
NE
E
SE
S
SW
W
NW
20
40
60
80
100
120
140
180
200
220
240
260
280
300
320
360

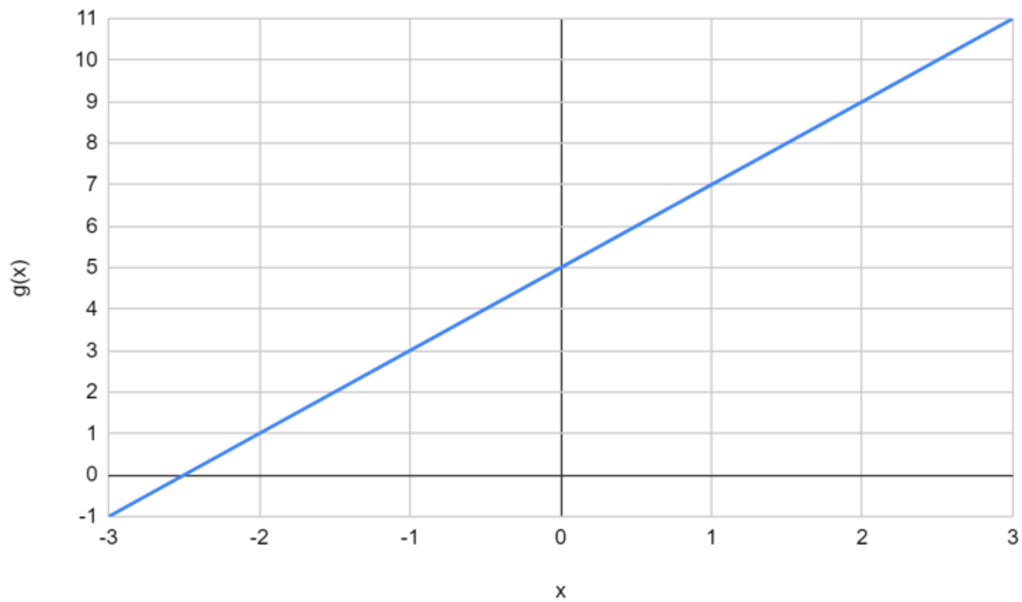
Calculus in Game and Decision Theory

A Very Mathematical Explanation of Derivatives

This post is meant for readers familiar with algebra and derivatives, but want to deepen their understanding and/or need a refresh.

Linear functions

Let's start with a family of very basic functions: the *linear functions*, expressed as $f(x) = ax + b$. You might remember its derivative is $f'(x) = a$, because x is multiplied by a and the constant b "disappears" when taking the derivative. This is correct, but let's actually calculate the derivative. Since $f(x)$ is a linear function, $f'(x)$ is the same for all x. That is, a linear function "goes up" with the same "speed" everywhere, as can be seen in the following graph for $g(x) = 2x + 5$:



For example, between $x = 0$ and $x = 1$, $f(x)$ increases with 2, just like it does between e.g. $x = 2$ and $x = 3$. Therefore, determining the average slope between x and $x + d$ will do. The average slope between x and $x + d$ is how much $f(x)$ increases between x and $x + d$, divided by the difference between x and $x + d$ (which is d). Let $d = 1$, in which case we don't have to do the division, as $f'(x) = \frac{f(x+1) - f(x)}{1} = f(x + 1) - f(x)$. Filling in $ax + b$ for $f(x)$ and $a(x + 1) + b$ for $f(x + 1)$, we get:

$$f'(x) = f(x + 1) - f(x) = (a(x + 1) + b) - (ax + b) = ax + a + b - ax - b = a$$

