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

回顾

上节课:

- 随机游走的碰撞时间(**Hitting time**)和返程时间(**commute time**)
- 线性规划 **LP**，顶点的不同定义

这节课:

- 对偶性
- 零和游戏

LP的顶点

考虑$\max\langle c, x\rangle$，约束在$P := \{Ax \le b\}$这个polytope内

顶点可以有3种等价的定义：

1. **<u>边角(corner)</u>**: 如果不存在$y \neq 0$ 使得 $x + y \in P$ and $x - y \in P$，则称点x是一个边角
2. **<u>极值点:</u>** 如果 $\exists c$ 使得x是该目标方向c的唯一最优解，则称点x是一个极值点.
3. **<u>基本解</u>**: 如果 $\langle A_i, x\rangle = b_i$ ，我们称第i个约束是<u>紧致 (tight)</u>的，其中 A_i 是第i行.

   对于给定的$x \in P$, 记 $A^=$ 为A 中关于 x 紧致的约束组成的子矩阵.

   如果$A^=$ 是满秩的, i.e. $rank(A^=) = n$，那么我们称x是一个基本解.

Simplex算法：从一个顶点开始；找下一个顶点，如果目标函数更优，则选择移到该顶点；重复；

邻居的选择: 最多(m-n)n

如果所有邻居都更差，则当前必定是最优的解：对于凸优化问题，局部最优即是全局最优

最坏情况下面， Simplex算法可能需要指数时间。但是实践中表现往往不错，smoothed analysis

多项式算法； Ellipsoid algorithm, interior point methods

Perfect bipartite matching

定理：考虑二分图完美匹配的线性规划：

max $\sum_{e\in E} c_e x_e$ subject to

$$\sum_{e\in\delta(v)} x_e = 1 \quad \forall v \in V$$

$$0 \le x_e \le 1 \quad \forall e \in E$$

该LP的最优解一定是整数解

注意：这个LP在一般图上面并不一定是整数解的

Ellipsoid algorithm, separation oracle (选讲)

最小生成树的一种LP写法

$$\max \sum_{e \in E} c_e x_e \quad \text{subject to}$$

- $\sum_{e \in E(S)} x_e \leq |S| - 1, \quad \forall S \subset V$
- $\sum_{e \in E(V)} x_e = |V| - 1$
- $0 \leq x_e \leq 1$

虽然是指数大小的LP，但是Ellipsoid algorithm只需要有separation oracle，也能多项式时间解出来

Duality：给目标函数证明上界

$$\max 5x_1 + 4x_2$$

$$2x_1 + x_2 \le 100$$

$$x_1 \le 30$$

$$x_2 \le 60$$

$$x_1, x_2 \ge 0.$$

- 如何证明目标函数的一个下界?

Duality：给目标函数证明上界

$$\max 5x_1 + 4x_2$$

$$2x_1 + x_2 \le 100 \quad (1)$$

$$x_1 \le 30 \quad (2)$$

$$x_2 \le 60 \quad (3)$$

$$x_1, x_2 \ge 0.$$

如何证明**目标函数的一个上界**?

- 考虑 $0\times(1) + 5\times(2) + 4\times(3)$
- 考虑 $3\times(1) + 0\times(2) + 1\times(3)$
- 考虑 $\frac{5}{2}\times(1) + 0\times(2) + \frac{3}{2}\times(3)$
- …
- 更一般地，考虑$y_1\times(1) + y_2\times(2) + y_3\times(3)$

Duality：给目标函数证明上界

$$\min 100y_1 + 30y_2 + 60y_3$$

$$2y_1 + y_2 \geq 5$$

$$y_1 + y_3 \geq 4$$

$$y_1, y_2, y_3 \geq 0.$$

$$5x_1 + 4x_2 \leq (2y_1 + y_2)x_1 + (y_1 + y_3)x_2 \leq 100y_1 + 30y_2 + 60y_3$$

- 拉格朗日乘数法
- 弱对偶性 (weak duality)

