeq 0$ 使

$$G = \frac{1}{2}(\mathbf{z}\mathbf{1}^\top + \mathbf{1}\mathbf{z}^\top - D) \succeq 0. \quad (8.8)$$

令 centering matrix $J = I - \mathbf{1}\mathbf{1}^\top/n$, 则等价于

$$D = D^\top, \quad D_{ii} = 0, \quad D_{ij} \geq 0, \quad JDJ \preceq 0,$$

并且 centered Gram matrix 为 $G_c = -\frac{1}{2}JDJ \succeq 0$ 。



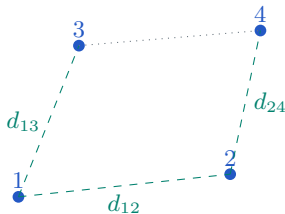
Boyd & Vandenberghe, *Convex Optimization*, §8.3.3, equation (8.8)

应用例 4: 传感器定位

四个 sensors 的部分 squared distances 已知, 未知 entries 组成矩阵 D 。可解

$$\begin{aligned} & \underset{D}{\text{minimize}} && \sum_{(i,j) \in \Omega} w_{ij} (D_{ij} - \hat{d}_{ij}^2)^2 \\ & \text{subject to} && D = D^\top, D_{ii} = 0, D_{ij} \geq 0, JDJ \preceq 0, \end{aligned}$$

再 factorize $G_c = -JDJ/2$ 恢复坐标。



自编示例: network localization via Euclidean distance matrix completion

应用例 5: 修复相关矩阵

样本缺失与独立 pairwise estimation 可能给出 indefinite matrix

$$\hat{C} = \begin{bmatrix} 1 & .9 & .9 \\ .9 & 1 & -.9 \\ .9 & -.9 & 1 \end{bmatrix}, \quad \lambda_{\min}(\hat{C}) = -0.8.$$

最近 correlation matrix 由 SDP 得到:

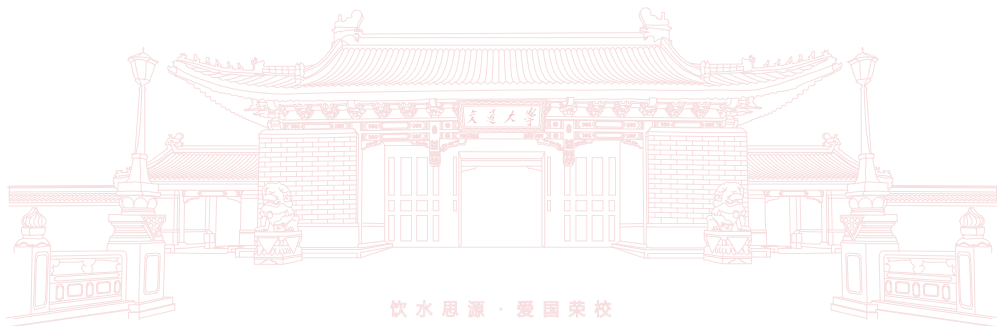
$$\begin{aligned} & \underset{C}{\text{minimize}} && \|C - \hat{C}\|_F^2 \\ & \text{subject to} && C \succeq 0, \quad C_{ii} = 1. \end{aligned}$$

该 projection 恢复 joint realizability; 在 portfolio risk、Gaussian model 与 kernel methods 中, PSD 不是数值修饰, 而是合法 covariance/Gram matrix 的必要条件。

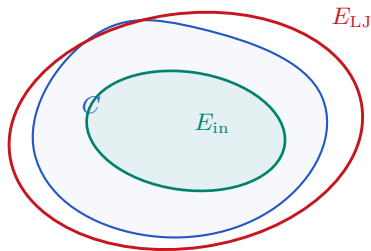
自编示例: nearest correlation matrix for inconsistent pairwise estimates

05

极值体积椭球



外接与内接椭球



Outer approximation

Löwner–John ellipsoid: 覆盖 C 的 minimum-volume ellipsoid。

Inner approximation

位于 C 内的 maximum-volume ellipsoid。

二者都对 affine coordinate changes 自然变换。

Boyd & Vandenberghe, *Convex Optimization*, §8.4

Löwner–John ellipsoid

用 inverse-image parameterization

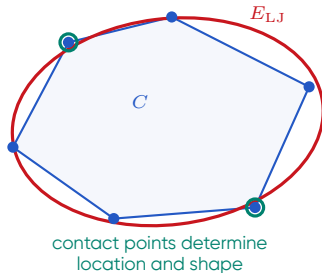
$$E = \{v \mid \|Av + b\|_2 \leq 1\}, \quad A \in \mathbb{S}_{++}^n.$$

其 volume 与 $\det A^{-1}$ 成正比。因此 minimum-volume cover 为

$$\begin{array}{ll} \underset{A \succ 0, b}{\text{minimize}} & \log \det A^{-1} \\ \text{subject to} & \sup_{v \in C} \|Av + b\|_2 \leq 1. \end{array}$$

(8.10)

目标与约束关于 (A, b) 均 convex; 难点是 containment constraint 中的 convex maximization, 只有特殊 C 可 tractably evaluate.

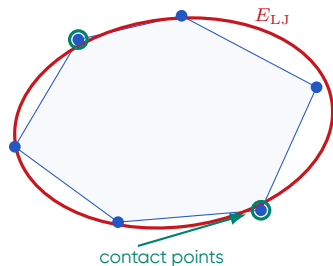


覆盖有限点集

若 $C = \{x_1, \dots, x_m\}$, 覆盖 C 等价于覆盖 $\text{conv } C$:

$$\begin{array}{ll} \underset{A \succ 0, \mathbf{b}}{\text{minimize}} & \log \det A^{-1} \\ \text{subject to} & \|A\mathbf{x}_i + \mathbf{b}\|_2 \leq 1, \\ & i = 1, \dots, m. \end{array} \quad (8.11)$$

active points are contact points; 它们决定最优椭球的中心、方向与轴长。



Boyd & Vandenberghe, *Convex Optimization*, §8.4.1, problem (8.11), Figure 8.3

覆盖椭球的并集

令

$$E_i = \{\mathbf{x} \mid \mathbf{x}^\top A_i \mathbf{x} + 2\mathbf{b}_i^\top \mathbf{x} + c_i \leq 0\}, \quad A_i \succ 0.$$

欲求覆盖 $\bigcup_i E_i$ 的最小体积椭球。对候选 $E = \{\mathbf{x} \mid \|A\mathbf{x} + \mathbf{b}\|_2 \leq 1\}$, S-procedure 把 $E_i \subseteq E$ 转成存在 $\tau_i \geq 0$ 的 matrix inequality。令 $Q = A^2$ 、 $\tilde{\mathbf{b}} = A\mathbf{b}$, 可写为

$$\begin{bmatrix} Q - \tau_i A_i & \tilde{\mathbf{b}} - \tau_i \mathbf{b}_i & 0 \\ (\tilde{\mathbf{b}} - \tau_i \mathbf{b}_i)^\top & -1 - \tau_i c_i & \tilde{\mathbf{b}}^\top \\ 0 & \tilde{\mathbf{b}} & -Q \end{bmatrix} \preceq 0,$$

