A Prisoner's Dilemma

Ajay Kumar

1 Introduction

In Game Theory, one of the most important thought experiments is the Prisoner's Dilemma. It entails two rational agents who have been arrested and are being interrogated separately (i.e., communication is not allowed between them). Each of these agents are posed with two possible options: they can betray their partner; they can co-operate with their partner. However, they don't know what decision their partner is making. After the agents have made their decisions, they will each receive a prison sentence accordingly.

Let's map this out in a table with two rational agents, let's call them $\boldsymbol{A}$ and $\boldsymbol{B}$, and every possible outcome with their respective sentences (see **Figure 1**).

	$\boldsymbol{A}$ Co-operates	$\boldsymbol{A}$ Betrays
$\boldsymbol{B}$ Co-operates	Both $\boldsymbol{A}$ and $\boldsymbol{B}$ receive small medium sized sentence (let's call this sentence α years)	$\boldsymbol{A}$ receives a minimum sentence (let's call this sentence β years) $\boldsymbol{B}$ receives a maximum sentence (let's call this sentence γ years)
$\boldsymbol{B}$ Betrays	$\boldsymbol{A}$ receives a maximum sentence (this sentence is γ years) $\boldsymbol{B}$ receives a minimum sentence (this sentence is β years)	Both $\boldsymbol{A}$ and $\boldsymbol{B}$ receive large medium sized sentence (let's call this sentence δ years)

Figure 1: a table with all possible outcomes and their respective sentences

In this table, $\gamma > \delta > \alpha > \beta$ and $\{\alpha, \beta, \gamma, \delta \in \mathbb{R}: \alpha, \beta, \gamma, \delta \geq 0\}$. To get a better understanding of this table numerically, we can set α, β, γ and δ to their typical values which are $\alpha = 1, \beta = 0, \gamma = 5$ and $\delta = 3$ (see **Figure 2**).

	$\boldsymbol{A}$ Co-operates	$\boldsymbol{A}$ Betrays
$\boldsymbol{B}$ Co-operates	Both $\boldsymbol{A}$ and $\boldsymbol{B}$ receive a 1-year sentence	$\boldsymbol{A}$ receives a 0-year sentence $\boldsymbol{B}$ receives a 5-year sentence
$\boldsymbol{B}$ Betrays	$\boldsymbol{A}$ receives a 5-year sentence $\boldsymbol{B}$ receives a 0-year sentence	Both $\boldsymbol{A}$ and $\boldsymbol{B}$ receive a 3-year sentence

Figure 2: a table with all possible outcomes and their typical respective sentences

We can now see why the prisoners would want to make a certain decision. If a prisoner wants to prioritise getting no sentence at all, they are most likely to take the risk of betraying their partner. If a prisoner is more ethical and believes that his partner is also ethical, they are most likely to co-operate, believing that they will both receive a minimal sentence.

This dilemma proposes an important question: what is the best choice? To answer this question, we will have to look at the probability of each outcome and create a sort of scoring system that enables the quantification of the data so we can determine, statistically, what the best option is for the rational agent.

2 The Probability of each Outcome

To analyse the problem quantitatively, let's set the probability of agent $\boldsymbol{A}$ co-operating equal to q and set the probability of agent $\boldsymbol{B}$ co-operating equal to r. This means that the probability of agent $\boldsymbol{A}$ betraying is equal to $1 - q$ and the probability of agent $\boldsymbol{B}$ betraying is equal to $1 - r$ (since the probability of co-operating and betraying for the agents must sum to 1).

$$\mathbb{P}(\boldsymbol{A\ co - operates}) = q \text{ where } \{q \in \mathbb{R}: 0 \leq q \leq 1\} \because it\ is\ a\ probabilty$$

$$\mathbb{P}(\boldsymbol{A\ betrays}) = 1 - q$$

$$\mathbb{P}(\boldsymbol{B\ co - operates}) = r\ where\ \{r \in \mathbb{R}: 0 \leq r \leq 1\} \because it\ is\ a\ probabilty$$

$$\mathbb{P}(\boldsymbol{B\ betrays}) = 1 - r$$

Now let's define the matrix M_P which encompasses the probabilities of all possible outcomes in a similar way to the table in **Figure 1**.

$$M_P = \begin{pmatrix} \text{Both } \boldsymbol{A} \text{ and } \boldsymbol{B} \text{ co} - \text{operate} & \boldsymbol{A} \text{ betrays and } \boldsymbol{B} \text{ co} - \text{operates} \\ \boldsymbol{A} \text{ co} - \text{operates and } \boldsymbol{B} \text{ betrays} & \text{Both } \boldsymbol{A} \text{ and } \boldsymbol{B} \text{ betray} \end{pmatrix}$$

$$M_P = \begin{pmatrix} \mathbb{P}(\boldsymbol{A\ co - operates}) \times \mathbb{P}(\boldsymbol{B\ co - operates}) & \mathbb{P}(\boldsymbol{A\ betrays}) \times \mathbb{P}(\boldsymbol{B\ co - operates}) \\ \mathbb{P}(\boldsymbol{A\ co - operates}) \times \mathbb{P}(\boldsymbol{B\ betrays}) & \mathbb{P}(\boldsymbol{A\ betrays}) \times \mathbb{P}(\boldsymbol{B\ betrays}) \end{pmatrix}$$

$$M_P = \begin{pmatrix} qr & (1-q)r \\ q(1-r) & (1-q)(1-r) \end{pmatrix}$$

3 The Scoring System

To create the scoring system, we need to come back to the prison sentences that were given for each outcome (see **Figure 3**).

