高级算法/ 2026 年秋季

线性规划与优化专题

半正定规划

矩阵松弛与Goemans–Williamson 舍入

第9 讲：100 分钟

本讲义仅覆盖栗师老师负责的线性规划专题。本专题共九讲，每讲100 分钟，讲次编号与课件一致。第一、第二部分是本讲内部的阅读划分；选读附录供课后拓展。

授课教师：栗师
南京大学

以授课教师的课件为依据，借助AI 将讲解、证明和课后练习展开为文字。中英文版本的数学内容相互对应。

学习目标与阅读路线

最后一讲将松弛方法从线性规划推广到半正定规划。第一部分建立矩阵语言与最大割松弛；第二部分用随机超平面对向量解舍入，并证明近似保证。半正定规划对偶作为选读内容提供。

- 理解半正定性的二次型、谱分解和Gram 矩阵刻画之间的等价性。
- 将最大割写成SDP，并用负特征值方向分离非半正定矩阵。
- 推导“夹角除以π”的切割概率与常数α_{GW}。

预备知识为初等线性代数、实对称矩阵的谱定理、期望，以及第4 讲的分离判定器观点。下文边权均非负；课件中的无权图即所有边权为一的特例。

目录

1 第一部分：从割到半正定矩阵 1
1.1 最大割与二分之一近似基准 . . . 1
1.2 朴素线性松弛为何不足 . . . 1
1.3 将正负符号替换为单位向量 . . . 1
1.4 半正定矩阵 . . . 2
1.5 矩阵表述与分离 . . . 2

2 第二部分：Goemans–Williamson 随机超平面舍入 2
2.1 舍入算法 . . . 2
2.2 夹角引理 . . . 3
2.3 近似常数 . . . 3
2.4 一个完整的三角形例子 . . . 4
2.5 精度、概率与困难性 . . . 4
2.6 课后练习 . . . 4

A 选读：半正定规划对偶 5
A.1 对偶锥与弱对偶 . . . 5
A.2 推导及其与LP 的关系 . . . 5

1 第一部分：从割到半正定矩阵

1.1 最大割与二分之一近似基准

给定无自环无向图$G=(V,E)$ 及边权$w_e \geq 0$，最大割问题要求将顶点划分为$(S, V \setminus S)$，使跨越两侧的边权和最大。记$W = \sum_e w_e$。互补问题Min-Uncut 最小化两端位于同侧的边权和，也就是使图成为二部图所需删除边的最小权重。二者精确最优值之和为W，但一个问题的乘法近似保证不能按同一因子直接转移到另一个问题。

命题1.1. 存在简单的随机1/2 近似算法，也存在具有同样保证的确定性贪心算法。

证明. 独立地以1/2 的概率将每个顶点放到任意一侧。每条边跨越两侧的概率为1/2，故期望割权重为$W/2 \geq \mathrm{OPT}/2$。贪心算法则按任意顺序处理顶点，将下一个顶点放在使其连接已处理顶点的边中至少一半权重被切开的那一侧。每条边在第二个端点被放置时计费，最终至少切开总权重的一半。 □

另一种证明使用局部搜索：若移动任意一个顶点都不能改善割，则每个顶点连接另一侧的边权至少等于连接同侧的边权。对所有顶点求和，得到割权重至少为$W/2$。对于无权图，每次改进都使整数割大小增加，因此最多进行$|E|$ 次移动。这个迭代次数界不能直接推广到任意二进制编码边权。

1.2 朴素线性松弛为何不足

一种看似自然的编码采用分侧变量$y_v \in [0,1]$ 和割变量x_{uv}，要求$x_{uv} \leq |y_u - y_v|$。这既不是线性约束，也不是凸约束：点$(y_u, y_v, x_{uv}) = (1,0,1/2)$ 和$(0,1,1/2)$ 都可行，而中点$(1/2,1/2,1/2)$ 不可行。若改为$x_{uv} \geq |y_u - y_v|$，虽然可以线性化，却反转了所需关系。一个有效LP 为每对顶点设置$x_{uv} \in [0,1]$，并加入

$$x_{uv} + x_{vw} + x_{uw} \leq 2 \qquad (u,v,w \text{ 互异}).$$

任意割至多切开一个三角形的两条边，所以所有割都可行。然而，在无三角形图上，将图中的边赋值为一、非边赋值为零，就满足所有这些约束，并取得目标值W。无权情形下，存在最大割任意接近$|E|/2$ 的无三角形图，故该松弛的整数间隙可接近二。这一存在性结论可由足够稀疏的随机图再删除三角形证明；并非每个无三角形图的最大割都很小，例如二部图可以切开全部边。