并最小化 $-\frac{1}{2} \log \det Q$ 。这是 $(Q, \tilde{\mathbf{b}}, \tau)$ 的 convex problem。

外接椭球的近似保证

设 E_{LJ} 的中心为 $\mathbf{x}_c$ 。对 bounded convex $C \subset \mathbb{R}^n$,

$$\mathbf{x}_c + \frac{1}{n}(E_{\text{LJ}} - \mathbf{x}_c) \subseteq C \subseteq E_{\text{LJ}}.$$

若 C 关于 $\mathbf{x}_c$ centrally symmetric, 则改进为

$$\mathbf{x}_c + \frac{1}{\sqrt{n}}(E_{\text{LJ}} - \mathbf{x}_c) \subseteq C \subseteq E_{\text{LJ}}.$$

紧性

一般集合的 $1/n$ 在 simplex 上达到; symmetric case 的 $1/\sqrt{n}$ 在 cube 上达到。保证只依赖 dimension, 不依赖集合的细节。

用 quadratic norm 近似任意 norm

令 $C = \{x \mid \|x\| \leq 1\}$ 为某个 norm 的 unit ball, 且

$$E_{LJ} = \{x \mid x^\top Q x \leq 1\}, \quad Q \succ 0.$$

定义 quadratic norm $\|x\|_{LJ} = (x^\top Q x)^{1/2}$ 。由于 C centrally symmetric,

$$\boxed{\|x\|_{LJ} \leq \|x\| \leq \sqrt{n} \|x\|_{LJ}} \quad \forall x.$$

几何上: 复杂 unit ball 被一个 ellipsoid 在 dimension-dependent factor 内逼近。

算法上: 可用 Q 作为 metric/preconditioner, 把非 Euclidean geometry 转换为 quadratic geometry。

Boyd & Vandenberghe, *Convex Optimization*, §8.4.1, approximating a norm by a quadratic norm

Maximum-volume inscribed ellipsoid

用 image parameterization

$$E = \{B\mathbf{u} + \mathbf{d} \mid \|\mathbf{u}\|_2 \leq 1\}, \quad B \in \mathbb{S}_{++}^n,$$

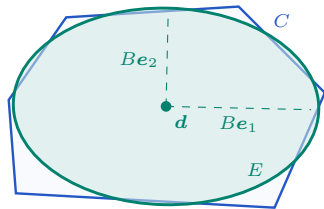
volume 与 $\det B$ 成正比。最大内接椭球由

$$\begin{aligned} & \underset{B \succ 0, \mathbf{d}}{\text{maximize}} && \log \det B \\ & \text{subject to} && B\mathbb{B}_2 + \mathbf{d} \subseteq C \end{aligned} \quad (8.14)$$

给出。若 $C = \{\mathbf{x} \mid \mathbf{a}_i^\top \mathbf{x} \leq b_i\}$, containment 精确化简为

$$\|B\mathbf{a}_i\|_2 + \mathbf{a}_i^\top \mathbf{d} \leq b_i, \quad i = 1, \dots, m. \quad (8.15)$$

这是 log-det objective 加 SOCP constraints 的 convex problem; active facets 与椭球接触。



改变 B 的 eigenvectors 与 eigenvalues, 分别调整椭球方向与半轴。

接触点

若第 i 个 facet active, 则最远方向 $\mathbf{u}_i = B\mathbf{a}_i / \|B\mathbf{a}_i\|_2$, 对应接触点 $\mathbf{x}_i^c = \mathbf{d} + B^2\mathbf{a}_i / \|B\mathbf{a}_i\|_2$ 。

内接于椭球交集

若 $C = \bigcap_i E_i$,

$$E_i = \{\mathbf{x} \mid \mathbf{x}^\top A_i \mathbf{x} + 2\mathbf{b}_i^\top \mathbf{x} + c_i \leq 0\}, \quad A_i \succ 0,$$

则 $E = \{B\mathbf{u} + \mathbf{d} \mid \|\mathbf{u}\|_2 \leq 1\} \subseteq E_i$ 可由 S-procedure 写成存在 $\lambda_i \geq 0$ 使

$$\begin{bmatrix} -\lambda_i - \mathbf{d}^\top A_i \mathbf{d} - 2\mathbf{b}_i^\top \mathbf{d} - c_i & (A_i \mathbf{d} + \mathbf{b}_i)^\top B \\ B(A_i \mathbf{d} + \mathbf{b}_i) & \lambda_i I - BA_i B \end{bmatrix} \succeq 0.$$

Schur complement 给出对 $(B, \mathbf{d}, \lambda_i)$ affine 的等价 LMI:

$$\begin{bmatrix} -\lambda_i - c_i + \mathbf{b}_i^\top A_i^{-1} \mathbf{b}_i & 0 & (\mathbf{d} + A_i^{-1} \mathbf{b}_i)^\top \\ 0 & \lambda_i I & B \\ \mathbf{d} + A_i^{-1} \mathbf{b}_i & B & A_i^{-1} \end{bmatrix} \succeq 0.$$

因此可在这些 LMI 下最小化 $-\log \det B$ 。

近似效率

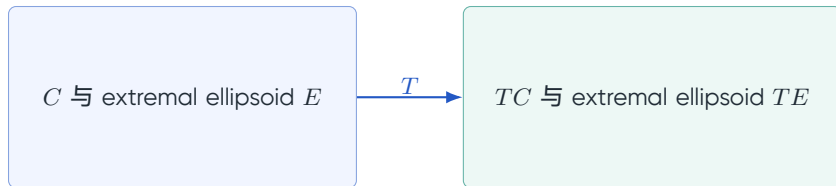
maximum-volume inner ellipsoid 扩大 n 倍覆盖一般 C ; centrally symmetric 时扩大 $\sqrt{n}$ 倍即可。

Boyd & Vandenberghe, *Convex Optimization*, §8.4.2, maximum-volume ellipsoid in an intersection of ellipsoids

Affine invariance

若 T nonsingular, 则

$$E_{\text{LJ}}(TC) = TE_{\text{LJ}}(C), \quad E_{\text{in}}(TC) = TE_{\text{in}}(C).$$



理由是 $E \mapsto TE$ 在 covering (或 inscribed) ellipsoids 之间一一对应, 且所有 volume 统一乘以 $|\det T|$, 故最优排序不变。