	$\boldsymbol{A}$ Co-operates	$\boldsymbol{A}$ Betrays
$\boldsymbol{B}$ Co-operates	Both $\boldsymbol{A}$ and $\boldsymbol{B}$ receive a α −year sentence	$\boldsymbol{A}$ receives a β −year sentence $\boldsymbol{B}$ receives a γ −year sentence
$\boldsymbol{B}$ Betrays	$\boldsymbol{A}$ receives a γ −year sentence $\boldsymbol{B}$ receives a β −year sentence	Both $\boldsymbol{A}$ and $\boldsymbol{B}$ receive a δ −year sentence

Figure 3: a table with all possible outcomes and their respective sentences where $\gamma > \delta > \alpha > \beta$

The simplest and best way to create a scoring system is to set the number of points for each agent equal to the size of their prison sentence (i.e., equal to α or β or γ or δ). This would mean that the aim for each agent is to receive the smallest possible number of points as that would equate to the shortest prison sentence. We can sum this up in a table (see **Figure 4**).

	$\boldsymbol{A}$ Co-operates	$\boldsymbol{A}$ Betrays
$\boldsymbol{B}$ Co-operates	$\boldsymbol{A's\ score}: \alpha$ $\boldsymbol{B's}$ score: α	$\boldsymbol{A's\ score}: \beta$ $\boldsymbol{B's}$ score: γ
$\boldsymbol{B}$ Betrays	$\boldsymbol{A's\ score}: \gamma$ $\boldsymbol{B's}$ score: β	$\boldsymbol{A's\ score}: \delta$ $\boldsymbol{B's}$ score: δ

Figure 4: a table with all possible outcomes and their scores where $\gamma > \delta > \alpha > \beta$

Let's now convert this table into matrix $M_{\boldsymbol{A}}$ and $M_{\boldsymbol{B}}$ to make the next section easier to calculate:

$$M_{\boldsymbol{A}} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

$$M_{\boldsymbol{B}} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}$$

4 Combining the Probabilities with the Scoring System

To combine the probability of each outcome with our scoring system, we can use the concept of expected value ($E[\boldsymbol{A}]$ *and* $E[\boldsymbol{B}]$). This is done by using the formula:

$$E[X] = \sum[(Score) \times (Probability\ of\ getting\ that\ score)]$$

To calculate $E[\boldsymbol{A}]$, we must combine matrix M_P with matrix $M_{\boldsymbol{A}}$:

$$E[\boldsymbol{A}] = \alpha qr + \beta(1-q)r + \gamma q(1-r) + \delta(1-q)(1-r)$$

To calculate $E[\boldsymbol{B}]$, we must combine matrix M_P with matrix $M_{\boldsymbol{B}}$:

$$E[\boldsymbol{B}] = \alpha qr + \gamma(1-q)r + \beta q(1-r) + \delta(1-q)(1-r)$$

But what does this expected value actually tell us? The expected value represents the average outcome over many repetitions of the game, weighted by the probability of each scenario (see **Section 5** but slightly different as we wouldn't know what the other prisoner previously did in the last round).

Imagine if we were in the shoes of agent $\boldsymbol{A}$. If we want to have a lower prison sentence compared to agent $\boldsymbol{B}$, our expected value should be lower than the expected value for agent $\boldsymbol{B}$ (i.e., $E[\boldsymbol{A}] < E[\boldsymbol{B}]$).

To help understand this, let's substitute the typical values for the length of the prison sentence into our expected value ($\alpha = 1, \beta = 0, \gamma = 5$ and $\delta = 3$):

$$E[\boldsymbol{A}] = qr + 5q(1-r) + 3(1-q)(1-r)$$

$$E[\boldsymbol{B}] = qr + 5(1-q)r + 3(1-q)(1-r)$$

Now we want our expected value to be less than the expected value of agent $\boldsymbol{B}$:

$$qr + 5q(1-r) + 3(1-q)(1-r) < qr + 5(1-q)r + 3(1-q)(1-r)$$

$$5q(1-r) < 5(1-q)r$$

$$q - qr < r - qr$$

$$q < r$$

This tells us that if we want our expected value to be lower than the expected value of agent $\boldsymbol{B}$, then we should betray them rather than co-operate since we want the probability of us co-operating to be lower than the probability of them co-operating (since $q < r$) – this is done by betraying them which sets $q = 0$.

