Stat 155 Lecture 8 Notes

Daniel Raban

February 13, 2018

1 Nash Equilibria, Linear Programming, and von Neumann's Minimax Theorem

1.1 Nash equilibria

1.1.1 Optimality of Nash equilibria

Definition 1.1. A pair $(x^*, y^*) \in \Delta_m \times \Delta_n$ is a *Nash equilibrium* for a payoff matrix $A \in \mathbb{R}^{m \times n}$ if

$$\max_{x \in \Delta_m} x^\top A y^* = (x^*)^\top A y^* = \min_{y \in \Delta_n} (x^*)^\top A y.$$

Think of these as locally optimal strategies. If Player 1 plays x^* and Player 2 plays y^*, neither player has an incentive to change. Given a pair of safety strategies, we can get a Nash equilibrium, but a Nash equilibrium is a priori not necessarily a pair of safety strategies. The difference is that we do not require $(x^*)^\top A y^*$ to be the value of the game. However, these are actually globally optimal strategies, as well.

Theorem 1.1. *The pair (x^*, y^*) is a Nash equilibrium iff x^* and y^* are optimal.*

Proof. ($\Longrightarrow$) This is the same as the proof for the optimality of a saddle point.

$$\begin{aligned}
\min_{y \in \Delta_n} \max_{x \in \Delta_m} x^\top A y &\geq \max_{x \in \Delta_m} \min_{y \in \Delta_n} x^\top A y \\
&\geq \min_{y \in \Delta_n} (x^*)^\top A y \\
&= (x^*)^\top A y^* \\
&= \max_{x \in \Delta_m} x^\top A y^* \\
&\geq \min_{y \in \Delta_n} \max_{x \in \Delta_m} x^\top A y.
\end{aligned}$$

($\Longleftarrow$) The von Neumann minimax theorem implies that

$$(x^*)^\top A y^* \geq \min_y (x^*)^\top A y$$

$$
\begin{aligned}
&= \max_x \min_y x^\top A y \\
&= \min_y \max_x x^\top A y \\
&= \max_x x^\top A y^* \\
&\geq (x^*)^\top A y^*. \qquad \square
\end{aligned}
$$

1.1.2 Indifference and Nash Equilibria

Assume that

$$
(x^*)^\top A = (a, \ldots, a), \qquad \begin{pmatrix} a \\ \vdots \\ a \end{pmatrix} = Ay^*
$$

for some constant a. Then

$$
\min_y (x^*)^\top A y = a = (x^*)^\top A y^* = \max_x x^\top A y^*,
$$

so (x^*, y^*) is a Nash equilibrium. So x^* an y^* are optimal.

1.2 Solving zero-sum games using matrix inversion

Here is a useful theorem that is a consequence of the principle of indifference. You can find the proof in the Ferguson book.

Theorem 1.2. *Suppose the square matrix A is nonsingular and $\mathbf{1}^\top A^{-1}\mathbf{1} \neq 0$. Then the game with matrix A has value $V = (\mathbf{1}^\top A^{-1}\mathbf{1})^{-1}$ and optimal strategies $(x^*)^\top = V\mathbf{1}^\top A^{-1}$ and $y^* = VA^{-1}\mathbf{1}$, provided both $x^* \geq 0$ and $y^* \geq 0$.*

Example 1.1. Let $A \in \mathbb{R}^{3\times 3}$ be

$$
A = \begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}
$$

with each $a_{i,i} > 0$. Using the theorem, we get

$$
\begin{aligned}
V &= (\mathbf{1}^\top A^{-1}\mathbf{1})^{-1} \\
&= \left[(1,1,1) \begin{pmatrix} 1/a_{1,1} & 0 & 0 \\ 0 & 1/a_{2,2} & 0 \\ 0 & 0 & 1/a_{3,3} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right]^{-1} \\
&= \frac{1}{1/a_{1,1} + 1/a_{2,2} + 1/a_{3,3}}.
\end{aligned}
$$

We also get

$$\begin{aligned}
(x^*)^\top &= V\mathbf{1}^\top A^{-1} \\
&= V(1,1,1)\begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix} \\
&= \frac{1}{1/a_{1,1} + 1/a_{2,2} + 1/a_{3,3}}(1/a_{1,1}, 1/a_{2,2}, 1/a_{3,3}),
\end{aligned}$$

$$\begin{aligned}
y^* &= VA^{-1}\mathbf{1} \\
&= V\begin{pmatrix} a_{1,1} & 0 & 0 \\ 0 & a_{2,2} & 0 \\ 0 & 0 & a_{3,3} \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
&= \frac{1}{1/a_{1,1} + 1/a_{2,2} + 1/a_{3,3}}(1/a_{1,1}, 1/a_{2,2}, 1/a_{3,3}).
\end{aligned}$$

1.3 Linear programming: an aside

Definition 1.2. A *linear program* is an optimization problem involving the choice of a real vector to maximize a linear objective subject to linear constraints:

$$\max_{x\in\mathbb{R}^n} x^\top b \quad \text{such that} \quad \begin{aligned} &d_1^\top \le c_1 \\ &\vdots \\ &d_k^\top \le c_k. \end{aligned}$$

Here, $b \in \mathbb{R}^n$ specifies the *linear objective* $x \to b^\top x$, and $d_i \in \mathbb{R}^n$ and $c_i \in \mathbb{R}$ specify the i-th constraint.

