Introduction to Game Theory
Matrix Games and Lagrangian Duality

1. Introduction

In this section we study only *finite, two person, zero-sum, matrix games*. We introduce the basics by studying a Canadian drinking game. One and two dollar coins are very popular in Canada. The one dollar coin, introduced in 1987, is adorned with the picture of the common loon an aquatic bird found throughout Canada. Quickly the one dollar coin was nicknamed the *loonie*. The two dollar coin was issued in 1996. At the time there was a national competition for its naming with *Nanuq* the winning name. But the popular name for the coin quickly became the portmanteau *toonie*. The game is played with loonies and toonies, and, so I suppose, should be called *loonie-toonie*. Actually, this is a version of an ancient game called *Morra* which dates back to at least Roman times and most probably much earlier. Regardless, the rules are as follows: each player chooses either the loonie or the toonie and places the single coin in their closed right hand with the choice hidden from their opponent. Each player then guesses the play of the other. If only one guesses correctly, then the other player pays to the correct guesser the sum of the coins in both their hands. If both guess incorrectly or both correctly, then there is no payoff. This is an example of a *zero-sum* game since in each case, what one player loses the other player gains. This is not always the case. For example, in a casino, the house always takes a commission from the winner. That is, the winner makes less than the loser loses.

We now model the game of Morra mathematically. The first step is to define the *payoff matrix*. Designate one of the players as the *column player* and the other the *row player*. The payoff matrix consists of the payoff to the column player based on the strategy employed by both players in a given round of play. The strategies for either player are the same and they consist of a pair of decisions. The first is the choice of coin to hide, and the second is the guess for the opponents hidden coin. We denote these decisions by (i, j) with $i = 1, 2$ and $j = 1, 2$. For example, the strategy $(2, 1)$ is to hide the toonie in your fist and to guess your opponent is hiding a loonie. The payoff matrix P to the column player is given by

$$\begin{array}{c|cccc} & (1,1) & (1,2) & (2,1) & (2,2) \\ \hline (1,1) & 0 & -2 & 3 & 0 \\ (1,2) & 2 & 0 & 0 & -3 \\ (2,1) & -3 & 0 & 0 & 4 \\ (2,2) & 0 & 3 & -4 & 0 \end{array} .$$

For example, if the row player plays strategy $(2, 2)$ while the column player uses strategy $(2, 1)$, then the column player must pay the row player \$4.

The elements of P are the payoffs for the use of a *pure* strategy. But this game is played over and over again. So it is advisable for the column player to use a different pure strategy on each play. How should these strategies be chosen? One possibility is for the column player to decide on a long run frequency of play for each strategy, or equivalently, to decide on a probability of play for each strategy on each play. This is called a *mixed* strategy which can be represented as a vector of probabilities in R^4:

$$0 \le x \quad \text{and} \quad \mathsf{e}^T x = 1, \tag{1}$$

where e always represents the vector of all ones of the appropriate dimension, in this case $\mathsf{e} = (1, 1, 1, 1)^T$. Given a particular mixed strategy, one can easily compute the expected payoff to the column player for each choice of pure strategy by the row player. For example, if the row player chooses pure strategy $(1,1)$, then the expected payoff to the column player is

$$0 \cdot x_1 - 2 \cdot x_2 + 3 \cdot x_3 + 0 \cdot x_4 = \sum_{j=1}^{4} P_{1j} x_j \ .$$

Now given that the column player will use a mixed strategy, what mixed strategy should be chosen? One choice is the strategy that maximizes the column player's minimum expected payoff over the range of the row player's pure strategies. This strategy can be found by solving the optimization problem

$$\max_{0 \le x,\ \mathsf{e}^T x = 1} \ \min_{i=1,2,3,4} \sum_{j=1}^{4} P_{ij} x_i \ . \tag{2}$$

Note that this problem is equivalent to the linear program

$$\begin{array}{lll} \mathcal{C} & \text{maximize} & \gamma \\ & \text{subject to} & \gamma \mathsf{e} \le Px, \\ & & \mathsf{e}^T x = 1 \\ & & 0 \le x. \end{array}$$

