计算方法

刘景铖

计算机软件新技术国家重点实验室
南京大学

补充：Von-Neumann Minimax Theorem

Von-Neumann Minimax Theorem.

$$\max_{x\in\Delta^m}\min_{y\in\Delta^n} x^T Ay = \min_{y\in\Delta^n}\max_{x\in\Delta^m} x^T Ay$$

证明(cont'd)：把它们都转化为标准形式

$$\begin{aligned} \max \quad & t \\ t-\sum_{i=1}^{m} x_i a_{ij} \le 0 \quad & \forall j=1,\dots,n \qquad \times y_j \\ \sum_{i=1}^{m} x_i = 1 & \qquad \times r \\ x_i \ge 0 \quad & \forall i=1,\dots,m \end{aligned}$$

$$\begin{aligned} \min \quad & r \\ r-\sum_{j=1}^{n} a_{ij}\, y_j \ge 0 \quad & \forall i=1,\dots,m \qquad \times x_i \\ \sum_{i=1}^{n} y_i = 1 & \qquad \times t \\ y_i \ge 0 \quad & \forall i=1,\dots,n \end{aligned}$$

Equality follows directly from strong duality!

Strong duality theorem

强对偶性：假设primal和dual LP都存在可行解, 则它们的最优解相等

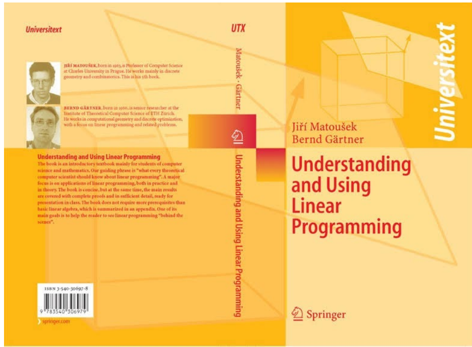
证明方法很多, 其中包括分析单纯形算法的

这节课：通过凸优化分析(Farkas lemma)给出证明

参考资料:

Understanding and Using Linear Programming

By Jiří Matoušek , Bernd Gärtner



Separation theorem

凸几何的性质：Separation theorem

定义：一个集合S，如果对$\forall x, y \in S, \forall \alpha \in [0,1]$都有$\alpha x + (1-\alpha)y \in S$，则称$S$为凸集(convex set)

定义：如果S中的所有的点列的极限点都在S中, 则称集合S为闭集(closed set)

给定一个闭的凸集$S \subseteq \mathbb{R}^n$，和集合外的点$v \notin S$。
$\exists w \in \mathbb{R}^n$ 使得$\langle w, v\rangle > \langle w, x\rangle$对任意$x \in S$成立

Separation theorem

给定一个闭的凸集$S \subseteq \mathbb{R}^n$，和集合外的点$v \notin S$。
$\exists w \in \mathbb{R}^n$ 使得$\langle w, v\rangle > \langle w, x\rangle$ 对任意$x \in S$成立
证明(sketch)：找到唯一的$x_* \in S$，使得与v最接近

1. 存在性：由于S是闭集，由Weierstrass定理，可知其最接近v的点可以在S中取到
2. 唯一性：假设有两个点，考虑它们的中点，只会与v更加接近
3. x_*为最接近的点当且仅当$\forall x \in S, \langle x - x_*, v - x_*\rangle \le 0$

   考虑$y \coloneqq (1-t)x_* + t\,x \in S$，则

$$
\begin{aligned}
\|y - v\|_2^2 &= \|x_* - v - t(x_* - x)\|_2^2 \\
&= \|v - x_*\|_2^2 - 2t\langle x - x_*, v - x_*\rangle + t^2\|x - x_*\|_2^2
\end{aligned}
$$

$$
\begin{aligned}
x_* \text{为最接近的点} &\Leftrightarrow -2t\langle x - x_*, v - x_*\rangle + t^2\|x - x_*\|_2^2 \ge 0,\ \forall x, t > 0 \\
&\Leftrightarrow \frac{t}{2}\|x - x_*\|_2^2 \ge \langle x - x_*, v - x_*\rangle,\ \forall x, t > 0 \\
&\Leftrightarrow 0 \ge \langle x - x_*, v - x_*\rangle,\ \forall x
\end{aligned}
$$