应用例 6: 压缩不确定性集合

工程参数的 10^4 个 feasible samples 为 $\{x_i\}$ 。用 finite-set LJ problem 得到

$$E = \{x \mid \|Ax + b\|_2 \leq 1\}$$

作为 compact uncertainty model。

数据压缩

存储 (A, b) 代替全部 samples; volume minimization 抑制过度保守。

若 samples 只是有限观测, coverage 仅对样本集合成立; 推广到真实分布仍需要统计置信保证。

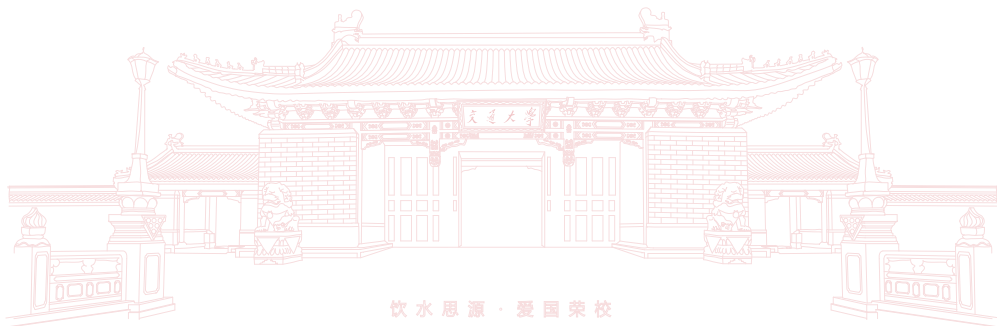
下游鲁棒优化

对 linear constraint, ellipsoidal uncertainty 的 worst case 变为一个 norm term, 通常得到 SOCP。

自编示例: minimum-volume ellipsoid as a compact uncertainty model

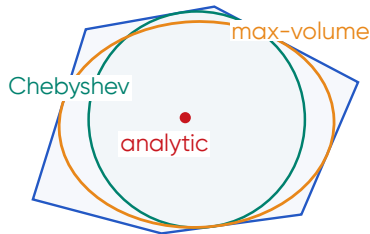
06

中心



饮水思源 · 爱国荣校

三种中心回答不同问题



Chebyshev center

最大化最小边界距离。

MVE center

最大化可容纳 ellipsoid 的体积。

Analytic center

最大化 inequalities slacks 的乘积。

三者通常不重合；所选中心应对应应用中真正关心的“居中”概念。

Boyd & Vandenberghe, *Convex Optimization*, §8.5, Figures 8.5–8.7; qualitative redraw

Chebyshev center

点 $x \in C$ 的 depth 定义为

$$\text{depth}(x, C) = \text{dist}(x, \mathbb{R}^n \setminus C).$$

Chebyshev center 是最深点, 也就是 C 内最大 norm ball 的中心:

$$x_{\text{cheb}} \in \arg \max_{x \in C} \text{depth}(x, C).$$

若 $C = \{x \mid f_i(x) \leq 0\}$, 则

$$\begin{aligned} & \underset{x, R}{\text{maximize}} && R \\ & \text{subject to} && g_i(x, R) := \sup_{\|u\| \leq 1} f_i(x + Ru) \leq 0, \end{aligned} \tag{8.16}$$

其中 g_i 是 convex function; 计算难点是其中的 convex maximization。

Polyhedron 的 Chebyshev center

若 $C = \{x \mid a_i^\top x \leq b_i\}$, 则

$$\sup_{\|u\| \leq 1} a_i^\top (x + Ru) - b_i = a_i^\top x + R\|a_i\|_* - b_i.$$

所以最大内接 norm ball 由 LP 给出:

$$\begin{array}{ll} \underset{x, R}{\text{maximize}} & R \\ \text{subject to} & a_i^\top x + R\|a_i\|_* \leq b_i, \quad i = 1, \dots, m, \\ & R \geq 0. \end{array}$$

Euclidean norm 时 $\|a_i\|_* = \|a_i\|_2$; 先把 normals normalized, 就得到“到每条 hyperplane 的距离至少 R ”。

Ellipsoids 交集的 Chebyshev center

若

$$C = \bigcap_i \{ \mathbf{x} \mid \mathbf{x}^\top A_i \mathbf{x} + 2\mathbf{b}_i^\top \mathbf{x} + c_i \leq 0 \}, \quad A_i \succ 0,$$

则 Euclidean ball $\mathbf{x} + R\mathbb{B}_2 \subseteq C$ 可经 S-procedure 与 Schur complement 转为 LMI: 存在 λ_i 使

$$\begin{bmatrix} -\lambda_i - c_i + \mathbf{b}_i^\top A_i^{-1} \mathbf{b}_i & 0 & (\mathbf{x} + A_i^{-1} \mathbf{b}_i)^\top \\ 0 & \lambda_i I & RI \\ \mathbf{x} + A_i^{-1} \mathbf{b}_i & RI & A_i^{-1} \end{bmatrix} \succeq 0,$$

对所有 i 成立, 并 maximize R 。因此原本的“ball contained in every ellipsoid”可由 SDP 精确求解。

Maximum-volume ellipsoid center

x_{mve} 定义为 C 内 maximum-volume ellipsoid 的中心。对 polyhedron

$$C = \{x \mid a_i^\top x \leq b_i, i = 1, \dots, m\},$$

求解

$$\begin{array}{ll} \underset{B \succ 0, d}{\text{maximize}} & \log \det B \\ \text{subject to} & \|Ba_i\|_2 + a_i^\top d \leq b_i \end{array}$$

后, $d^* = x_{\text{mve}}$ 。

与 Chebyshev center 的差异

Chebyshev center 只优化最短半径; MVE center 同时调整所有 principal axes, 因此会利用长而窄的方向, 并具有 affine invariance。

Analytic center

对

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \quad F\mathbf{x} = \mathbf{g},$$

analytic center 定义为 logarithmic barrier 的 minimizer:

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & -\sum_{i=1}^m \log(-f_i(\mathbf{x})) \\ \text{subject to} & F\mathbf{x} = \mathbf{g}. \end{array} \quad (8.18)$$

等价地，在 equality constraints 上最大化 slacks 的乘积

$$\prod_{i=1}^m [-f_i(\mathbf{x})].$$

需要 strict feasibility; bounded feasible set 通常保证 barrier bounded below。

Analytic center 依赖表示

同一集合用不同 inequalities 描述, analytic center 可能改变。例如区间 $[-1, 1]$:

$$x \leq 1, -x \leq 1 \implies x_{\text{ac}} = 0.$$

若重复约束 $x \leq 1$ 一次, barrier 变为

$$-2 \log(1 - x) - \log(1 + x),$$

一阶条件给出 $x_{\text{ac}} = -\frac{1}{3}$ 。

仍保持的不变性

analytic center 对 coordinate 的 nonsingular affine change、每个 inequality 的 positive scaling, 以及 equality constraints 的等价重参数化不变。

Linear inequalities 的 analytic center

对 $P = \{\mathbf{x} \mid \mathbf{a}_i^\top \mathbf{x} \leq b_i\}$, 求解

$$\underset{\mathbf{x}}{\text{minimize}} \quad - \sum_{i=1}^m \log(b_i - \mathbf{a}_i^\top \mathbf{x}). \quad (8.19)$$

若令 $d_i = (b_i - \mathbf{a}_i^\top \mathbf{x}_{\text{ac}})^{-1}$, 则 barrier Hessian 为

$$H = \sum_{i=1}^m d_i^2 \mathbf{a}_i \mathbf{a}_i^\top.$$

它定义 Dikin inner ellipsoid 与 outer ellipsoid:

$$E_{\text{in}} = \{\mathbf{x} \mid (\mathbf{x} - \mathbf{x}_{\text{ac}})^\top H (\mathbf{x} - \mathbf{x}_{\text{ac}}) \leq 1\},$$

$$E_{\text{out}} = \{\mathbf{x} \mid (\mathbf{x} - \mathbf{x}_{\text{ac}})^\top H (\mathbf{x} - \mathbf{x}_{\text{ac}}) \leq m(m-1)\},$$

