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

零和游戏

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

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$-A_{r,c}$

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$x^T Ay$

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

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

不管谁先宣布自己的策略，都会达到均衡解

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:

左边: 行玩家先选定 x

列玩家在给定$x^T A$后，如何选择最优的y？

$x^T A$为一行向量， 注意到y为概率分布

令 $(x^T A)^{(j)}$为$x^T A$的第 j列，则有

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

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

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

Von-Neumann Minimax Theorem

Von-Neumann Minimax Theorem.

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

证明(cont'd)：

左边可化简为$\max\limits_{x\in\Delta^m}\min\limits_{j}(x^T A)^{(j)}$，这可以通过LP表示

引入辅助变量 $t = \min\limits_{j}(x^T A)^{(j)}$

$x \in \Delta^m$

$$\max \quad t$$

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

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

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

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)：

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 $r = \max_i (Ay)_i$

$y \in \Delta^n$

$$\min \quad r$$

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

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

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

Von-Neumann Minimax Theorem

Von-Neumann Minimax Theorem.

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

证明(cont'd)：把它们都转化为标准形式

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

$\boxed{\times\, y_j}$ $\boxed{\times\, r}$

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

$\boxed{\times\, x_i}$ $\boxed{\times\, t}$

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

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

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

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

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

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*

A detour: Online decision problems

Imagine playing a repeated game against an environment

Paging algorithms: Virtual memory/page table management

Caching: cache eviction algorithms in CPU

Compiling codes: register allocations

Networking: routing algorithms

Virtual machine/runtime environment: garbage collection

Database systems: B-tree

Scheduling jobs

Resource management

Online Expert model

	1	2	3		n
1	1	1	0		
2	1	0	1		
3	0	0	1		
T					

更一般地，每次决策之后的收益可能都不一样

- Hedge algorithm，对冲/避险
- Boosting
- Fictitious play
- Multiplicative weight update method

- Applications in zero-sum game and solving LP

Online Expert model

	1	2	3		n
1	1	1	0		
2	1	0	1		
3	0	0	1		
T					

假设有n个“专家”

- 天气预报
- 股市走势预测
- 资源调度
- 进化中的基因

第t天，专家i可能会出错，导致损失$l_i^{(t)} \in \{0,1\}$

问：该听信哪个专家？如何比较不同专家的表现？

如果在第t天，选择专家i，损失$l_i^{(t)}$；
在T天之后，

总的遗憾(Regret)=我们的选择导致的损失 – 表现最好的专家导致的损失

Online Expert model

	1	2	3		n
1	1	1	0		
2	1	0	1		
3	0	0	1		
T					

假设有n个“专家”（天气预报，股市走势预测，资源调度，进化中的基因）
第t天，专家i可能会出错，导致损失$l_i^{(t)} \in \{0,1\}$

问：该听信哪个专家？
如果在第t天，选择专家i，损失$l_i^{(t)}$；
在T天之后，

总的遗憾(Regret)=我们的选择导致的损失 – 表现最好的专家导致的损失

直观上，想要找“最少犯错的”

目前从未犯过错的专家中，选择一个听取？

- 注意：我们并不假设过去的表现可以预测未来的表现
- 专家的预测可能是任意的(adversarial)，或者是恶意的(malicious)
- 对手的策略：观察你听信哪个专家，然后去收买
- 犯错次数：n-1

Online Expert model

假设有n个专家（天气预报，股市走势预测）

第t天，专家i可能会出错，导致损失$l_i^{(t)} \in \{0,1\}$

直观上，想要找“最少犯错的”

听取目前从未犯过错的专家中，多数派的意见？

- 犯错次数$\leq \log n$
- 每当我们听信的意见错了，意味着“从未犯过错的专家”这个集合的大小要减半了
- 如果至少存在一个完美的专家，至多$\log n$次就能找到

Online Expert model

假设有n个专家（天气预报，股市走势预测）
第t天，专家i可能会出错，导致损失$l_i^{(t)} \in \{0,1\}$

直观上，想要找“最少犯错的”
如果没有完美的专家，所有专家迟早都会犯错呢？

从所有专家中，加权多数派的意见？
- 最开始，每个专家有$w_i^{(t)} = 1$票
- **听取**加权多数派的意见
- 如果专家i出错了，$w_i^{(t+1)} = w_i^{(t)}/2$