Primal-Dual

原始(Primal)问题：

$$\max_x \begin{pmatrix} \cdots & c^\top & \cdots \end{pmatrix} \begin{pmatrix} \vdots \\ x \\ \vdots \end{pmatrix}$$

$$\text{subject to} \begin{pmatrix} & \vdots & \\ \cdots & A & \cdots \\ & \vdots & \end{pmatrix} \begin{pmatrix} \vdots \\ x \\ \vdots \end{pmatrix} \leq \begin{pmatrix} \vdots \\ b \\ \vdots \end{pmatrix}$$

$$\begin{pmatrix} \vdots \\ x \\ \vdots \end{pmatrix} \geq 0$$

对偶(Dual)问题：

$$\min_y \begin{pmatrix} \cdots & b^\top & \cdots \end{pmatrix} \begin{pmatrix} \vdots \\ y \\ \vdots \end{pmatrix}$$

$$\text{subject to} \begin{pmatrix} & \vdots & \\ \cdots & A^\top & \cdots \\ & \vdots & \end{pmatrix} \begin{pmatrix} \vdots \\ y \\ \vdots \end{pmatrix} \geq \begin{pmatrix} \vdots \\ c \\ \vdots \end{pmatrix}$$

$$\begin{pmatrix} \vdots \\ y \\ \vdots \end{pmatrix} \geq 0$$

Weak duality

$$\max \ \langle c, x\rangle$$
$$Ax \le b$$
$$x \ge 0.$$

(Primal LP)

$$\max \ \langle c, x\rangle$$
$$\langle a_1, x\rangle \le b_1$$
$$\cdots$$
$$\langle a_m, x\rangle \le b_m$$
$$x \ge 0.$$

$$\min \ \langle b, y\rangle$$
$$\sum_{i=1}^{m} y_i a_i \ge c$$
$$y \ge 0.$$

$$\min \ \langle b, y\rangle$$
$$A^\top y \ge c$$
$$y \ge 0.$$

(Dual LP)

弱对偶性 (weak duality)定理: 对于任意原始的最大化LP中的可行解x，与对偶的最小化LP中的可行解y, 都有

$$\langle c, x\rangle \le \langle b, y\rangle$$

即，原始最大值≤对偶最小值.

证明：

$$\langle c, x\rangle = c^\top x \le (y^\top A)x = y^\top (Ax) \le y^\top b = \langle b, y\rangle.$$

Complementary Slackness Conditions

如何证明x是最优解？

- 弱对偶定理告诉我们，一个证明方法是：找到对偶LP中的y使得$\langle c, x\rangle = \langle b, y\rangle$

这样的y是否总是存在？

- 强对偶性定理 (strong duality)

什么时候可以取到等号？

- $c^\top x \le (y^\top A)x$**要取到等号**：如果$x_j > 0$, 则$\langle a^j, y\rangle = c_j$, 其中 a^j 是A的第j列.
- $y^\top(Ax) \le y^\top b$**要取到等号**：如果$y_i > 0$, 则$\langle a_i, x\rangle = b_i$, 其中 a_i 是A的第i行.

匈牙利算法

Min-max定理例子：组合优化

二分图中的最大匹配和最小顶点覆盖 (Kőnig's theorem)

$$\max \sum_{e \in E} x_e \qquad\qquad \min \sum_{v \in V} y_v$$

$$x(\delta(v)) \le 1 \quad \forall v \in V \qquad\qquad y_u + y_v \ge 1 \quad \forall e \in E$$

$$x_e \ge 0 \quad \forall e \in E \qquad\qquad x_v \ge 0 \quad \forall v \in V$$

整数解分别对应最大匹配和最小顶点覆盖

练习： 在二分图中，可以证明它们的最优解是整数解(integral)
拓展： 试推导组合中的Hall’s theorem

因此强对偶定理告诉我们

二分图中的最大匹配的大小 = 最小顶点覆盖的大小

Min-max定理例子：最大流最小割

$$\max f_{ts}$$

$$f\left(\delta^{in}(v)\right) - f\left(\delta^{out}(v)\right) \leq 0 \qquad \forall v \in V$$