且 $E_{\text{in}} \subseteq P \subseteq E_{\text{out}}$ 。

LMI 的 analytic center

对 spectrahedron

$$\mathcal{C} = \{x \mid B - \sum_{i=1}^n x_i A_i \succeq 0\},$$

analytic center 定义为

$$\underset{x}{\text{minimize}} \quad -\log \det \left(B - \sum_{i=1}^n x_i A_i \right).$$

当 smallest eigenvalue 接近零, $-\log \det$ 发散,
阻止 iterates 接触 PSD boundary。

这正是 Lecture 06 中 semidefinite interior-point method 的 barrier; “中心”与 central path 直接连接。

Boyd & Vandenberghe, *Convex Optimization*, §8.5.3, analytic center of a linear matrix inequality

应用例 7: 制造公差中心

设计变量 x 满足若干 normalized linear tolerances

$$\mathbf{a}_i^\top \mathbf{x} \leq b_i, \quad \|\mathbf{a}_i\|_2 = 1.$$

若生产误差满足 $\|\Delta x\|_2 \leq r$, Chebyshev center 求得最大共同容差:

$$\max_{\mathbf{x}, r} r \quad \text{s.t.} \quad \mathbf{a}_i^\top \mathbf{x} + r \leq b_i.$$

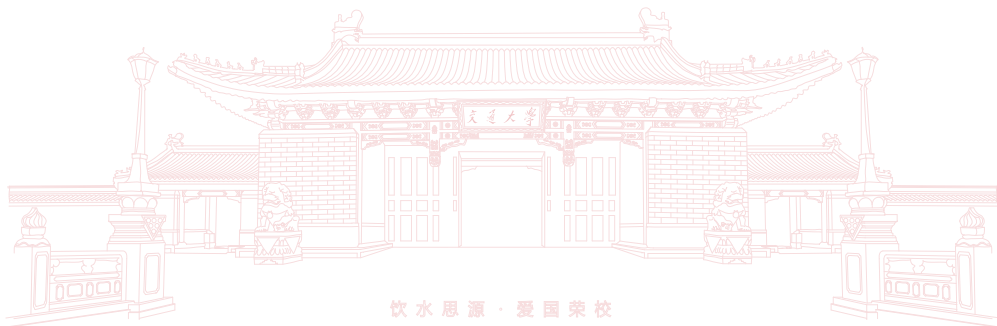
决策含义

r^* 是所有方向上都可承受的 worst-case manufacturing error; active faces 指出最先限制公差的规格。
若误差方向有不同尺度, 可改用 maximum-volume ellipsoid center.

自编示例: design centering under manufacturing tolerances

07

分类

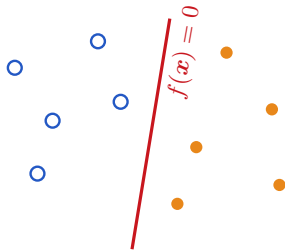


饮水思源 · 爱国荣校

分类的几何定义

给定两组 points $\{x_i\}_{i=1}^N$ 与 $\{y_j\}_{j=1}^M$, 寻找函数 f 使

$$f(x_i) > 0, \quad f(y_j) < 0.$$



若 f 对参数 affine, 则所有 sample sign constraints 对参数也是 linear。模型是否 convex 取决于参数化, 而边界可对输入 x 非线性。

Boyd & Vandenberghe, *Convex Optimization*, §8.6, Figure 8.8; qualitative redraw

Linear discrimination

令 $f(x) = a^\top x - b$ 。strict separation

$$a^\top x_i - b > 0, \quad a^\top y_j - b < 0 \quad (8.20)$$

由于对 (a, b) homogeneous, 等价于 linear feasibility

$$a^\top x_i - b \geq 1, \quad a^\top y_j - b \leq -1. \quad (8.21)$$

几何判据

两组 points 可被 affine hyperplane 严格分开, 当且仅当

$$\text{conv}\{x_i\} \cap \text{conv}\{y_j\} = \emptyset.$$

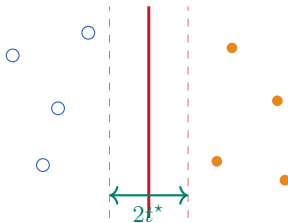
若 convex hulls 相交, 交点的两种 convex combinations 形成 alternative certificate。

Maximum-margin separation

为排除任意 scaling, 约束 $\|a\|_2 \leq 1$:

$$\begin{array}{ll}\underset{a,b,t}{\text{maximize}} & t \\ \text{subject to} & a^\top x_i - b \geq t, \\ & a^\top y_j - b \leq -t, \\ & \|a\|_2 \leq 1.\end{array}\tag{8.23}$$

若可严格分离, 则最优点 $\|a^*\|_2 = 1$, 而 $2t^*$ 是 separating slab thickness。



Boyd & Vandenberghe, *Convex Optimization*, §8.6.1, problem (8.23), Figure 8.9

Margin 等于 convex hull distance 的一半

maximum-margin problem 的 dual 可解释为

$$t^* = \frac{1}{2} \min_{\substack{\mathbf{u} \geq 0, \mathbf{1}^\top \mathbf{u} = 1 \\ \mathbf{v} \geq 0, \mathbf{1}^\top \mathbf{v} = 1}} \left\| \sum_i u_i \mathbf{x}_i - \sum_j v_j \mathbf{y}_j \right\|_2.$$

右侧两点分别属于两组 convex hull; 因此

$$2t^* = \text{dist}(\text{conv}\{\mathbf{x}_i\}, \text{conv}\{\mathbf{y}_j\}).$$

Support vectors

最优 convex combinations 中 positive weights 对应 closest faces; primal classifier 仅由这些靠近 slab 的 points 决定。

不可分数据: LP slack model

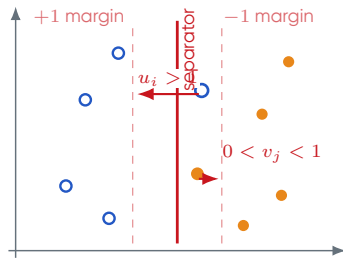
引入 $u_i, v_j \geq 0$:

$$\mathbf{a}^\top \mathbf{x}_i - b \geq 1 - u_i, \quad \mathbf{a}^\top \mathbf{y}_j - b \leq -(1 - v_j). \quad (8.24)$$

最小化总 violation:

$$\begin{aligned} & \underset{\mathbf{a}, b, \mathbf{u}, \mathbf{v}}{\text{minimize}} && \mathbf{1}^\top \mathbf{u} + \mathbf{1}^\top \mathbf{v} \\ & \text{subject to} && (8.24), \quad \mathbf{u}, \mathbf{v} \geq 0. \end{aligned} \quad (8.25)$$