为了简单起见，这里假设每一轮都只有两派意见

Online Expert model

从所有的专家中，加权多数派的意见？

- 最开始，每个专家有$w_i^{(1)}=1$，“票数”
- 听取加权多数派的意见
- 如果专家i出错了，$w_i^{(t+1)}=w_i^{(t)}/2$

分析：

- 假设表现最好的专家会犯错m次
- 考虑$\Phi^{(t)}=\sum_{i=1}^{n} w_i^{(t)}$，注意到$\Phi^{(1)}=n$
- 表现最好的专家最后的“票数”$w_i\geq\frac{1}{2^m}$
- 每当我们听信的意见错了：$\Phi^{(t+1)}\leq\Phi^{(t)}\cdot\frac{3}{4}$
- 如果我们一共犯错的次数为M，则有

$$\frac{1}{2^m}\leq\sum_i w_i\leq\left(\frac{3}{4}\right)^M n$$

对$\frac{1}{2^m}\leq\left(\frac{3}{4}\right)^M n$取对数可得：

$$M\leq\frac{m+\log n}{\log\frac{4}{3}}\leq 2.41(m+\log n)$$

如果不是票数减半，而是乘上$(1-\epsilon)$，类似可得$(1-\epsilon)^{\mathrm{m}}\leq\left(1-\frac{\epsilon}{2}\right)^M n$，进而$M\leq 2(1+\epsilon)m+\frac{2\ln n}{\epsilon}$

Online Expert model

从所有的专家中，加权多数派的意见？

如果不是票数减半，而是乘上$(1-\epsilon)$，类似可得$(1-\epsilon)^m \le \left(1-\frac{\epsilon}{2}\right)^M n$，进而$M \le 2(1+\epsilon)m + \frac{2\ln n}{\epsilon}$

依然是接近双倍的犯错次数！而且这是确实会发生的，考虑两个专家A和B：

- A在偶数天是正确的，B在奇数天是正确的
- A在前面一半的时间是正确的，B在后面一半的时间是正确的

最佳的专家：都是T/2次犯错
而我们的算法在前面的情形下面会犯T-1次错误，双倍的犯错

Online Expert model

从所有的专家中，加权多数派的意见？

如果不是票数减半，而是乘上$(1-\epsilon)$，类似可得$(1-\epsilon)^m \le \left(1-\frac{\epsilon}{2}\right)^M n$，进而$M \le 2(1+\epsilon)m + \frac{2\ln n}{\epsilon}$

依然是接近双倍的犯错次数！而且这是确实会发生的，考虑两个专家A和B：

- A在偶数天是正确的，B在奇数天是正确的
- A在前面一半的时间是正确的，B在后面一半的时间是正确的

更好的策略：随机化/混合策略，而不是单纯的策略
以概率$\propto w_i^{(t)}$**选择相信专家** i
考虑同样的第一种情形，会以接近1/2的概率选择相信专家A或者专家B
我们的期望犯错次数也将是 T/2次，接近最优的专家！

对于股票的例子中，也可以想象持有上市公司i的股票$\propto w_i^{(t)}$

Multiplicative weight update

MWU算法

固定参数$\eta \leq 1/2$. 令$w_i^{(1)} = 1, \forall 1 \leq i \leq n$.

For $1 \leq t \leq T$ do

1. 记$\Phi^{(t)} = \sum_{i=1}^{n} w_i^{(t)}$. 令 $p_i^{(t)} = w_i^{(t)} / \Phi^{(t)}$.
2. 以概率 $p_i^{(t)}$听取专家 i 的意见.
3. 得到当天专家们的表现 $l^{(t)}$.
4. 更新 $w_i^{(t+1)} \coloneqq \left(1 - \eta l_i^{(t)}\right) \cdot w_i^{(t)}$.

MWU Theorem

更新方法：$w_i^{(t+1)} \coloneqq \left(1-\eta l_i^{(t)}\right) \cdot w_i^{(t)}$.

MWU定理：假设$l_i^{(t)} \in [-1,+1]$和$\eta \le 1/2$. 在T天之后，MWU算法满足$\forall i$

$$\text{期望损失} = \sum_{t=1}^{T} \langle l^{(t)}, p^{(t)} \rangle \le \sum_{t=1}^{T} l_i^{(t)} + \eta \sum_{t=1}^{T} \left| l_i^{(t)} \right| + \frac{\ln n}{\eta}.$$

即总的遗憾(regret) $\le \eta T + \frac{\ln n}{\eta}$. 取 $\eta = \sqrt{\frac{\ln n}{T}}$则有 总的遗憾(regret) $\le O\left(\sqrt{T \ \ln n}\right)$.

换言之，平均每天的遗憾为$O\left(\sqrt{\frac{\ln n}{T}}\right)$

MWU Analysis（选讲）

更新方法：$w_i^{(t+1)} \coloneqq \left(1-\eta l_i^{(t)}\right)\cdot w_i^{(t)}$.