Now what if we wanted to do this generally without giving the prison sentences specific values:

$$\alpha qr + \beta(1-q)r + \gamma q(1-r) + \delta(1-q)(1-r) < \alpha qr + \gamma(1-q)r + \beta q(1-r) + \delta(1-q)(1-r)$$

$$\beta(1-q)r + \gamma q(1-r) < \gamma(1-q)r + \beta q(1-r)$$

$$\beta r - \beta qr + \gamma q - \gamma qr < \gamma r - \gamma qr + \beta q - \beta qr$$

$$\beta r + \gamma q < \gamma r + \beta q$$

$$\gamma q - \beta q < \gamma r - \beta r$$

$$q(\gamma - \beta) < r(\gamma - \beta)$$

$$q < r \; [we\ can\ divide\ by\ (\gamma - \beta) \because \gamma > \beta]$$

This gives us the exact same outcome as when we used specific values for the prison sentence.

This logically makes sense for a single round of the prisoner's dilemma. By betraying our partner (though this is not a very ethical choice), the possible sentences we could receive are β and γ years (0 and 3 years) whereas by co-operating with our partner, the possible sentences we could receive are α and δ years (1 and 5 years) – we are essentially minimising our own prison sentence, selfishly. Furthermore, if we expect our partner to have the same selfish thoughts, then they are also most likely to betray us in which case our best choice is still to betray them.

This result is an example of a Nash Equilibrium in which betraying your partner leads to a better or equal outcome and so if you individually tried to change your decision from betraying to co-operating, you wouldn't get a better result.

5 The Iterated Prisoner's Dilemma

Let's say that, suddenly, the prisoners can magically go back in time and change their decisions, each knowing what decision the other prisoner made previously. But this repeated round of the prisoner's dilemma gets added to their original prison sentence. What is the best strategy now, knowing that they will have to face their partner again in another round of the prisoner's dilemma?

Essentially, the addition of more rounds helps the prisoners gain a better insight into who their partner is and allows an element of communication. Intuition tells us that it is better for us to co-operate with our partner initially in the first round so that they can create mutual trust in any upcoming rounds.

Let's try and explain this through our probabilities which are q for us and r for the other agent for co-operation. Say we decide to betray our partner on the first round. This means that in the next round, our partner is going to have less trust in us so the probability of them betraying us increases (i.e., $(1-r)$ increases) so r decreases. This means that in future rounds, they are less likely to co-operate with us. However, if we do the opposite and co-operate, they trust us more and so the probability of them co-operating with us increases and therefore r increases. This means that you both receive prison sentences of α-years (1 year) for the first round rather than receiving an unequal number of years and regretting it in upcoming rounds.

Therefore, while betrayal is optimal in a single round, cooperation can be seen as a more effective long-term strategy in repeated interaction.

6 Maths, Ethics and Psychology

However, the strategies I have mentioned above highlight the question: when harmful outcomes could occur, do we prioritise our own selfish needs over being moral?

The overlay between Maths, Ethics and Psychology help understand how to approach this question. For example, say we wanted to reach an equilibrium within the question in which we equally prioritised our own desires whilst remaining ethical. In a single round of the prisoner's dilemma, this can be achieved by setting our expected values equal ($E[\boldsymbol{A}] = E[\boldsymbol{B}]$) so that both agents receive the same prison sentence. This would tell us, mathematically, that $q = r$ which says that the probability that we co-operate with them is the same as the probability of them co-operating with us. Therefore, the equilibrium is reached when both agents either co-operate with each other or both betray each other. But this equilibrium relies on the ability of one agent to predict the other agent's decision. Thus, in a situation like this, the optimal result could be achieved by understanding the statistics and behaviour of your partner.

7 Conclusion

To summarise, the prisoner's dilemma teaches us that in a single round, where you will never see your partner again, the best choice you can make for yourself is to betray them. In multiple rounds of the prisoner's dilemma, the best choice is to co-operate with them to foster mutual trust.

Ultimately, the Prisoner's Dilemma demonstrates how mathematical modelling can be used to analyse decision-making under uncertainty. It shows the versatility and vastness of maths as it can be applied in such seemingly foreign subjects. Thank you for reading this essay and I hoped you have enjoyed it.

8 References

M. V. Rajeev Gowda. *Teaching the Prisoners' Dilemma*. *Journal of Policy Analysis and Management* 15, no. 4 (1996): 646–53. http://www.jstor.org/stable/3326054.

Lombrozo, T. (2013) *The 'prisoner's dilemma' tests women in and out of jail*, *NPR*. Available at: https://www.npr.org/sections/13.7/2013/07/29/206513599/the-prisoner-s-dilemma-goes-to-jail-and-some-science-reporting-should-follow.

Wikipedia Contributors (2025). *Prisoner's dilemma*. [online] Wikipedia. Available at: https://en.wikipedia.org/wiki/Prisoner%27s_dilemma#Real-life_examples.

Chen, J. (2024). *Nash Equilibrium: How It Works in Game Theory, Examples, Plus Prisoner's Dilemma*. [online] Investopedia. Available at: https://www.investopedia.com/terms/n/nash-equilibrium.asp.