The set of values x that satisfy the constraints is a polytope (an intersection of half spaces). From the perspective of the row player, a two player zero-sum game is an optimization problem of the form

$$\max_{x\in\mathbb{R}^n} \min_{i\in\{1,\dots,n\}} x^\top Ae_i \quad \text{such that} \quad \begin{aligned} &x_1^\top \le 0 \\ &\vdots \\ &x_k^\top \le 0 \\ &\mathbf{1}^\top x = 1. \end{aligned}$$

This is not a linear program; the constrants are linear, but hte objective is not. But we can convert it to a linear program by introducting the slack variable $Z = \min_{i\in\{1,\dots,n\}} x^\top Ae_i$.

There are efficient (polynomial time) algorithms for solving linear programs. The column player's linear program is the dual of the row player's linear program. In fact, for any concave maximization problem, like the row player's linear program (we'll call it the *primal problem*), it is possible to define a dual convex minimization problem, like the column player's linear program. This dual problem has a value that is at least as large the value of the primal problem.

In many important cases (such as our linear program), these values are the same. In optimization, this is called strong duality. This is von Neumann's minimax theorem. The principle of indifference is a general property of dual optimization problems (called *complementary duality*).

1.4 Proof of von Neumann's minimax theorem

We want to prove the following theorem:

Theorem 1.3. *For any two-person zero-sum game with payoff matrix $A \in \mathbb{R}^{m\times n}$,*

$$\min_{y\in\Delta_n} \max_{x\in\Delta_m} x^\top Ay = \max_{x\in\Delta_m} \min_{y\in\Delta_n} x^\top Ay.$$

The textbook proves this theorem using the separating hyperplane theorem. We will prove this theorem in a more algorithmic way, developing an optimal strategy by learning from the other player's optimal moves against ours.

Consider a two-player zero-sum game that is repeated for T rounds. At teach round, the row player chooses an $x_t \in \Delta_m$. Then the columns player chooses a $y_t \in \Delta_n$, and the row player receives a payoff of $x_t^\top Ay_t$.

The row player's regret after T rounds is how much its total payoff falls short of the best in retrospect that it could have achieved against the column player's choices with a fixed mixed strategy:

$$R_T = \max_{x\in\Delta_m} \sum_{t=1}^{T} x^\top Ay_t - \sum_{t=1}^{T} x_t^\top Ay_t.$$

We will see that there are learning algorithms that have low regret against any sequence played by the column player. These learning algorithms don't need to know anything about the game in advance; they just need to see, after each round, the column vector of payoffs corresponding to the column player's choice.

Lemma 1.1. *The existence of a row player with low regret ($R_T/T \to 0$ as $T \to \infty$) implies the minimax theorem.*

Proof. Define $\bar{x} = T^{-1}\sum_{t=1}^{T} x_t$. Suppose that the column player plays a best response y_t against the row player's choice x_t:

$$x_t^\top Ay_t = \min_{y\in\Delta_n} x_t^\top Ay.$$

Define $\bar{y} = T^{-1} \sum_{t=1}^{T} y_t$. We then have

$$\begin{aligned}
\max_{x \in \Delta_m} \min_{y \in \Delta_n} x^\top A y &\geq \min_{y \in \Delta_n} \bar{x}^\top A y \\
&= \min_{y \in \Delta_n} \frac{1}{T} \sum_{t=1}^{T} x_t^\top A y \\
&\geq \frac{1}{T} \sum_{t=1}^{T} \min_{y \in \Delta_n} x_t^\top A y \\
&= \frac{1}{T} \sum_{t=1}^{T} x_t^\top A y_t \\
&= \max_{x \in \Delta_m} \frac{1}{T} \sum_{t=1}^{T} x^\top A y_t - \frac{R_T}{T} \\
&= \max_{x \in \Delta_m} x^\top A \bar{y} - \frac{R_T}{T} \\
&\geq \min_{y \in \Delta_n} \max_{x \in \Delta_m} x^\top A y - \frac{R_T}{T} \\
&\to \min_{y \in \Delta_n} \max_{x \in \Delta_m} x^\top A y
\end{aligned}$$

as $T \to \infty$. □

The proof shows that $\bar{x}$ and $\bar{y}$ are asymptotically optimal, in the sense that the gain of $\bar{x}$ and the loss of $\bar{y}$ approach the value of the game. Next lecture, we'll consider a specific low regret learning algorithm: gradient ascent.