On the flip side, the row player can also chose a mixed strategy of play, $0 \le y,\ \mathsf{e}^T y = 1$. In this case, the expected payoff to the column player when the column player uses the pure strategy $(2,1)$ is

$$3 \cdot y_1 + 0 \cdot y_2 + 0 \cdot y_3 - 4 \cdot y_4 = \sum_{i=1}^{4} P_{i3} y_i \ .$$

How should the row player decide on their strategy? One approach is for the row player to minimize the maximum expected payoff to the column player:

$$\min_{0 \le y,\ \mathsf{e}^T y = 1} \ \max_{j=1,2,3,4} \sum_{i=1}^{4} P_{ij} y_i. \tag{3}$$

This problem is equivalent to the linear program

$$\begin{array}{lll} \mathcal{R} & \text{minimize} & \eta \\ & \text{subject to} & P^T y \le \eta \mathsf{e}, \\ & & \mathsf{e}^T y = 1 \\ & & 0 \le y. \end{array}$$

Both the column player's problem $\mathcal{C}$ and the row player's problem $\mathcal{R}$ are linear programming problems. Let us pause for a moment to consider their dual linear programs. We begin with the column player's problem by putting it into our general standard form so that we can immediately write down its dual LP. Rewriting we have

$$\begin{array}{llll} \mathcal{C} & \text{maximize} & \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} \gamma \\ x \end{pmatrix} & \\ & \text{subject to} & \begin{bmatrix} 0 & \mathsf{e}^T \end{bmatrix} \begin{pmatrix} \gamma \\ x \end{pmatrix} = 1 & (\tau) \\ & & \begin{bmatrix} \mathsf{e} & -P \end{bmatrix} \begin{pmatrix} \gamma \\ x \end{pmatrix} \le 0 & (y) \\ & & 0 \le x & . \end{array}$$

The dual problem becomes

$$\begin{array}{ll} \text{minimize} & \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} \tau \\ y \end{pmatrix} \\ \text{subject to} & \begin{bmatrix} 0 & \mathsf{e}^T \end{bmatrix} \begin{pmatrix} \tau \\ y \end{pmatrix} = 1 \\ & \begin{bmatrix} \mathsf{e} & -P^T \end{bmatrix} \begin{pmatrix} \tau \\ y \end{pmatrix} \geq 0 \\ & 0 \leq y. \end{array}$$

Rewriting this dual, we have the LP

$$\begin{array}{ll} \text{minimize} & \tau \\ \text{subject to} & P^T y \leq \tau \mathsf{e} \\ & \mathsf{e}^T y = 1 \\ & 0 \leq y\ . \end{array}$$

But this is just the row player's problem $\mathcal{R}$! That is, the row player's and the column players problems are dual to each other! Also, observe that the feasible regions for both the primal and dual problems are always nonempty (why?) and bounded in the variables x and y, respectively (why?), and so the optimal values of both are necessarily bounded (why?). Hence, by the Strong Duality Theorem solutions to both the primal and dual problems exist with the optimal values coinciding.

Let us now apply this structure to the loonie-toonie game. One way to compute an optimal solution is to guess what it is and then use the Weak Duality Theorem to verify. In this game neither the column player nor the row player has a clear advantage, so it is reasonable to guess that their optimal strategies should be the same giving the same optimal value of zero. One guess that corresponds to this intuition is

$$\bar{\gamma} = 0, \quad \bar{x} = \begin{pmatrix} 0 \\ 3/5 \\ 2/5 \\ 0 \end{pmatrix} \quad \text{and} \quad \bar{\eta} = 0, \quad \bar{y} = \begin{pmatrix} 0 \\ 4/7 \\ 3/7 \\ 0 \end{pmatrix},$$

with

$$P = \begin{bmatrix} 0 & -2 & 3 & 0 \\ 2 & 0 & 0 & -3 \\ -3 & 0 & 0 & 4 \\ 0 & 3 & -4 & 0 \end{bmatrix}.$$