令 $w = v - x_*$，由$\|v - x_*\|_2^2 > 0$可知 $\langle v, v - x_*\rangle > \langle x_*, v - x_*\rangle$，进而$\langle v, w\rangle > \langle x_*, w\rangle$
另一方面，前面已经证明$\forall x \in S, \langle x - x_*, v - x_*\rangle \le 0$
这意味着$\langle x, v - x_*\rangle \le \langle x_*, v - x_*\rangle$，即$\langle x, w\rangle \le \langle x_*, w\rangle$，因此$\langle v, w\rangle > \langle x_*, w\rangle \ge \langle x, w\rangle, \forall x \in S$

Farkas Lemma

$Ax = b, x \geq 0$无解当且仅当存在 y 使得$y^T A \geq 0$ 且 $y^T b < 0$

证明：(<=)如果这样的 y 存在，显然无解，否则$y^T Ax \geq 0$ 与$y^T b < 0$ 矛盾

(=>)考虑 $S = \{Ax: x \geq 0\}$. S是闭的凸集. 无解意味着$b \notin S$.

由Separation theorem，存在$w \in \mathbb{R}^n$ 使得$\langle w, b\rangle > \langle w, s\rangle$ 对任意$s \in S$ 成立。令$y = -w$，则有 $y^T b < y^T Ax, \forall x \geq 0$

由于 $0 \in S$，因此$y^T b < 0$

另一方面，$y^T A \geq 0$，否则可以找到$x \geq 0$使得 $y^T Ax = -\infty$，与 $y^T b < y^T Ax$矛盾

Strong LP duality

假设primal和dual LP都存在可行解, 则它们的最优解相等

证明：

$$\max \ \langle c, x\rangle \qquad\qquad \min \ \langle b, y\rangle$$

$$Ax = b \qquad\qquad y^T A \geq c$$

$$x \geq 0.$$

只需要证明，如果primal的objective小于t，则dual objective也小于t

“primal objective小于t”本身可以表示成LP

$$\langle c, x\rangle - s = t$$

$$Ax = b$$

$$x \geq 0, s \geq 0$$

$$\boxed{\begin{matrix} \begin{bmatrix} A & 0 \\ c^T & -1 \end{bmatrix} \begin{pmatrix} x \\ s \end{pmatrix} = \begin{pmatrix} b \\ t \end{pmatrix} \\ x \geq 0, s \geq 0 \end{matrix}}$$

这个x和s的方程组有解当且仅当“primal objective大于等于t”

这个x和s的方程组无解当且仅当“primal objective小于t”

Strong LP duality

假设primal和dual LP都存在可行解, 则它们的最优解相等

证明(cont'd)： 由Farkas lemma

$$\begin{bmatrix} A & 0 \\ c^T & -1 \end{bmatrix} \begin{pmatrix} x \\ s \end{pmatrix} = \begin{pmatrix} b \\ t \end{pmatrix}$$
$$x \geq 0, s \geq 0$$

无解当且仅当存在 y, z 使得

$$(y^T \quad z) \begin{bmatrix} A & 0 \\ c^T & -1 \end{bmatrix} \geq 0$$
$$(y^T \quad z) \begin{pmatrix} b \\ t \end{pmatrix} < 0$$

即$y^T A + zc^T \geq 0,\ -z \geq 0, y^T b + zt < 0$

考虑$z = 0$，则有$y^T A \geq 0, y^T b < 0$，primal not feasible;