$0 < u_i < 1$ 表示 point 仍正确分类但落入 margin slab;
 $u_i > 1$ 表示 misclassification。 ℓ_1 objective 是
misclassification count 的 convex surrogate。



Slack 的数值读法

对正类 score $s_i = \mathbf{a}^\top \mathbf{x}_i - b$, 最小可行 slack 为 $u_i = \max\{0, 1 - s_i\}$: $s_i = 0.3$ 给出 $u_i = 0.7$; $s_i = -0.4$ 给出 $u_i = 1.4$, 点已越过 classifier。

Support vector classifier

标准 soft-margin model 为

$$\begin{array}{ll}\underset{\mathbf{a}, b, \mathbf{u}, \mathbf{v}}{\text{minimize}} & \|\mathbf{a}\|_2 + \gamma(\mathbf{1}^\top \mathbf{u} + \mathbf{1}^\top \mathbf{v}) \\ \text{subject to} & \mathbf{a}^\top \mathbf{x}_i - b \geq 1 - u_i, \\ & \mathbf{a}^\top \mathbf{y}_j - b \leq -(1 - v_j), \quad \mathbf{u}, \mathbf{v} \geq 0.\end{array}$$

$\|\mathbf{a}\|_2$ 越小, slab width $2/\|\mathbf{a}\|_2$ 越大。

γ 越大, violations 越昂贵, 模型更强调 training fit。

与常见 $\frac{1}{2}\|\mathbf{a}\|_2^2 + C \sum \xi_i$ 形式相比, 损失含义相同, 但 regularizer 的具体尺度不同。

Logistic discrimination

模型

$$\Pr(z = 1 \mid \mathbf{u}) = \frac{e^{\mathbf{a}^\top \mathbf{u} - b}}{1 + e^{\mathbf{a}^\top \mathbf{u} - b}}, \quad \Pr(z = 0 \mid \mathbf{u}) = \frac{1}{1 + e^{\mathbf{a}^\top \mathbf{u} - b}} \quad (8.26)$$

通过 convex maximum likelihood 求 $(\mathbf{a}, b)$; decision boundary 仍是 $\mathbf{a}^\top \mathbf{u} = b$ 。

与 margin model 的区别

SVC 直接优化 geometric margin 与 hinge-like violations; logistic model 优化 probabilistic log-loss, 并输出 class probability。若 data strictly separable 且无 regularization, logistic MLE 不存在有限解。

Boyd & Vandenberghe, *Convex Optimization*, §8.6.1, equations (8.26)–(8.27), Figure 8.12; Remark 8.1

Quadratic discrimination

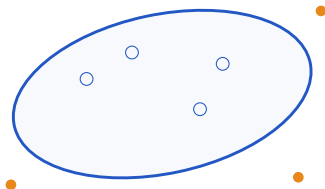
令

$$f(x) = x^\top P x + q^\top x + r.$$

sample sign constraints 对参数 (P, q, r) 仍为 linear:

$$x_i^\top P x_i + q^\top x_i + r \geq 1, \quad y_j^\top P y_j + q^\top y_j + r \leq -1.$$

若要求 $P \preceq -I$, 则 positive region 是一个 ellipsoid, 可用 SDP 寻找“包含所有 x_i 、排除所有 y_j ”的 ellipsoidal classifier。



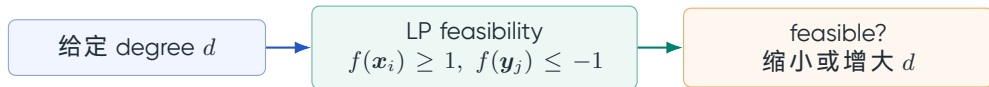
Boyd & Vandenberghe, *Convex Optimization*, §8.6.2, quadratic discrimination, Figure 8.13

Polynomial discrimination

degree at most d 的 polynomial

$$f(x) = \sum_{|\alpha| \leq d} c_\alpha x^\alpha$$

在 sample points 上对 coefficients c_α 为 affine。因此固定 d 时，separation 是 LP feasibility。



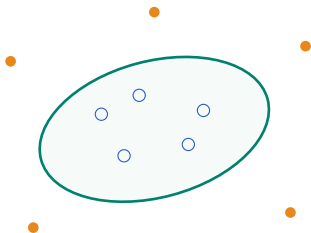
minimum separating degree 可通过整数 d 上的 bisection/search 求得；模型可能过拟合，实际应用需 regularization 或验证集。

应用例 8: 质量缺陷筛查

二维 features 为振动幅值与温升, 正常样本位于中央、缺陷样本围绕外侧。linear classifier 无法分离。几何上可写成

$$f(\mathbf{x}) = r^2 - (\mathbf{x} - \mathbf{c})^\top Q(\mathbf{x} - \mathbf{c}), \quad Q \succ 0$$

一个 ellipsoidal acceptance region。

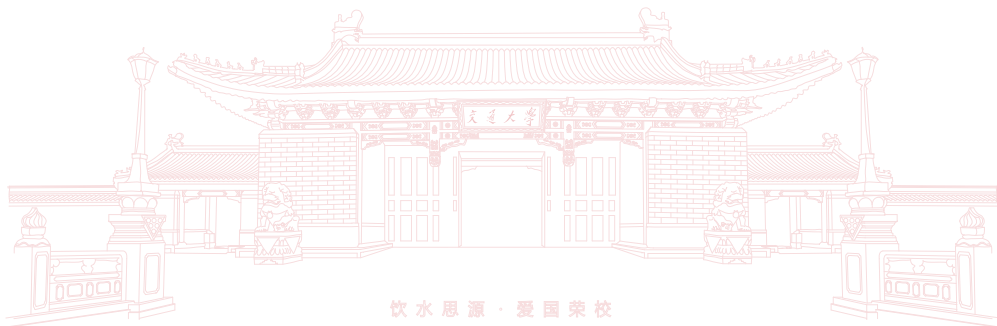


为保持训练问题 convex, 使用一般形式 $g(\mathbf{x}) = \mathbf{x}^\top P\mathbf{x} + \mathbf{q}^\top \mathbf{x} + r_0$, 对系数施加 $P \preceq -I$, 并加入 sample slacks; 求解 SDP 后再配方恢复 $(Q, \mathbf{c}, r)$ 。输出边界可解释为可接受工况 envelope。

自编示例: ellipsoidal classifier for condition-monitoring data

08

布局与选址

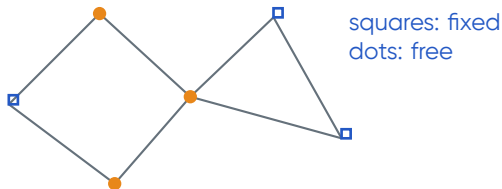


Placement 的 graph model

节点 i 的位置为 $\mathbf{x}_i \in \mathbb{R}^k$ (通常 $k = 2$ 或 3), links 为 graph edge set $\mathcal{A}$ 。若部分位置固定, 其余为变量, 则

$$\underset{\{\mathbf{x}_i\}}{\text{minimize}} \quad \sum_{(i,j) \in \mathcal{A}} f_{ij}(\mathbf{x}_i, \mathbf{x}_j)$$