Observe that

$$P\bar{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1/7 \end{pmatrix} \quad \text{and} \quad P^T \bar{y} = \begin{pmatrix} -1/7 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

so $(\bar{\gamma}, \bar{x})$ is primal feasible while $(\bar{\eta}, \bar{y})$ is dual feasible and their optimal values coincide at zero. Therefore, by the Weak Duality Theorem, they are optimal for their respective problems. Note that although the game seems to be structured in a way that it makes no difference who the column and row players are, the optimal solutions given above are not the same strategy. Is there an issue yet to be resolved? What can be said about the strategies

$$\bar{\gamma} = 0, \quad \bar{x} = \begin{pmatrix} 0 \\ 4/7 \\ 3/7 \\ 0 \end{pmatrix} \quad \text{and} \quad \bar{\eta} = 0, \quad \bar{y} = \begin{pmatrix} 0 \\ 3/5 \\ 2/5 \\ 0 \end{pmatrix}?$$

Can more be said about the structure of the solution sets for both the primal and dual games in this case?

Any matrix $P \in \mathrm{R}^{m \times n}$ can define a matrix game of the form $\mathcal{C}$ having dual $\mathcal{R}$. These games are always feasible with bounded feasible region. Therefore, by the Strong Duality Theorem, both $\mathcal{C}$ and $\mathcal{R}$ always have optimal solutions with a common optimal value. This common optimal value is called the value of the game. If the value of the game is zero, then it is said to be a *fair game* since neither the column or row player has an advantage. Games such as Morra are said to be symmetric since their payoff matrix is skew symmetric, i.e., $P^T = -P$. Symmetric games are always fair (why?). A pair of strategies for the column and row players are said to constitute a *Nash equilibrium* if prior knowledge of the mixed strategy of ones opponent has no effect on ones own choice of strategy. Is the case for the strategies provided by our minimax approach to the game of Morra?

2. Equilibria and Minimax Problems

We can change the way we represent matrix games by making a simple observation about how the maximum or minimum of a finite set of objects can be represented. If $\{= a_1, a_2, \ldots, a_N\}$ is any collection of real numbers, and $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_N)^T$ is any N-dimensional probability vector, then

$$\min\{a_1, a_2, \ldots, a_N\} \;\le\; \sum_{i=1}^{N} \lambda_i a_i \;\le\; \max\{a_1, a_2, \ldots, a_N\}.$$

That is, the expected value, or average, of the a_i's in any discrete probability distribution always lies between the minimum and the maximum values of the a_i's. Consequently,

$$\min\{a_1, a_2, \ldots, a_N\} = \min_{0 \le y, \mathrm{e}^T y = 1} y^T a \quad \text{and} \quad \max\{a_1, a_2, \ldots, a_N\} = \max_{0 \le y, \mathrm{e}^T y = 1} y^T a.$$

By applying this observation to (2) and (3), we obtain the following representations for $\mathcal{C}$ and $\mathcal{R}$, respectively:

$$\max_{0 \le x, \mathrm{e}^T x = 1} \; \min_{0 \le y, \mathrm{e}^T y = 1} \; y^T P x \;, \tag{4}$$

and

$$\min_{0 \le y, \mathrm{e}^T y = 1} \; \max_{0 \le x, \mathrm{e}^T x = 1} \; y^T P x \;. \tag{5}$$

This implies that the difference between the column player's problem and the row player's problem is simply reversing the order in which the min andmax are taken. The fact that the optimal values in (4) and (5) coincide is an instance of what is known as a *Minimax Theorem*. Such theorems play an important role in several areas of application, particularly in economics and game theory. The first big theorem of this type was proven by John von Neumann of which the following theorem is an elementary special case.

Theorem 1. *Given any matrix $P \in \mathrm{R}^{m \times n}$, one has*

$$\max_{0 \le x, \mathrm{e}^T x = 1} \; \min_{0 \le y, \mathrm{e}^T y = 1} \; y^T P x \;=\; \min_{0 \le y, \mathrm{e}^T y = 1} \; \max_{0 \le x, \mathrm{e}^T x = 1} \; y^T P x \;.$$

3. Lagrangian Duality

A general minimax problem can be obtained from any function $L\colon \mathrm{R}^n \times \mathrm{R}^m \mapsto \mathrm{R}$ and two sets $X \subset \mathrm{R}^n$ and $Y \subset \mathrm{R}^m$ and writing the two problems

$$\max_{x\in X}\ \min_{y\in Y}\ L(x,y) \qquad \text{and} \qquad \min_{y\in Y}\ \max_{x\in X}\ L(x,y).$$

