Linear Optimization
Matrix Games
and
Lagrangian Duality

A Canadian Drinking Game: Morra

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.

Mathematical Model of Morra

We define a *payoff matrix* for this game.
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.

Mathematical Model of Morra

The payoff matrix P to the column player is given by

$$\begin{array}{ccccc} & (1,1) & (1,2) & (2,1) & (2,2) \\ (1,1) & 0 & -2 & 3 & 0 \\ (1,2) & 2 & 0 & 0 & -3 \\ (2,1) & -3 & 0 & 0 & 4 \\ (2,2) & 0 & 3 & -4 & 0 \end{array} .$$

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.

Pure and Mixed Strategies

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?

Pure and Mixed Strategies

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 $\mathbb{R}^4$:

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

where e always represents the vector of all ones of the appropriate dimension, in this case $\mathsf{e} = (1, 1, 1, 1)^T$.

Expected Return

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

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

Optimal Mixed Strategy: Column Player

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

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

Optimal Mixed Strategy: Column Player

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

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

Optimal Mixed Strategy: Row Player

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

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

The Duals for $\mathcal{C}$ and $\mathcal{R}$

Both the column player's problem $\mathcal{C}$ and the row player's problem $\mathcal{R}$ are linear programming problems. What are their duals?

The Duals for $\mathcal{C}$ and $\mathcal{R}$

Both the column player's problem $\mathcal{C}$ and the row player's problem $\mathcal{R}$ are linear programming problems. What are their duals?

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

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

The Duals for $\mathcal{C}$ and $\mathcal{R}$

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

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

The Duals for $\mathcal{C}$ and $\mathcal{R}$

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 $\mathcal{R}$!

Observations

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.

Solution for Morra

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

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.

Solution for Morra

What can be said about the strategies

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

Equilibria and Minimax Problems

Observation: Given

$$\{a_1, a_2, \ldots, a_N\} \subset \mathbb{R} \quad \text{and} \quad \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_N)^T \in \mathbb{R}^N$$

with $0 \leq \lambda$ and $\mathrm{e}^T\lambda = 1$. Then

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

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.

Equilibria and Minimax Problems

Observation: Given

$$\{a_1, a_2, \ldots, a_N\} \subset \mathbb{R} \quad \text{and} \quad \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_N)^T \in \mathbb{R}^N$$

with $0 \leq \lambda$ and $\mathsf{e}^T\lambda = 1$. Then

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

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,

$$\min\{a_1, a_2, \ldots, a_N\} = \min_{0 \leq y, \mathsf{e}^T y = 1} y^T a$$

and

$$\max\{a_1, a_2, \ldots, a_N\} = \max_{0 \leq y, \mathsf{e}^T y = 1} y^T a.$$

Equilibria and Minimax Problems

Theorem

Given any matrix $P \in \mathbb{R}^{m \times n}$*, one has*

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

Lagrangian Duality

A general minimax problem can be obtained from any function $L : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$ and two sets $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ 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 $L(x, y) = y^T P x$.

Lagrangian Duality

A general minimax problem can be obtained from any function $L : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$ and two sets $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ 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 $L(x, y) = y^T P x$.
Define the function $p : \mathbb{R}^n \mapsto \mathbb{R} \cup \{-\infty\}$ and $d : \mathbb{R}^m \mapsto \mathbb{R} \cup \{+\infty\}$ 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*.

Lagrangian Duality

Note that for every pair $(\bar{x}, \bar{y}) \in X \times Y$,

$$p(\bar{x}) = \min_{y \in Y} L(\bar{x}, y) \leq L(\bar{x}, \bar{y}) \leq \max_{x \in X} L(x, \bar{y}) = d(\bar{y}). \tag{4}$$

The inequality (4) is called the *Weak Duality Theorem* for minimax problems of this type.

Theorem (Weak Duality for Minimax)

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

$$p(x) \leq d(y). \tag{5}$$

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 $\mathcal{D}$.

Lagrangian Duality

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

$$L(x, \bar{y}) \leq L(\bar{x}, \bar{y}) \leq L(\bar{x}, y) \quad \forall \ (x, y) \in X \times Y. \tag{6}$$

Theorem (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.*

Lagrangian Duality

Theorem (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.*

Lagrangian Duality

Theorem (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

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

By combining this with (6), we obtain

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

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

Linear Programming Duality

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$, and define

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

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

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

Linear Programming Duality

Therefore, the primal problem has the form

$$\mathcal{P} \qquad \max_{0 \le x} p(x) \;=\; \max \; c^T x \quad \text{s.t.} \quad Ax \le b,\; 0 \le x,$$

Linear Programming Duality

Therefore, the primal problem has the form

$$\mathcal{P} \qquad \max_{0\le x} p(x) \;=\; \max \quad c^T x \qquad \text{s.t.} \quad Ax \le b,\ 0 \le x,$$

while the dual problem takes the form

$$\mathcal{D} \qquad \min_{0\le y} d(y) \;=\; \min \quad b^T y \qquad \text{s.t.} \quad A^T y \ge c,\ 0 \le y.$$

Linear Programming Duality

Therefore, the primal problem has the form

$$\mathcal{P} \qquad \max_{0 \le x} p(x) \;=\; \max \; c^T x \quad \text{s.t.} \quad Ax \le b,\; 0 \le x,$$

while the dual problem takes the form

$$\mathcal{D} \qquad \min_{0 \le y} d(y) \;=\; \min \; b^T y \quad \text{s.t.} \quad A^T y \ge c,\; 0 \le y.$$

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.

Quadratic Programming Duality

Consider the convex quadratic program

$$\begin{array}{lll} \mathcal{P} & \text{minimize} & \frac{1}{2}x^T Q x + c^T x \\ & \text{subject to} & Ax \le b,\ 0 \le x\ , \end{array}$$

where $Q \in \mathbb{R}^{n \times n}$ is symmetric and positive definite.

Quadratic Programming Duality

Consider the convex quadratic program

$$\begin{array}{lll} \mathcal{P} & \text{minimize} & \frac{1}{2}x^TQx + c^Tx \\ & \text{subject to} & Ax \le b,\ 0 \le x\ , \end{array}$$

where $Q \in \mathbb{R}^{n\times n}$ is symmetric and positive definite.
The Lagrangian is given by

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

The dual objective function is

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

Quadratic Programming Duality

We need a closed form expression for g.
We obtain this from the first-order optimality condition on

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

which is

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

Quadratic Programming Duality

We need a closed form expression for g.
We obtain this from the first-order optimality condition on

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

which is

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

Since Q is invertible, we have

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

Quadratic Programming Duality

We need a closed form expression for g.
We obtain this from the first-order optimality condition on

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

which is

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

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

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

Quadratic Programming Duality

The dual problem is

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

Moreover, $(\bar{y}, \bar{v})$ solves 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.