Matrix games

Any matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ defines a game, and we call $\mathbf{A}$ the *payout matrix*. The row player is *Alice* and each row is called a *pure strategy* for Alice. The column player is *Bob* and each column is a *pure strategy* for Bob. If Alice plays pure strategy i and Bob play pure strategy j, Alice receives $a_{i,j}$ dollars from Bob. If $a_{i,j} < 0$, then this means that Alice pays Bob $-a_{i,j}$ dollars.

A vector $\mathbf{y}$ is *stochastic* if $\mathbf{y} \geq \mathbf{0}$ and $\sum_{i=1}^{m} y_i = 1$. Throughout, assume $\mathbf{x}, \mathbf{y}$ are stochastic vectors in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

Let $\mathbf{A} = \{a_{ij}\}$ be an $m \times n$ payoff matrix. Any $\mathbf{y}$ defines the *mixed strategy* for the row player, Alice, where she chooses each pure strategy i with probability y_i, and any $\mathbf{x}$ defines a *mixed strategy* for the column player, Bob, where he chooses each pure strategy j with probability x_j. If Alice plays mixed strategy $\mathbf{y}$ and Bob plays mixed strategy $\mathbf{x}$, then the *expected payout* is

$$\sum_{i=1}^{m}\sum_{j=1}^{n} y_i a_{ij} x_j = \mathbf{y}^T A\mathbf{x}.$$

We say that stochastic $\tilde{\mathbf{y}} \in \mathbb{R}^m$ or stochastic $\tilde{\mathbf{x}} \in \mathbb{R}^n$ is *optimal* if it maximizes $\min_{\mathbf{x}} \tilde{\mathbf{y}}^T A\mathbf{x}$ or minimizes $\max_{\mathbf{y}} \mathbf{y}^T A\tilde{\mathbf{x}}$, respectively. It is not hard to see that for any stochastic $\bar{\mathbf{y}} \in \mathbb{R}^m$ and stochastic $\bar{\mathbf{x}} \in \mathbb{R}^n$,

$$\min_{\mathbf{x}} \bar{\mathbf{y}}^T A\mathbf{x} \leq \mathbf{y}^T A\mathbf{x} \leq \max_{\mathbf{y}} y^T A\bar{\mathbf{x}}. \tag{1}$$

Theorem (The Minimax Theorem)**.** *For every $A \in \mathbb{R}^{m\times n}$ there exists stochastic $\tilde{\mathbf{y}} \in \mathbb{R}^m$ and stochastic $\tilde{\mathbf{x}} \in \mathbb{R}^n$ such that*

$$\min_{\mathbf{x}} \tilde{\mathbf{y}}^T A\mathbf{x} = \tilde{\mathbf{y}}^T A\tilde{\mathbf{x}} = \max_{\mathbf{y}} \mathbf{y}^T A\tilde{\mathbf{x}}.$$

Note that, using (1), it can be shown that the mixed strategies $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{x}}$ from the theorem are optimal. With $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{x}}$ from the theorem, the number $\tilde{\mathbf{y}}^T A\tilde{\mathbf{x}}$ is called the *value* of the game. We first prove the following proposition ($\mathbf{e_i}$ is the ith standard basis vector).

Proposition. *For any $\mathbf{A} \in \mathbb{R}^{m\times n}$ and stochastic $\mathbf{y} \in \mathbb{R}^m$,*

$$\min_{\mathbf{x}} \mathbf{y}^T A\mathbf{x} = \min_{j} \mathbf{y}^T A\mathbf{e_j} = \min_{j} \sum_{i=1}^{m} y_i a_{i,j}.$$

For any stochastic $\mathbf{x} \in \mathbb{R}^n$,

$$\max_{\mathbf{y}} \mathbf{y}^T A\mathbf{x} = \max_{i} \mathbf{e_i}^T A\mathbf{x} = \max_{i} \sum_{j=1}^{n} a_{i,j} x_j.$$

Proof. Clearly, $\min_j \mathbf{y}^T A\mathbf{e_j} = \min_j \sum_{i=1}^{m} y_i a_{i,j}$ and $\min_{\mathbf{x}} \mathbf{y}^T A\mathbf{x} \leq \min_j \mathbf{y}^T A\mathbf{e_j}$ since every standard basis vector is stochastic. So we now only need to show $\min_{\mathbf{x}} \mathbf{y}^T A\mathbf{x} \geq \min_j \sum_{i=1}^{m} y_i a_{i,j}$ to prove the first sentence. Let $t = \min_j \sum_{i=1}^{m} y_i a_{i,j}$. For any stochastic $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{y}^T A\mathbf{x} = \sum_{j=1}^{n}\sum_{i=1}^{m} y_i a_{i,j} x_j = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} y_i a_{i,j}\right) \geq \left(\sum_{j=1}^{n} x_j\right) \cdot t = t.$$

This implies that $\min_{\mathbf{x}} \mathbf{y}^T A\mathbf{x} \geq t$, which proves the first sentence. The proof of the second sentence is similar. $\square$

Proof of the Minimax Theorem. We want to find an optimal $\mathbf{x}$ which, by definition, is a stochastic vector $\mathbf{x}$ that minimizes $\max_{\mathbf{y}} \mathbf{y}^T \mathbf{A}\mathbf{x}$. By the proposition, this is equivalent finding a vector $\mathbf{x}$ that minimizes $\max_i \sum_{j=1}^{m} a_{i,j} x_j$ subject to $x_1 + \cdots + x_n = 1$ and $\mathbf{x} \geq \mathbf{0}$. By adding a variable w, we can write an LP whose solutions gives us the optimal $\mathbf{x}$ as follows,

$$
\begin{array}{lll}
\text{Minimize} & w & \\
\text{subject to} & x_1 + \cdots + x_n = 1 & \\
 & w \geq \sum_{j=1}^{n} a_{i,j} x_j & \text{for all } i \in [m] \\
 & \mathbf{x} \geq \mathbf{0} &
\end{array}
\tag{2}
$$

Similarly, we can write an LP whose solutions gives us an optimal $\mathbf{y}$ as follows

$$
\begin{array}{lll}
\text{Maximize} & z & \\
\text{subject to} & y_1 + \cdots + y_m = 1 & \\
 & z \leq \sum_{i=1}^{m} a_{i,j} y_i & \text{for all } j \in [n] \\
 & \mathbf{y} \geq \mathbf{0} &
\end{array}
\tag{3}
$$

It is a good exercise to show that (2) and (3) are dual linear programs (and this was done in class). Since both linear program are feasible, this observation and strong duality give us the theorem. $\square$