$$f_e \leq 1 \qquad \forall e \in E$$

$$f_e \geq 0 \qquad \forall e \in E$$

$$\min \sum_{e \in E} d_e$$

$$d_{uv} + y_u - y_v \geq 0 \qquad \forall uv \in E$$

$$y_s - y_t \geq 1$$

$$y_v \geq 0 \qquad \forall v \in V$$

- **原始问题：** 最大流
- 对偶问题：最小割
- 可证明最优解是整数解

零和游戏

一个矩阵A

- “行玩家”选一行r
- “列玩家”选一列c

行玩家得到的回报是$A_{r,c}$，列玩家得到$-A_{r,c}$

	P	S	R
P	0	-1	1
S	1	0	-1
R	-1	1	0

行玩家的目标是最大化 $A_{r,c}$，而列玩家目标是最小化$A_{r,c}$

纳什均衡：即使一个玩家知道对方的策略之后，他/她也不能找到比当前策略严格更优的策略。

单纯的策略(pure strategy)：选一行/列

混合的策略(mixed strategy)：单纯策略的一个概率分布

Von-Neumann Minimax Theorem

给定$A \in \mathbb{R}^{m\times n}$
假设行玩家策略的概率分布是 $x \in \Delta^m$
列玩家策略的概率分布是$y \in \Delta^n$

那么它们玩下来的期望回报是$x^\top Ay$

<u>**Von-Neumann Minimax Theorem**</u>.

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

不管谁先宣布自己的策略，都会达到均衡解

Von-Neumann Minimax Theorem

<u>Von-Neumann Minimax Theorem</u>.

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

证明：

左边: 行玩家先选定 x

列玩家在给定$x^\top A$后，如何选择最优的y？

$x^\top A$为一行向量， 注意到y为概率分布

令 $(x^\top A)^{(j)}$为$x^\top A$的第 j列，则有

$$\min_j (x^\top A)^{(j)} \le x^\top Ay \le \max_j (x^\top A)^{(j)}$$

且等号可以取到。因此$\min_{y\in\Delta^n} x^\top Ay = \min_j (x^\top A)^{(j)}$

左边即可化简为$\max_{x\in\Delta^m}\min_j (x^\top A)^{(j)}$

Von-Neumann Minimax Theorem

<u>Von-Neumann Minimax Theorem</u>.

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

证明(cont'd)：

左边可化简为$\max_{x\in\Delta^m}\min_j (x^\top A)^{(j)}$，这可以通过LP表示

引入辅助变量 $t = \min_j (x^\top A)^{(j)}$

$$\max \quad t$$

$$(x^\top A)^{(j)} = \sum_{i=1}^{m} x_i a_{ij} \ge t \qquad \forall j = 1, \dots, n$$

$x \in \Delta^m$

$$\sum_{i=1}^{m} x_i = 1$$

$$x_i \ge 0 \qquad \forall i = 1, \dots, m$$

<u>Von-Neumann Minimax Theorem</u>.

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

证明(cont'd)：

类似地，右边可化简为$\min_{y\in\Delta^n}\max_i (Ay)_i$，进而通过LP表示

引入辅助变量 $r = \max_i (Ay)_i$

$$\min \quad r$$

$$(Ay)_i = \sum_{j=1}^{n} a_{ij}\, y_j \le r \qquad \forall i = 1, \dots, m$$

$y \in \Delta^n$

$$\sum_{i=1}^{n} y_i = 1$$

$$y_i \ge 0 \qquad \forall i = 1, \dots, n$$

Von-Neumann Minimax Theorem

<u>**Von-Neumann Minimax Theorem**</u>.

$$\max_{x\in\Delta^m}\min_{y\in\Delta^n} x^\top A y = \min_{y\in\Delta^n}\max_{x\in\Delta^m} x^\top A y$$

证明(cont'd)：把它们都转化为<u>**标准形式**</u>

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

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

Equality follows directly from strong duality!

Yao’s Minimax Principle

一个随机算法的最坏运行时间，是它在所有输入实例中期望运行时间的最大值

试考虑这样的一个零和游戏：

“算法”玩家需要从不同的算法中选择，目标是最小化运行时间

“攻击”玩家需要从不同的输入实例中选择，目标是最大化运行时间

“**算法**”的**混合策略**：随机算法

“攻击”的混合策略：输入实例的概率分布

要成功地“攻击”所有的随机算法，等价的问题是：

给出一个输入实例的概率分布，使得所有的确定性算法的期望运行时间都比较大

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$\{q_i\}$
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*