在所有 f_{ij} jointly convex 时是 convex problem。



Linear facility location

最基本 cost 是 weighted link length:

$$\underset{\{\mathbf{x}_i\}}{\text{minimize}} \quad \sum_{i < j} w_{ij} \|\mathbf{x}_i - \mathbf{x}_j\|, \quad w_{ij} \geq 0. \quad (8.28)$$

Euclidean norm

直接距离，适合仓库运输、无线节点或能够沿直线连接的系统；模型可写为 SOCP。

w_{ij} 可表示 transport volume、wire multiplicity 或 unit-length cost。

Manhattan norm

$\|\mathbf{x}_i - \mathbf{x}_j\|_1$ 描述只能水平/垂直布线或街区路网；模型可写为 LP。

Boyd & Vandenberghe, *Convex Optimization*, §8.7.1, equation (8.28)

教材例 8.4: 一个自由点

给定 K 个 fixed points (u_i, v_i) , 选择 (u, v) 最小化距离和。

ℓ_1 distance

$$\min_{u,v} \sum_i (|u - u_i| + |v - v_i|).$$

任意 coordinate-wise median 是最优解; K 为偶数时可出现一整块 optimal rectangle。

mean、median 与 geometric median 对应不同 loss; 不能仅凭“中心”一词互换。

Euclidean distance

$$\min_{u,v} \sum_i \sqrt{(u - u_i)^2 + (v - v_i)^2}.$$

最优点称为 Weber point 或 geometric median, 通常需迭代算法求解。

Placement constraints

保持 convexity 的常见约束包括:

- ▣ x_i 位于 line、box、ellipsoid 等 convex region;
- ▣ pairwise distance upper bounds;
- ▣ relative position, 例如 $x_i \leq x_j$;
- ▣ linear equalities 描述 alignment。

minimum distance constraint $\|x_i - x_j\| \geq d$ 一般非凸; 它要求排除一个 convex ball。

限制 points $x_1, \dots, x_p$ 的 bounding box perimeter: 引入 u, v ,

$$u \leq x_i \leq v, \quad 2\mathbf{1}^\top(v - u) \leq P_{\max}.$$

这是 linear constraints。

Nonlinear facility costs

更一般地, 若 $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ convex nondecreasing, 则

$$\min \sum_{i < j} w_{ij} h(\|x_i - x_j\|)$$

仍为 convex problem。常见选择:

$h(t)$	行为	布局倾向
t	linear penalty	允许少量较长 links
t^2	quadratic penalty	更均衡地压缩长 links
t^4	high-order penalty	强烈抑制 maximum link length
deadzone	短距离零成本	小移动不敏感
quadratic-linear	近处平方、远处线性	平滑且降低 outlier influence

教材例 8.5: 平方距离得到均值

只有一个 free point x 时,

$$\min_x \sum_{i=1}^K \|x - x_i\|_2^2.$$

展开并令 gradient 为零:

$$2Kx - 2 \sum_{i=1}^K x_i = 0 \implies \boxed{x^* = \frac{1}{K} \sum_{i=1}^K x_i.}$$

与例 8.4 对比

ℓ_1 distance sum 给 coordinate median; Euclidean distance sum 给 geometric median; squared Euclidean distance sum 给 arithmetic mean。三者对 outliers 的敏感性逐步增加。

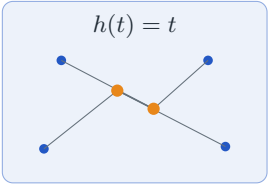
教材例 8.6: penalty 改变布局

教材比较同一 graph 上

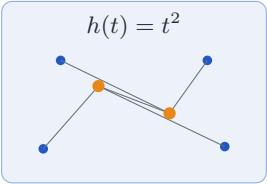
$$\sum_{(i,j) \in \mathcal{A}} \|x_i - x_j\|_2,$$

$$\sum_{(i,j) \in \mathcal{A}} \|x_i - x_j\|_2^2,$$

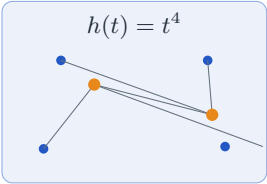
$$\sum_{(i,j) \in \mathcal{A}} \|x_i - x_j\|_2^4.$$



集中; 长尾



较均衡



压制最长 links

随着 power 增大, 长 link 的 marginal cost 增长更快; free nodes 更分散, maximum length 缩短。

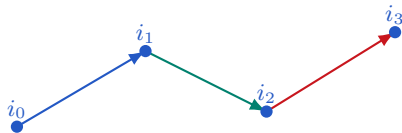
Boyd & Vandenberghe, *Convex Optimization*, §8.7.3, Example 8.6, Figures 8.15–8.17; schematic redraw

Path constraints

path $i_0, \dots, i_p$ 的长度

$$L = \sum_{k=1}^p \| \mathbf{x}_{i_k} - \mathbf{x}_{i_{k-1}} \|$$

是 node locations 的 convex function, 因此 $L \leq L_{\max}$ 是 convex constraint。



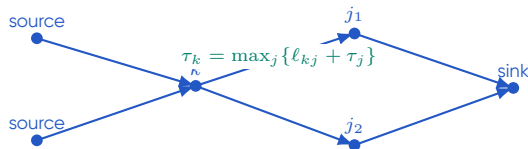
限制 path sum 与逐条限制 edge length 不同: 前者允许在路径内部重新分配 delay budget。

Minimax delay 与 dynamic programming

在 DAG 中, 令 τ_k 为从 node k 到任一 sink 的最大总 delay。sink 上 $\tau_k = 0$; 其余节点满足 Bellman recursion

$$\tau_k = \max_{(k,j) \in \mathcal{A}} \{ \ell_{kj} + \tau_j \}. \quad (8.30)$$

最大 source-sink delay 是 source nodes 上的 $\max \tau_k$ 。



计算结构

按 reverse topological order 计算 τ , 只需约 $|\mathcal{A}|$ 次 edge relaxations, 避免枚举可能指数多的 source-sink paths。

Minimax delay 的凸 reformulation

引入 upper bounds $T_k \geq \tau_k$:

$$\begin{array}{ll}\text{minimize} & t \\ \text{subject to} & T_k = 0, \quad k \text{ sink}, \\ & T_k \geq \|x_j - x_k\| + T_j, \quad (k, j) \in \mathcal{A}, \\ & T_k \leq t, \quad k \text{ source}.\end{array}$$

每条 edge 只产生一个 convex inequality; 目标压紧所有 upper bounds, 因此最优 t 等于真实 maximum path delay。

建模模式

用 epigraph inequalities 取代 max recursion 的 equality, 是把 dynamic programming relation 嵌入 convex optimization 的典型方法。

