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

Announcement

习题课：16周/周六/5-6节，逸C-115

对偶LP在经济学中的解释

设想我们为自己设计一份菜单

目标：满足每天营养需求、尽量的廉价

营养需求：500卡的蛋白质，100卡的碳水和400卡的脂肪

	肉类	主食	乳制品
蛋白质	500	50	300
碳水	0	300	100
脂肪	500	25	200
价格	**5**	**2**	**4**

假设每类食物都是无限可分的。

求解：这3类食物的一个组合，既满足每天的营养要求，同时越经济越好。

对偶LP在经济学中的解释

令 x: 肉类, y: 主食, z: 乳制品, 有如下的线性规划.

$$
\begin{aligned}
\min\ & 5x + 2y + 4z && \text{(花费)} \\
\text{s.t. } & 500x + 50y + 300z \geq 500 && \text{(每日蛋白质要求)} \\
& 0x + 300y + 100z \geq 100 && \text{(每日碳水要求)} \\
& 500x + 25y + 200z \geq 400 && \text{(每日脂肪要求)} \\
& x, y, z \geq 0.
\end{aligned}
$$

对偶规划如下.

$$
\begin{aligned}
\max\ & 500a + 100b + 400c \\
\text{s.t. } & 500a + 0b + 500c \leq 5 \\
& 50a + 300b + 25c \leq 2 \\
& 300a + 100b + 200c \leq 4 \\
& a, b, c \geq 0
\end{aligned}
$$

新产品达到肉类同样的营养的所需花费不应该超过肉类

否则客户只会直接购买肉类

对偶规划：设想一家未来的制药公司有蛋白质，碳水和脂肪三种药丸作为新产品。应该如何为这些药丸定价？

Earth mover’s distance (选讲)

“**愚公移山**”距离

考虑在直线上的一座“山”$\{p_i\}$，希望通过搬运，变成另外的形状$\{q_i\}$；直线上的距离为$\{\boldsymbol{d_{i,j}}\}$

求最小的搬运方案

$$\boldsymbol{min_{\vec{f}} \sum_i \sum_j f_{i,j} d_{i,j}}$$

subject to

$$\sum_i f_{i,j} \leq q_j, \forall j$$

$$\sum_j f_{i,j} \leq p_i, \forall i$$

$$\sum_i \sum_j f_{i,j} = \sum_i p_i = \sum_j q_j$$

Kantorovich(-Rubinstein) 对偶性:

在最优的(经过充分博弈的)传输方案之中，“自己搬运”的开销，与“最佳外包”的开销是一样的

Tolls for Multicommodity Flow (选讲)

交通道路或者是在网络中，如何最大化地利用现有的基础设施?

假设有n个用户，其中i 希望从 s_i 到 t_i 发送d_i 单位的数据

If we let the users to choose their own paths to send their information, they may all send along the shortest paths, but this may cause high congestions on some edges.

In a dual view: Instead of directly controlling their behaviors, what we could do is to give prices on each edge, and charge the users on the edges they used, and the hope is that it is possible to do it in a way to avoid congestion, and thus to achieve a better social welfare.

This may seem like a really difficult problem to solve.

But it turns out that if we write an LP to find a global optimal solution to minimize congestions, and then we look at the dual LP, the prices are just the dual variables on the edges!

See *Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games*

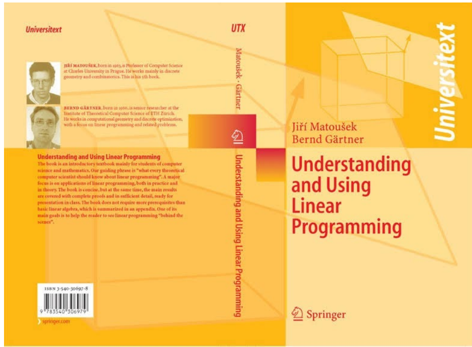
Strong duality theorem

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
LP的几何特性：凸性

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

   $$\|y - v\|_2^2 = \|x_* - v - t(x_* - x)\|_2^2$$
   $$= \|v - x_*\|_2^2 - 2t\langle x - x_*, v - x_*\rangle + t^2\|x - x_*\|_2^2$$

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

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$y^\top A \geq 0$ 且 $y^\top b < 0$

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

由于 $0 \in S$，因此$y^\top b < 0$

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

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

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

$$x\geq 0.$$

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

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

$$Ax=b$$

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

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

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

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

无解当且仅当存在 y, z 使得

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

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

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

考虑$z \neq 0$，则有$\frac{1}{-z} y^\top A \geq c^T, \frac{1}{-z} y^\top b < t$，因此$\frac{1}{-z} y^\top$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
  – 正交变换的特性、 QR分解、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/commute 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

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