1.3 将正负符号替换为单位向量

用$s_v \in \{-1,1\}$ 表示顶点v 所在的一侧。于是$(1 - s_u s_v)/2$ 恰好是边uv 被切开的示性函数。因此

$$\mathrm{OPT} = \max_{s_v \in \{-1,1\}} \frac{1}{2} \sum_{uv \in E} w_{uv}(1 - s_u s_v).$$

将每个符号松弛为单位向量$y_v \in \mathbb{R}^n$，其中$n = |V|$:

$$\max \ \frac{1}{2} \sum_{uv \in E} w_{uv}(1 - \langle y_u, y_v \rangle) \quad \text{约束为} \|y_v\|_2 = 1 \quad (v \in V). \tag{1.1}$$

将y_v 取为$s_v e_1$，即可嵌入任意割，所以松弛最优值至少为OPT。上述向量表述包含二次表达式；通过Gram 矩阵换元，才能直接看出其凸性。

1.4 半正定矩阵

定义1.2. 若实对称矩阵X 对每个实向量z 都满足$z^{\mathsf{T}}Xz \geq 0$，则称其为半正定矩阵，记为$X \succeq 0$。当$X - Y \succeq 0$ 时记$X \succeq Y$。正定性记为$X \succ 0$，要求对每个非零z 都严格为正。

定理1.3. 对实对称$n \times n$ 矩阵X，以下条件等价：

1. $X \succeq 0$.
2. X 的所有特征值均非负。
3. $X = V^{\mathsf{T}}V$，其中实矩阵V 至多有n 行。
4. $X = \sum_{k=1}^{n} \lambda_k q_k q_k^{\mathsf{T}}$，其中$\lambda_k \geq 0$，且$q_k$ 构成标准正交基。

证明. 谱定理给出$X = Q\Lambda Q^{\mathsf{T}}$。将每个特征向量代入二次型，便知半正定性蕴含$\lambda_k \geq 0$。反过来，非负特征值给出$z^{\mathsf{T}}Xz = \sum_k \lambda_k (q_k^{\mathsf{T}}z)^2 \geq 0$。取$V = \Lambda^{1/2}Q^{\mathsf{T}}$ 得到Gram 分解，其中零行可删去。最后，$X = V^{\mathsf{T}}V$ 蕴含$z^{\mathsf{T}}Xz = \|Vz\|_2^2 \geq 0$。谱展开式只是将同一个分解逐项写出。 □

V 的列向量即y_v，满足$X_{uv} = \langle y_u, y_v \rangle$。因此，任意更高维空间中的$n$ 个向量，都可由$\mathbb{R}^n$ 中的向量保持所有两两内积来表示。半正定不等于逐元素非负：$\left(\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right)$ 是半正定矩阵；逐元素非负也不蕴含半正定，因为$\left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$ 有特征值-1。

1.5 矩阵表述与分离

对称矩阵的内积定义为$A \bullet X = \operatorname{tr}(A^{\mathsf{T}}X) = \sum_{u,v} A_{uv}X_{uv}$。在这个完整矩阵内积中，非对角元素计数两次。一般SDP 具有线性矩阵目标、线性标量约束和半正定约束，例如

$$\min C \bullet X \quad \text{约束为} A_k \bullet X \geq b_k \ (k \in [m]), \quad X \succeq 0.$$

课件中的数值例子取$X = \left(\begin{smallmatrix} z_1 & z_2 \\ z_2 & z_3 \end{smallmatrix}\right)$，目标为$5z_1+6z_2+7z_3$，约束为$z_1+3z_2+4z_3 \geq 5$、$2z_1+3z_2+z_3 \geq 10$ 和$3z_1 + 2z_2 + 2z_3 \geq 7$。例如，其目标矩阵为$C = \left(\begin{smallmatrix} 5 & 3 \\ 3 & 7 \end{smallmatrix}\right)$，非对角元素应为三，而不是六。

最大割的SDP

$$\text{SDP} = \max \ \frac{1}{2} \sum_{uv \in E} w_{uv}(1 - X_{uv}) \quad \text{约束为} X_{vv} = 1 \ (v \in V), \quad X \succeq 0. \tag{1.2}$$