In the case of matrix games, we have $L(x,y) = y^T Px$. Returning to the general case, define the function $p\colon \mathrm{R}^n \mapsto \mathrm{R}\cup\{-\infty\}$ and $d\colon \mathrm{R}^m \mapsto \mathrm{R}\cup\{+\infty\}$ by

$$p(x) := \min_{y\in Y} L(x,y) \qquad \text{and} \qquad d(y) := \max_{x\in X} L(x,y)\ .$$

We call p the primal objective function and d the dual objective, and we call the problem

$$\mathcal{P} \qquad \max_{x\in X} p(x)$$

the *Primal Problem* and

$$\mathcal{D} \qquad \min_{y\in Y} d(y)$$

the *Dual Problem.* Note that for every pair $(\bar{x},\bar{y}) \in X \times Y$,

$$p(\bar{x}) = \min_{y\in Y} L(\bar{x},y) \le L(\bar{x},\bar{y}) \le \max_{x\in X} L(x,\bar{y}) = d(\bar{y}). \tag{6}$$

The inequality (6) is called the *Weak Duality Theorem* for minimax problems of this type.

Theorem 2 (Weak Duality for Minimax)**.** *Let L, p, and d, be as defined above. Then for every $(x,y) \in X \times Y$,*

$$p(x) \le d(y). \tag{7}$$

Moreover, if $(\bar{x},\bar{y})$ are such that $p(\bar{x}) = d(\bar{y})$, then $\bar{x}$ solves $\mathcal{P}$ and $\bar{y}$ solves $\mathcal{D}$.

We call a point $(\bar{x},\bar{y}) \in X \times Y$ a *saddle point* for L, if

$$L(x,\bar{y}) \le L(\bar{x},\bar{y}) \le L(\bar{x},y) \quad \forall\ (x,y) \in X \times Y. \tag{8}$$

Theorem 3 (Saddle Point Theorem)**.** *Let L, p, and d, be as defined above.*

(i) *If $(\bar{x},\bar{y})$ is a saddle point for L, the $\bar{x}$ solves $\mathcal{P}$ and $\bar{y}$ solves $\mathcal{D}$ with the optimal value in both $\mathcal{P}$ and $\mathcal{D}$ equal to the saddle point value $L(\bar{x},\bar{y})$.*
(ii) *If $\bar{x}$ solves $\mathcal{P}$ and $\bar{y}$ solves $\mathcal{D}$ with the optimal values coinciding, then $(\bar{x},\bar{y})$ is a saddle point for L.*

Proof. (i) Suppose $(\bar{x},\bar{y})$ is a saddle point for L. Let $\epsilon > 0$ and choose $(x_\epsilon, y_\epsilon) \in X \times Y$ so that

$$d(\bar{y}) - \epsilon \le L(x_\epsilon,\bar{y}) \qquad \text{and} \qquad L(\bar{x},y_\epsilon) \le p(\bar{x}) + \epsilon.$$

By combining this with (8), we obtain

$$d(\bar{y}) - \epsilon \le L(\bar{x},\bar{y}) \le p(\bar{x}) + \epsilon.$$

Since this holds for all $\epsilon > 0$, we have $d(\bar{y}) \le L(\bar{x},\bar{y}) \le p(\bar{x})$. But then, by the Weak Duality Theorem, $p(\bar{x}) \le d(\bar{y}) \le L(\bar{x},\bar{y}) \le p(\bar{x}) \le d(\bar{y})$ which, again by the Weak Duality Theorem, proves the result. $\square$

3.1. **Linear Programming Duality.** Let $A \in \mathrm{R}^{m\times n}$, $b \in \mathrm{R}^m$, and $c \in \mathrm{R}^n$, and define

$$L(x,y) := c^T x + y^T b - y^T Ax,$$

with $X := \mathrm{R}^n_+$ and $Y = \mathrm{R}^m_+$. Then

$$\begin{aligned} p(x) &= \min_{0\le y} L(x,y) = \min_{0\le y} c^T x + y^T(b - Ax) \\ &= c^T x + \min_{0\le y} y^T(b-Ax) \\ &= c^T x + \begin{cases} 0 & , \ Ax \le b, \\ -\infty & , \text{else.} \end{cases} \end{aligned}$$

and

$$\begin{aligned} d(y) &= \max_{0\le x} L(x,y) = \max_{0\le x} y^T b + (c - A^T y)^T x \\ &= y^T b + \max_{0\le x} (c - A^T y)^T x \\ &= y^T b + \begin{cases} 0 & , \ A^T y \ge c, \\ +\infty & , \text{else.} \end{cases} \end{aligned}$$