应用例 9: 应急设施选址

居民点 x_i 的需求权重为 w_i , 选一个仓库位置 z 。

平均运输成本:

$$\min_{z \in C} \sum_i w_i \|z - x_i\|_2.$$

得到 weighted geometric median, 强调总成本。

最坏响应时间:

$$\min_{z \in C} \max_i \frac{\|z - x_i\|_2}{s_i}.$$

得到 weighted 1-center, 强调最慢服务点。

实际约束

C 可加入 land-use polygon、河流一侧、已有道路附近等 convex approximations; 公平性与效率对应不同 objective。

自编示例: single-facility emergency logistics: median versus minimax center

应用例 10: 芯片互连布局

modules 的位置 $\mathbf{x}_i = (x_i, y_i)$, netlist 权重 w_{ij} 近似 signal traffic。quadratic placement

$$\min \sum_{i < j} w_{ij} \|\mathbf{x}_i - \mathbf{x}_j\|_2^2$$

形成 sparse graph Laplacian linear system; 再加入 bounding box、region 与 symmetry constraints。

优势

大规模、平滑、可用 sparse linear algebra 快速求解, 适合 global placement 初始阶段。

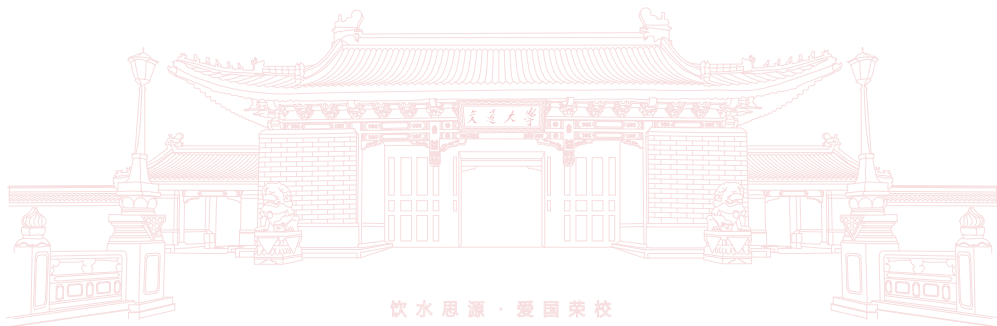
局限

points 没有面积, 不能阻止 modules overlap; 需要后续 legalization 或进入 floor-planning model。

自编示例: quadratic placement for integrated-circuit netlists

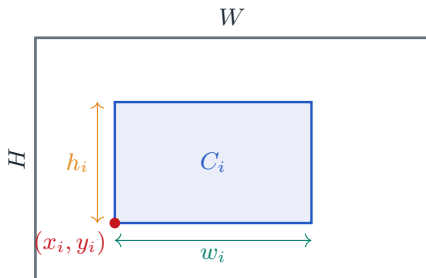
09

平面规划



Floor-planning variables

cell C_i 用 lower-left corner (x_i, y_i) 、width w_i 、height h_i 表示; bounding box 为 $W \times H$ 。



containment constraints:

$$x_i, y_i \geq 0,$$

$$x_i + w_i \leq W, \quad y_i + h_i \leq H. \quad (8.31)$$

variables 还允许 cells 在
area/aspect constraints 内重新配置。

Non-overlap 是 disjunction

对任意 $i \neq j$ ，不重叠要求以下四式至少一个成立：

$$\underbrace{x_i + w_i \leq x_j}_{i \text{ left of } j} \quad \vee \quad \underbrace{x_j + w_j \leq x_i}_{j \text{ left of } i} \quad \vee \quad \underbrace{y_i + h_i \leq y_j}_{i \text{ below } j} \quad \vee \quad \underbrace{y_j + h_j \leq y_i}_{j \text{ below } i}. \quad (8.32)$$

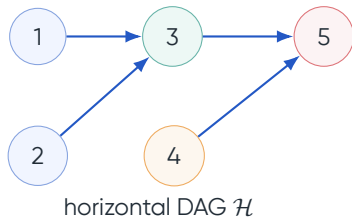
组合性来源

每个 inequality 本身 linear；困难来自“or”。若 relative order 未知，需 discrete variables/search。若预先指定每对 cells 的相对位置，剩余 continuous sizing/placement 可转为 convex optimization。

Relative positioning graphs

用 DAG $\mathcal{H}$ 表示 left-of, $\mathcal{V}$ 表示 below:

$$x_i + w_i \leq x_j \quad (i, j) \in \mathcal{H}, \quad y_i + h_i \leq y_j \quad (i, j) \in \mathcal{V}. \quad (8.34)$$



transitive relation 的所有 inequalities 不必显式保留; 只需 DAG edges。path $i \rightarrow j$ 自动推出 C_i 位于 C_j 左侧或下方。每对 cells 必须在 $\mathcal{H}$ 或 $\mathcal{V}$ 的某个方向上可达, 才能保证 non-overlap。

Bounding constraints 也可稀疏化

利用 partial order, 只需对 extremal cells 施加外框约束:

$$\begin{aligned}x_i &\geq 0, & i \text{ minimal in } \mathcal{H}, \\x_i + w_i &\leq W, & i \text{ maximal in } \mathcal{H}, \\y_i &\geq 0, & i \text{ minimal in } \mathcal{V}, \\y_i + h_i &\leq H, & i \text{ maximal in } \mathcal{V}.\end{aligned}\tag{8.35}$$

其他 cells 的 containment 由 path inequalities 与 extremal constraints 推出。

算法意义

用 transitive reduction 保留必要 edges, 可显著减少 linear constraints; 这不改变 feasible set, 只改变表示规模与数值求解成本。

Convex floor-planning model

给定 $\mathcal{H}, \mathcal{V}$ 后, 可最小化 bounding-box perimeter:

$$\min 2(W + H)$$

subject to relative-position 与 bounding constraints, 以及:

▣ spacing: $x_i + w_i + \rho \leq x_j$;

▣ minimum area: $w_i h_i \geq A_i$;

▣ aspect ratio: $\ell_i w_i \leq h_i \leq u_i w_i$.

minimum-area constraint 是 convex, 因为

$$w_i \geq A_i/h_i, \quad w_i, h_i > 0,$$

或等价地 $\log w_i + \log h_i \geq \log A_i$.

Boyd & Vandenberghe, *Convex Optimization*, §8.8.2, minimum spacing, area, and aspect-ratio constraints

Alignment、symmetry 与 similarity

这些设计规则在 relative order 固定后可用 linear equalities:

$$\begin{array}{ll}\text{cell } i \text{ center line aligns with cell } j \text{ top:} & y_i + h_i/2 = y_j + h_j, \\ \text{vertical symmetry:} & x_{\text{axis}} - (x_i + w_i/2) = x_j + w_j/2 - x_{\text{axis}}, \\ \text{fixed-scale similarity:} & w_i = aw_j, \quad h_i = ah_j.\end{array}$$

边界

若 similarity scale a 也是变量, $w_i = aw_j$ 是 bilinear, 不能直接放入本节的 convex formulation; GP parameterization 则可把未知 positive scale 作为 monomial relation 处理。

containment in a polyhedron 与指定 cell 覆盖某个 point 也只需 linear inequalities。

Cell 间距约束

center distance 可直接约束:

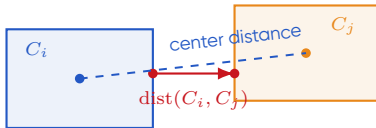
$$\|(x_i + w_i/2, y_i + h_i/2) - (x_j + w_j/2, y_j + h_j/2)\| \leq D_{ij}.$$