MWU定理： 假设$l_i^{(t)} \in [-1,+1]$ 和 $\eta \le 1/2$. 在T天之后，MWU算法满足$\forall i$

$$\text{期望损失} = \sum_{t=1}^{T}\left\langle l^{(t)}, p^{(t)}\right\rangle \le \sum_{t=1}^{T} l_i^{(t)} + \eta \sum_{t=1}^{T}\left|l_i^{(t)}\right| + \frac{\ln n}{\eta}.$$

证明：与此前类似，我们将会给$\Phi^{(t)}$一个上界和下界的分析。

对于上界：

$$\Phi^{(t+1)} = \sum_{i=1}^{n} w_i^{(t+1)} = \sum_{i=1}^{n} w_i^{(t)} - \eta\left\langle l^{(t)}, w^{(t)}\right\rangle = \Phi^{(t)}\left(1-\eta\left\langle l^{(t)}, p^{(t)}\right\rangle\right) \le \Phi^{(t)} e^{-\eta\left\langle l^{(t)}, p^{(t)}\right\rangle}$$

(to be cont'd..)

MWU Analysis （选讲）

更新方法：$w_i^{(t+1)} \coloneqq \left(1-\eta l_i^{(t)}\right)\cdot w_i^{(t)}$.

MWU定理： 假设$l_i^{(t)} \in [-1,+1]$ 和 $\eta \leq 1/2$. 在T天之后，MWU算法的期望损失满足$\forall i$

$$\sum_{t=1}^{T}\langle l^{(t)}, p^{(t)}\rangle \leq \sum_{t=1}^{T} l_i^{(t)} + \eta \sum_{t=1}^{T}\left|l_i^{(t)}\right| + \frac{\ln n}{\eta}.$$

证明(cont'd)：对于下界：

$$\Phi^{(T+1)} = \sum_{i=1}^{n} w_i^{(T+1)} \geq w_i^{(T+1)} = \prod_{t=1}^{T}\left(1-\eta l_i^{(t)}\right) \geq (1-\eta)^{\sum_{\geq 0} l_i^{(t)}}(1+\eta)^{-\sum_{<0} l_i^{(t)}}$$

这里使用到$1-\eta x \geq \begin{cases}(1-\eta)^x, & x\in[0,1]\\ (1+\eta)^x, & x\in[-1,0]\end{cases}$

合在一起有 $ne^{-\eta\sum_{t=1}^{T}\langle l^{(t)},p^{(t)}\rangle} \geq \Phi^{(T+1)} \geq (1-\eta)^{\sum_{\geq 0} l_i^{(t)}}(1+\eta)^{-\sum_{<0} l_i^{(t)}}$

(to be cont'd..)

MWU Analysis （选讲）

更新方法：$w_i^{(t+1)} := \left(1 - \eta l_i^{(t)}\right) \cdot w_i^{(t)}$.

MWU定理： 假设$l_i^{(t)} \in [-1, +1]$ 和 $\eta \leq 1/2$. 在T天之后，MWU算法的期望损失满足$\forall i$

$$\sum_{t=1}^{T} \langle l^{(t)}, p^{(t)} \rangle \leq \sum_{t=1}^{T} l_i^{(t)} + \eta \sum_{t=1}^{T} \left| l_i^{(t)} \right| + \frac{\ln n}{\eta}.$$

证明(cont'd)：对以下的取对数： $n e^{-\eta \sum_{t=1}^{T} \langle l^{(t)}, p^{(t)} \rangle} \geq (1-\eta)^{\sum_{\geq 0} l_i^{(t)}} (1+\eta)^{-\sum_{<0} l_i^{(t)}}$， 即有

$$\ln n - \eta \sum_{t=1}^{T} \langle l^{(t)}, p^{(t)} \rangle \geq \ln(1-\eta) \sum_{\geq 0} l_i^{(t)} - \ln(1+\eta) \sum_{<0} l_i^{(t)} \geq (-\eta - \eta^2) \sum_{\geq 0} l_i^{(t)} - (\eta - \eta^2) \sum_{<0} l_i^{(t)}$$

最后的不等式可以通过对 $\ln(1-\eta)$进行Taylor展开得到（假设$\eta \leq 1/2$）

化简得到 $\eta \sum_{t=1}^{T} \langle l^{(t)}, p^{(t)} \rangle \leq \ln n + \eta \sum_{t=1}^{T} l_i^{(t)} + \eta^2 \sum_{t=1}^{T} \left| l_i^{(t)} \right|$， 两边除以$\eta$即得到结论

Online Expert model

一些评论：

- **我们考虑的是一个简化的模型**：专家们只提供最终的意见，没有论证过程；决策者也只看专家们的意见，而没有利用自己的专业判断；
- **进行对比的**，是专家库里面表现最好的专家；“股市有风险，入市需谨慎”
- 假设中每天的损失$l_i^{(t)}$**都是有界的**，否则某一天的损失，就可能导致前功尽弃
- 通过引入随机化/混合的策略，起到了类似“不测之威”的效果；但并不是所有场合都适合引入随机化的

- 在“听取加权平均的意见”算法之中，我们假设了每天都只有两种不同的意见（这并不意味着一共只有两种不同的意见）；更一般地，也可以假设，每天有最多k种不同的意见，但是对应的近似比会随着k而增大（练习：找出这个近似比）
- 注意：MWU的算法之中，并不需要上述的假设！专家们可以有完全不一样的意见。

MWU vs. LP （选讲）

假设给定线性规划：$\min\langle c,x\rangle\ s.t.\ \ Ax\geq b, x\geq 0$，其中 $A\in\mathbb{R}^{m\times n}, c,x\in\mathbb{R}^n, b\in\mathbb{R}^m$.