半正定锥是凸的，因为固定z 后，每个不等式$z^{\mathsf{T}}Xz \geq 0$ 都关于X 线性。(1.2) 的可行域是紧的，因为所有元素均满足$|X_{uv}| \leq \sqrt{X_{uu}X_{vv}} = 1$。

要将对称矩阵X 与半正定锥分离，可寻找一个负特征值对应的特征向量z。此时$z^{\mathsf{T}}Xz < 0$，而任意半正定矩阵Y 都满足$(zz^{\mathsf{T}}) \bullet Y \geq 0$，这便是关于矩阵元素的分离线性不等式。其他线性约束可以直接检查。这解释了SDP 与椭球法的联系，但数值精度与条件数仍需考虑。一般SDP 并不继承有理LP 的全部精确可行性及最优值达到性质。对于最大割，I 在对角线等式确定的仿射空间内严格可行，且可行域有界。标准SDP 算法可按指定精度计算近似最优可行解。下面先分析理想最优解，再说明目标误差的影响。

2 第二部分：Goemans–Williamson 随机超平面舍入

2.1 舍入算法

将一个可行SDP 解分解为单位向量y_v，并为所有顶点使用同一个随机方向。共享该方向能够保留两两夹角的信息；若为每个顶点独立随机选边，则会丢失这些信息。

随机超平面舍入

1. 随机生成$g \in \mathbb{R}^n$，各坐标独立满足$g_k \sim N(0,1)$。

2. 令$S = \{v : \langle g, y_v \rangle \geq 0\}$，返回$(S, V \setminus S)$。

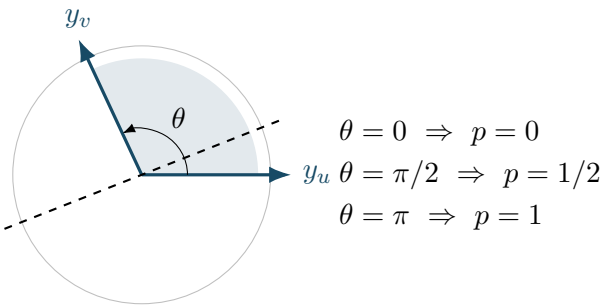
对任意固定单位向量，落在边界上的概率为零，所以等号如何处理不影响分析。无需将g 归一化，因为只使用内积的正负号。

2.2 夹角引理

引理2.1. 若单位向量y_u, y_v 的夹角为$\theta \in [0, \pi]$，则

$$\mathbb{P}[uv \text{ 被切开}] = \frac{\theta}{\pi} = \frac{\arccos\langle y_u, y_v \rangle}{\pi}.$$

证明. 先假设两向量线性独立，将g 投影到它们张成的二维平面上。两个内积的正负仅依赖这一投影。高斯分布的旋转对称性使投影方向在圆周上均匀分布。分别与两个向量内积非负的方向各构成一个半圆，这两个半圆的对称差总角度为2θ，故概率为$2\theta/(2\pi)$。若两向量相同，概率为零；若方向相反，概率为一。这也覆盖了其余情形。 □



2.3 近似常数

对一条边，记$t = \langle y_u, y_v \rangle = \cos\theta$。乘以边权之前，其SDP 贡献为$(1-t)/2$，而舍入后的期望贡献为$\arccos(t)/\pi$。定义

$$\alpha_{\mathrm{GW}} = \min_{-1 \leq t < 1} \frac{2\arccos t}{\pi(1-t)} = \min_{0 < \theta \leq \pi} \frac{2\theta}{\pi(1-\cos\theta)} \approx 0.878567.$$

该比值在$\theta \to 0$ 时趋于无穷，在$\theta = \pi$ 时等于一。其导数的符号与下式相同：

$$h(\theta) = 1 - \cos\theta - \theta\sin\theta, \qquad h'(\theta) = -\theta\cos\theta.$$

因此，h 在$(0, \pi/2)$ 上从零递减为负，在$(\pi/2, \pi)$ 上严格递增至2，后一个区间恰有一个零点，对应唯一极小值。等价地，最小化夹角满足$\tan(\theta/2) = \theta$，约为2.33112 弧度。以最小值给出的定义是精确的，小数只是其数值近似。

定理2.2 (Goemans–Williamson 保证)**.** 对最大割SDP 的最优解进行随机超平面舍入，得到的割C 满足

$$\mathbb{E}[w(C)] \geq \alpha_{\mathrm{GW}}\,\mathrm{SDP} \geq \alpha_{\mathrm{GW}}\,\mathrm{OPT}\,.$$