若限制 sets 的 minimum distance, 可引入 $p_i \in C_i, p_j \in C_j$:

$$p_i \in C_i, \quad p_j \in C_j, \quad \|p_i - p_j\| \leq D_{ij}.$$

方向要分清

$\text{dist}(C_i, C_j) \leq D$ 可由存在一对近点表达, 因此 convex; 而要求 $\text{dist}(C_i, C_j) \geq D$ 是 clearance/non-overlap 型约束, 通常非凸, 需借助已知 relative positioning 才能线性化。

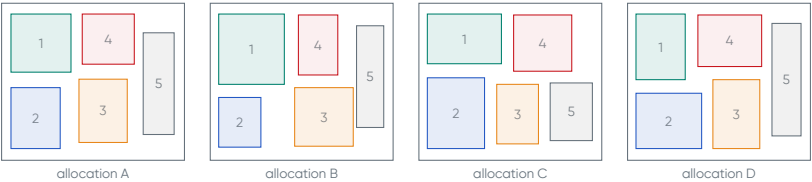


教材例 8.7: 五个 cells

教材固定 Figure 8.19 的 relative-position graphs, 统一要求

$$\frac{1}{5} \leq \frac{w_i}{h_i} \leq 5,$$

并使用相同 minimum spacing; 四组 minimum areas A_i 的总和相同, 但分配不同。



最优 perimeter 与 aspect ratios 会随 area allocation 改变; 相同 total area 不意味着相同 geometry 或 bounding box。

Floor planning via geometric programming

所有 variables positive 时, relative-position inequality

$$x_i + w_i \leq x_j \iff \frac{x_i + w_i}{x_j} \leq 1$$

是 posynomial inequality; $w_i h_i = A_i$ 是 monomial equality。

GP 能直接处理

最小化 area WH ; 固定 cell area;
unknown-scale similarity; relative order 与
bounding inequalities。

convex 与 GP formulations 各自保留不同的设计语言, 应根据所需 constraints 选择。

本 formulation 不便处理

alignment、additive symmetry 与一般 distance
constraints 不符合 monomial/posynomial 结构。

应用例 11: 数据中心机柜布局

每个 rack 是 rectangle; relative graphs 固定冷热通道方向。目标综合 perimeter 与 cable cost:

$$\min 2\alpha(W + H) + \sum_{(i,j) \in \mathcal{A}} w_{ij} \|\mathbf{c}_i - \mathbf{c}_j\|_1.$$

加入

$$w_i h_i \geq A_i, \quad \ell_i w_i \leq h_i \leq u_i w_i, \quad \text{aisle spacing } \rho_{ij}.$$

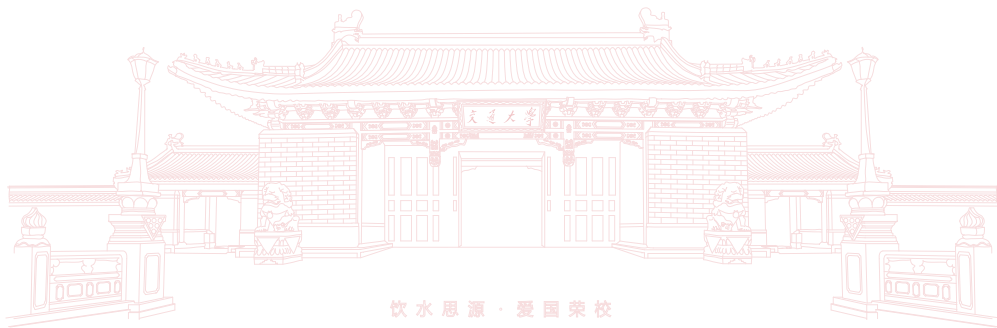
模型边界

给定 relative order 后这是 convex model; 若 aisle topology、rack orientation 或 pairwise order 也要选择, 则应在外层加入 mixed-integer decisions。

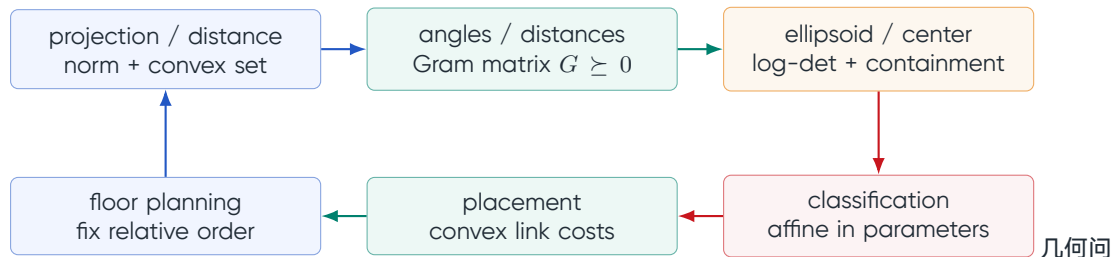
自编示例: data-center rack floor planning with cable and aisle costs

10

总结



Chapter 8 的建模链



题的核心工作是选择 representation: 坐标适合位置, Gram matrix 适合不变量, ellipsoid parameters 适合 containment, graph order 适合消除 disjunction。

公式索引

对象	关键公式
Projection	$P_C(\mathbf{x}_0) \in \arg \min_{\mathbf{x} \in C} \ \mathbf{x} - \mathbf{x}_0\ $
Set distance dual	$\max_{\ \mathbf{z}\ _* \leq 1} -S_C(-\mathbf{z}) - S_D(\mathbf{z})$
Gram geometry	$G = A^\top A \succeq 0, d_{ij}^2 = G_{ii} + G_{jj} - 2G_{ij}$
Euclidean distance matrix	$D_{ii} = 0, D_{ij} \geq 0, JDJ \preceq 0$
Outer ellipsoid	$\min \log \det A^{-1} \text{ s.t. } \sup_{\mathbf{x} \in C} \ A\mathbf{x} + \mathbf{b}\ _2 \leq 1$
Inner ellipsoid	$\max \log \det B \text{ s.t. } B\mathbb{B}_2 + \mathbf{d} \subseteq C$
Chebyshev center	$\max R \text{ s.t. } \mathbf{x} + R\mathbb{B} \subseteq C$
Analytic center	$\min -\sum_i \log(-f_i(\mathbf{x}))$
Floor planning	relative order converts non-overlap disjunctions into linear inequalities

模型选择

已知信息	变量	首选结构
point coordinates 与 convex regions	coordinates	norm/QP/SOCP
pairwise angles/distances	Gram or EDM	SDP
cover/containment geometry	ellipsoid parameters	log-det + SOCP/SDP
class labels	boundary parameters	LP/QP/SDP
network links	node coordinates	separable convex costs
rectangles + fixed order	positions and sizes	LP/SOCP/GP
rectangles + unknown order	continuous + discrete	mixed-integer/outer search

先识别 invariant、containment 和 disjunction；它们决定能否得到 convex formulation。



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

自然科学研究院
INSTITUTE OF NATURAL SCIENCES

Fin