我们可以通过二分搜索转化成一个约束可满足问题：$Ax\geq b, x\geq 0$

考虑满足所有约束的一个加权平均：$p^{(t)}Ax^{(t)}\geq p^{(t)}b$

- 最开始, $p^{(1)}=\vec{1}/m$.
- 在第t轮, 找$x^{(t)}$ 使得$p^{(t)}Ax^{(t)}\geq p^{(t)}b$ 并且 $x\geq 0$.
- 如果$x^{(t)}$ 不存在, 则LP不可解. （可以证明，不可解当且仅当存在$p^{(t)}A\leq 0$且 $p^{(t)}b>0$ ）
- 一般来说，$x^{(t)}$ 可能满足一些约束，违反另一些约束.
- 每个约束可以看成是一个“专家”，它预测导致的损失$l_i^{(t)}=A_ix^{(t)}-b_i$
- 可以用MWU来更新$p^{(t+1)}$使得更多的权重被放到那些被违反了的约束.
- 因此当$x^{(t+1)}$ 满足 $p^{(t+1)}Ax^{(t+1)}\geq p^{(t+1)}b$，它更关注那些之前被违反了的约束
- 不断重复，最后返回$\bar{x}=\frac{1}{T}\sum_{t=1}^{T}x^{(t)}$

直观的来说，我们把每个约束看成“专家”。约束被违反得越多的话，我们认为这个专家表现得越好，就越想要选择这样的约束。

MWU保证了我们平均的损失，和被违反得最多的约束相比，是趋向于零的。

MWU vs. Zero sum game

回顾Minimax定理：

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

左边：x先手，y后手；右边：y先手，x后手

后手的优势(LP弱对偶定理)说的是

$$\max_{x\in\Delta^m}\min_{y\in\Delta^n} x^T A y \le \min_{y\in\Delta^n}\max_{x\in\Delta^m} x^T A y$$

另一个方向可以通过几何证明（Farkas' lemma），博弈论或者说是拓扑的证明（不动点定理），或者**LP的强对偶形式**（上节课）证明

这节课：通过MWU定理的给出的算法证明

MWU vs. Zero sum game

假设两个玩家不断地玩同一个零和游戏，但是每一轮都会以MWU来更新自己的策略

记$x^{(t)}$和$y^{(t)}$为他们第t轮的（混合）策略
在第t轮结束的时候，行玩家会得知自己的损失向量是$-Ay^{(t)}$，列玩家也会得知自己的损失向量$A^T x^{(t)}$
接下来，他们用MWU来更新得到$x^{(t+1)}$和$y^{(t+1)}$

我们考虑在T轮之后，他们平均的损失
为了记号的方便，下面开始我们假设$x^{(t)}$是行向量的形式， $y^{(t)}$为列向量的形式

由MWU定理可得：

$$\sum_{t=1}^{T} x^{(t)}(-A)y^{(t)} \le \min_{x\in\Delta} \sum_{t=1}^{T} x(-A)y^{(t)} + O\left(\sqrt{T \ln n}\right)$$

$$\sum_{t=1}^{T} x^{(t)} A y^{(t)} \le \min_{y\in\Delta} \sum_{t=1}^{T} x^{(t)} A y + O\left(\sqrt{T \ln n}\right)$$

合并，并记$\bar{x} = \frac{1}{T}\sum_{t=1}^{T} x^{(t)}$， $\bar{y} = \frac{1}{T}\sum_{t=1}^{T} y^{(t)}$ 可得：

$$\max_{x\in\Delta} xA\bar{y} \le \min_{y\in\Delta} \bar{x}Ay + O\left(\sqrt{\frac{\ln n}{T}}\right)$$

因此也有

$$\min_{y\in\Delta^n} \max_{x\in\Delta^m} xAy \le \max_{x\in\Delta^m} \min_{y\in\Delta^n} xAy + O\left(\sqrt{\frac{\ln n}{T}}\right)$$

因为不等式对于任意T成立，因此令T趋向正无穷有$\min_{y\in\Delta^n} \max_{x\in\Delta^m} xAy \le \max_{x\in\Delta^m} \min_{y\in\Delta^n} xAy$，正是minimax定理的另外一个方向