考虑$z \neq 0$，则有$\frac{1}{-z} y^T A \geq c^T, \frac{1}{-z} y^T b < t$，因此$\frac{1}{-z} y^T$为dual LP中的可行解，且目标函数值$< t$。

课程回顾

核心思想：

- 收敛性：能合理近似的算法，至少需要收敛
- 复杂性：计算复杂性
- 条件性 (鲁棒性)
  - 回顾条件数的定义
- 压缩性
  - SVD
- **正交性**
  - 正交多项式
  - 共轭

课程回顾

课程话题

- 插值与拟合
  - 方程求根：二分法，不动点迭代法，牛顿法
  - 拉格朗日插值：误差分析
  - Chebyshev插值
  - Chebyshev多项式：极值性质
  - 函数空间上的线性代数、不同的范数及比较
  - 最小二乘法：几何推导、微积分推导
  - 正交化过程：Gram-Schmidt
  - 正交系统中的最小二乘法：三角级数、函数逼近
  - 傅里叶变换：多项式的两种表示，正交性与FFT算法
- 特征值与随机游走（线性代数方法）
- 最优化与线性规划

课程回顾

课程话题

- 插值与拟合
- 特征值与随机游走（线性代数方法）
  - 求解线性方程组——一种特殊的方程求根问题
  - 解决线性方程组的通用方法—高斯消元法：pivoting
  - 矩阵的算子范数：作为线性映射对长度的改变
  - 线性方程组的条件数（比较：根的敏感性）
  - 解决线性方程组的迭代方法：Jacobi，Gauss-Seidel，Richardson，Conjugate-gradient
  - 分析线性迭代方程：谱半径
  - 特征值的min-max刻画（Courant-Fischer）
  - 矩阵的多项式：线性迭代方程，Cayley-Hamilton，逆矩阵的多项式，幂迭代
  - 计算特征值与特征向量、Singular value decomposition (SVD)
  - …
- 最优化与线性规划

课程回顾

课程话题

- 插值与拟合
- 特征值与随机游走（线性代数方法）
  - 图上的随机游走：概率转移方程（线性迭代），稳态分布（特征向量），mixing time（收敛速度）与spectral gap
  - 马尔可夫链：基本定理、周期性、不可约；Pagerank
  - 图的谱理论：拉普拉斯矩阵，特征值、特征向量与图的连通结构，图上随机游走的谱分析
  - 电阻电路网络：拉普拉斯线性方程组，等效电阻，电势能，等效电阻距离
  - hitting time：拉普拉斯线性方程组，cover time
  - Spectral embedding：使用特征向量进行“嵌入”
- 最优化与线性规划

课程回顾

课程话题

- 插值与拟合
- 特征值与随机游走（线性代数方法）
- 最优化与线性规划
  - 线性规划的不同形式
  - 整数线性规划与松弛化
  - LP的顶点：整数规划的最优解与松弛化的LP最优解
  - Duality(对偶性)：经济学解释，Min-max定理与对偶性，组合优化，零和博弈，随机算法分析
  - * 在线决策问题与Multiplicative weight update：LP与minimax定理

What we did not cover (in detail)

- Expander codes
  - Further reading: Essential coding theory by V. Guruswami, A. Rudra, M. Sudan
- Low rank matrix approximation (via SVD)
  - Compression
  - De-noising
  - Matrix completion
  - Further reading: Algorithmic Aspects of Machine Learning by Ankur Moitra; A simpler approach to matrix completion by Ben Recht
- Generalizations of Cheeger's inequality
  - Further reading: Eigenvalues and polynomials by Lap Chi Lau
- Graph sparsification
  - Further reading: Lx=b by Nisheeth Vishnoi
- Markov chain Monte Carlo and high dimensional integral
  - Further reading: Techniques in Optimization and Sampling by Yin Tat Lee and Santosh Vempala
  - The Markov Chain Monte Carlo Revolution by Persi Diaconis, and the references therein