证明. 对每条满足$t<1$ 的边，由α_{GW} 的定义有$\arccos(t)/\pi \geq \alpha_{\mathrm{GW}}(1-t)/2$。若$t=1$，两边均为零。乘以非负边权$w_{uv}$ 后求和。期望的线性性不要求各条边被切开的事件相互独立：

$$\begin{aligned}\mathbb{E}[w(C)] &= \sum_{uv\in E} w_{uv}\mathbb{P}[uv \text{ 被切开}] \\ &\geq \alpha_{\mathrm{GW}}\frac{1}{2}\sum_{uv\in E} w_{uv}(1-\langle y_u, y_v\rangle) \\ &= \alpha_{\mathrm{GW}}\,\mathrm{SDP} \geq \alpha_{\mathrm{GW}}\ \mathrm{OPT}\,.\end{aligned}$$

□

2.4 一个完整的三角形例子

对无权三角形，每个非平凡割的值为二。将三个单位向量分别放在角度0、$2\pi/3$ 和$4\pi/3$ 处，每对内积为$-1/2$，SDP 目标值为$3\cdot(3/4)=9/4$。这已经最优，因为对任意三个单位向量，

$$\frac{1}{2}\sum_{u<v}(1-\langle y_u, y_v\rangle) = \frac{9-\|y_1+y_2+y_3\|_2^2}{4} \leq \frac{9}{4}.$$

舍入时每条边以$2/3$ 的概率被切开，因此期望割值恰为二。事实上，几乎每个超平面都会产生非平凡割。该实例中舍入值与SDP 值的比值为$2/(9/4)=8/9$，略大于通用常数。

2.5 精度、概率与困难性

该定理保证期望值，并不声称每次采样的超平面都能达到近似因子。独立重复并保留最好的割，不会降低期望质量。更定量地，当$\mathrm{OPT}>0$ 且$0<\varepsilon<\alpha_{\mathrm{GW}}$ 时，由$0\leq w(C)/\,\mathrm{OPT}\leq 1$ 可得

$$\mathbb{P}\big[w(C)\geq(\alpha_{\mathrm{GW}}-\varepsilon)\,\mathrm{OPT}\big] \geq \frac{\varepsilon}{1-\alpha_{\mathrm{GW}}+\varepsilon}.$$

若成功概率为p，则归一化割值的期望至多为$p+(1-p)(\alpha_{\mathrm{GW}}-\varepsilon)$，移项便得到上述下界。重复采样可以放大成功概率。若$\mathrm{OPT}=0$，则每个割都是最优的。若计算得到的可行SDP 解的目标值至少为$\mathrm{SDP}-\delta$，同一证明给出$\mathbb{E}[w(C)]\geq\alpha_{\mathrm{GW}}(\mathrm{OPT}-\delta)$。当$W>0$ 时，基准界$\mathrm{OPT}\geq W/2$ 允许选取与W 成比例的精度，以控制额外相对误差。实际调用数值求解器时，必须一致地处理Gram 矩阵的可行性及归一化。原算法来自Goemans 与Williamson（1995）。课件还提到一个条件最优性结论：假设唯一博弈猜想成立，达到任何固定的、严格大于α_{GW} 的近似因子都是NP 困难的，参见Khot、Kindler、Mossel 与O'Donnell 的论文。据此排除多项式时间算法还使用通常的$P\neq NP$ 假设；这不是无条件最优性定理。

2.6 课后练习

1. 用具体的W、最大割最优值及算法割值，说明最大割的常数近似比为何不能保证Min-Uncut 的常数近似比。
2. 证明：对称矩阵$\left(\begin{smallmatrix}a & b\\ b & c\end{smallmatrix}\right)$ 半正定，当且仅当$a,c\geq 0$ 且$ac-b^2\geq 0$。
3. 利用负特征值，为$X=\left(\begin{smallmatrix}1 & 2\\ 2 & 1\end{smallmatrix}\right)$ 找到一个与半正定锥分离的不等式。
4. 写出数值SDP 例子中的三个约束矩阵，并检查非对角元素中的二倍系数。
5. 验证相同、正交、反向三种向量对的切割概率。解释为什么各边被切开的事件不一定独立。
6. 改用Gram 矩阵推导三角形SDP 的最优值，再寻找达到同一上界的对偶证书。
7. 利用上面的成功概率下界，计算为了达到$\alpha_{\mathrm{GW}}-\varepsilon$ 近似比，并将失败概率控制在$\beta\in(0,1)$ 以内，需要独立重复多少次?