Therefore, the primal problem has the form

$$\mathcal{P} \qquad \max_{0\le x} p(x) = \begin{array}{ll} \max & c^T x \\ \text{s.t.} & Ax \le b, \ 0 \le x, \end{array}$$

while the dual problem takes the form

$$\mathcal{D} \qquad \min_{0\le y} d(y) = \begin{array}{ll} \min & b^T y \\ \text{s.t.} & A^T y \ge c, \ 0 \le y. \end{array}$$

In this case, the function L is called the *Lagrangian*, and this development is an instance of *Lagrangian duality*. Observe that if $\bar{x}$ solves $\mathcal{P}$ and $\bar{y}$ solves $\mathcal{D}$, then the Saddle Point Theorem tells us that $p(\bar{x}) = L(\bar{x},\bar{y}) = d(\bar{y})$, or equivalently,

$$c^T\bar{x} = c^T\bar{x} + b^T\bar{y} - \bar{y}^T A\bar{x} = b^T\bar{y},$$

or equivalently,

$$\bar{y}^T(b - A\bar{x}) = 0 \qquad \text{and} \qquad \bar{x}^T(c - A^T\bar{y}) = 0,$$

which is just the Complementary Slackness Theorem.

3.2. **Convex Quadratic Programming Duality.** One can also apply the Lagrangian Duality Theory in the context of Convex Quadratic Programming. To see how this is done let $Q \in \mathrm{R}^{n\times n}$ be symmetric and positive definite, and let $c \in \mathrm{R}^n$. Consider the convex quadratic program

$$\mathcal{D} \quad \begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Qx + c^T x \\ \text{subject to} & Ax \le b, \ 0 \le x \ . \end{array}$$

The Lagrangian is given by

$$L(x,y,v) = \frac{1}{2}x^T Qx + c^T x + y^T(A^T x - b) - v^T x \quad \text{where } 0 \le y, \ 0 \le v.$$

The dual objective function is

$$g(y,v) = \min_{x\in\mathrm{R}^n} L(x,y,v) \ .$$

The goal is to obtain a closed form expression for g with the variable x removed by using the first-order optimality condition $0 = \nabla_x L(x, y, v)$. This optimality condition completely identifies the solution since L is convex in x. We have

$$0 = \nabla_x L(x, y, v) = Qx + c + A^T y - v.$$

Since Q is invertible, we have

$$x = Q^{-1}(v - A^T y - c).$$

Plugging this expression for x into $L(x, y, v)$ gives

$$\begin{aligned}
g(y, v) &= L(Q^{-1}(v - A^T y - c), y, v) \\
&= \frac{1}{2}(v - A^T y - c)^T Q^{-1}(v - A^T y - c) \\
&\qquad + c^T Q^{-1}(v - A^T y - c) + y^T(AQ^{-1}(v - A^T y - c) - b) - v^T Q^{-1}(v - A^T y - c) \\
&= \frac{1}{2}(v - A^T y - c)^T Q^{-1}(v - A^T y - c) - (v - A^T y - c)^T Q^{-1}(v - A^T y - c) - b^T y \\
&= -\frac{1}{2}(v - A^T y - c)^T Q^{-1}(v - A^T y - c) - b^T y\ .
\end{aligned}$$

Hence the dual problem is

$$\begin{array}{ll}
\text{maximize} & -\frac{1}{2}(v - A^T y - c)^T Q^{-1}(v - A^T y - c) - b^T y \\
\text{subject to} & 0 \le y,\ 0 \le v\ .
\end{array}$$

Moreover, $(\bar{y}, \bar{v})$ solve the dual problem if an only if $\bar{x} = Q^{-1}(\bar{v} - A^T\bar{y} - c)$ solves the primal problem with the primal and dual optimal values coinciding.

All of this is just a glimpse into what is possible!