A 选读：半正定规划对偶

A.1 对偶锥与弱对偶

引理A.1. 若对称矩阵$X, Z \succeq 0$，则$X \bullet Z \geq 0$。反过来，若对称矩阵Z 与每个半正定矩阵的内积均非负，则$Z \succeq 0$。

证明. 由迹的循环性质，$\mathrm{tr}(XZ) = \mathrm{tr}(Z^{1/2} X Z^{1/2}) \geq 0$，因为最后一个迹中的矩阵半正定。反向只需对每个向量z 取$X = zz^{\mathsf{T}}$，便得到$z^{\mathsf{T}} Z z \geq 0$。 □

考虑如下原始与对偶规划：

$$\begin{aligned} \text{原始:} \quad & \min C \bullet X \\ & A_k \bullet X \geq b_k \quad (k \in [m]), \quad X \succeq 0; \\ \text{对偶:} \quad & \max b^{\mathsf{T}} y \\ & y \geq 0, \quad Z := C - \textstyle\sum_k y_k A_k \succeq 0. \end{aligned}$$

对可行的X, y，目标差为

$$C \bullet X - b^{\mathsf{T}} y = Z \bullet X + \sum_k y_k (A_k \bullet X - b_k) \geq 0.$$

这直接证明了弱对偶，也给出了目标相等时的互补松弛条件：$Z \bullet X = 0$，并且每个k 都有$y_k (A_k \bullet X - b_k) = 0$。

A.2 推导及其与LP 的关系

这些不等式的Lagrange 函数为

$$\mathcal{L}(X, y) = C \bullet X + \sum_k y_k (b_k - A_k \bullet X) = b^{\mathsf{T}} y + Z \bullet X, \qquad y \geq 0.$$

若$Z \succeq 0$，其在$X \succeq 0$ 上的下确界为$b^{\mathsf{T}} y$。若Z 存在负方向z，则令$X = tzz^{\mathsf{T}}$ 并使$t \to \infty$，可使下确界为$-\infty$。最大化这一目标下界，就得到上述对偶。这个有限维推导避免了将无限多个半正定不等式直接当成普通有限LP 的问题。课件通过将半正定矩阵表示为秩一矩阵zz^{T} 的非负组合，得到同一个对偶。由谱定理，这种表示至多需要n 项，无需真正的无限求和。若所有数据矩阵都是对角矩阵，则$X \succeq 0$ 保证对角元素非负，而将X 替换为其对角部分保持所有目标与约束不变。此时原始与对偶恰好退化为熟悉的LP 对：$\min\{c^{\mathsf{T}} x : Ax \geq b, x \geq 0\}$ 与$\max\{b^{\mathsf{T}} y : A^{\mathsf{T}} y \leq c, y \geq 0\}$。

注A.2 (强对偶需要额外条件)**.** 与有限LP 不同，若缺少正则性条件，一般SDP 可能存在对偶间隙或最优值无法达到的现象。对于上述不等式形式的原始问题，一个充分的Slater 条件是存在$X \succ 0$，使每个$A_k \bullet X > b_k$。若最优值有限，该条件保证零对偶间隙，且对偶最优值能够达到。等式约束应保留为等式，并采用相应的相对内部条件；其对偶乘子不受符号限制。

对于最大割，令L 为加权图Laplace 矩阵。目标可写为$L \bullet X/4$，对应对偶为

$$\min \sum_v z_v \quad \text{约束为 } \mathrm{Diag}(z) - L/4 \succeq 0, \qquad z \in \mathbb{R}^n.$$

原始矩阵I 在对角线等式确定的仿射空间内严格可行；将所有z_v 取为足够大的同一个数，则对偶严格可行。因此，这个具体SDP 具备用数值上、下界证书进行验证所需的强对偶性质。在无权三角形上，令所有$z_v = 3/4$，对偶值为$9/4$，与向量构造相匹配。

来源与阅读说明

主要依据：授课教师的课件文件夹

`9.SDP`

`Topics.tex` 与`Optional.tex`.

选读部分对应课件中的选讲内容（如有）。补充的例子和练习服务于课后阅读，不占用新的授课讲次。

课程网站. 网站上前段课程的讲义用于参考组织方式与展开程度，本专题的内容仍以线性规划课件为准。