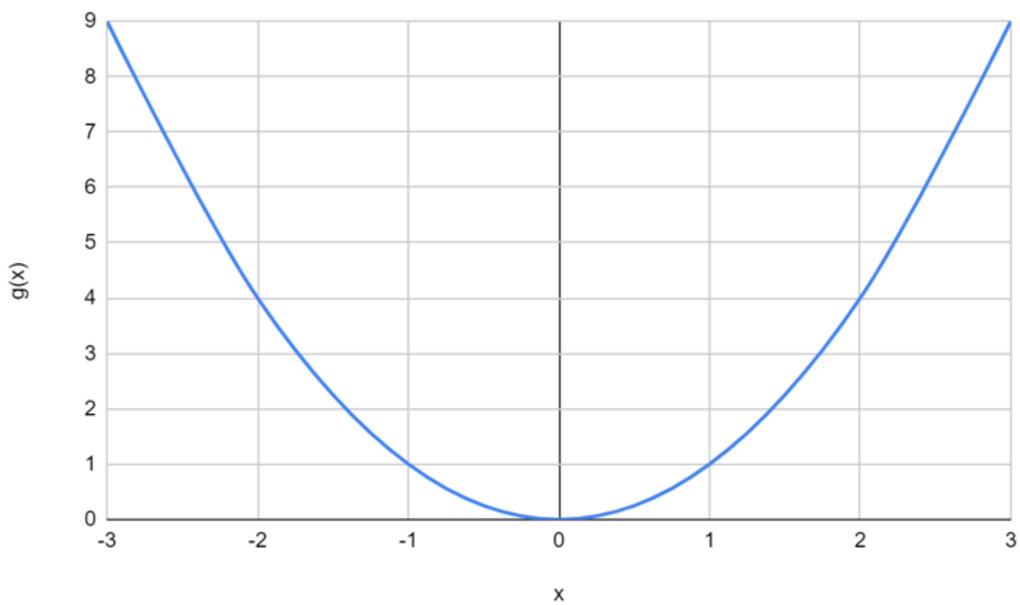
There it is! $f'(x) = a$. So for $g(x) = 2x + 5$, where $a = 2$, this means $g'(x) = 2$.

Polynomials (and more)

Polynomials are functions with the following form:

$$f(x) = a_0 x^n + a_1 x^{n-1} + ... + a_{n-1} x + a_n$$

Determining their derivative is a bit more tricky than determining the derivative of a linear function, because now, the derivative isn't necessarily the same everywhere. After all, take $g(x) = x^2$:



We can see this is a curved line, and so the derivative is constantly changing. We can sill do something like the "trick" we did with linear functions, but we can't determine $f'(x)$by looking how $f(x)$ changes between x and $x + 1$: that would assume $f'(x)$ is the same between x and $x + 1$, which isn't true. For x and $x + 0.001$, we would have a better estimate of $f'(x)$, but we'd still assume $f'(x)$ to be constant between these values. We need to determine how $f(x)$ changes between x and some $x + d$, where d needs to approach zero: the smaller d gets, the more accurate our calculation for $f'(x)$ becomes. We can do this using limits:

$$f'(x) = \lim_{d \to 0} \frac{f(x+d) - f(x)}{d}$$

(Since d isn't 1 now, we need to do the division.) We can read this as follows: what value does $\frac{f(x+d)+f(x)}{d}$ approach when d approaches 0?

Let's do this for the simple polynomial $g(x) = x^2$:

$$g'(x) = \lim_{d \to 0} \frac{g(x+d) - g(x)}{d} = \lim_{d \to 0} \frac{(x+d)^2 - x^2}{d} = \lim_{d \to 0} \frac{x^2 + 2dx + d^2 - x^2}{d} = \lim_{d \to 0} \frac{2dx + d^2}{d}$$

$$g'(x) = \lim_{d \to 0} \left(\frac{2dx}{d} + \frac{d^2}{d}\right) = \lim_{d \to 0} (2x + d)$$

When d approaches 0, $2x + d$ becomes $2x$:

$g'(x) = \lim_{d \to 0}(2x + d) = 2x.$

So for $g(x) = x^2$, $g'(x) = 2x$. You might have learned the general rule:

For $f(x) = ax^b$, $f'(x) = abx^{b-1}$

This is known as the *power rule*, and indeed works for $g(x) = x^2$, where $a = 1$ and $b = 2$and $1 * 2 * x^{2-1} = 2x^1 = 2x$. It also works for the linear function $h(x) = 2x$, where $a = 2$ and $b = 1$: $h'(x) = 1 * 2x^{1-1} = 1 * 2x^0 = 1 * 2 * 1 = 2$. But does it work in general? Yes, and we can proof it.

Let's first proof it works for all natural numbers (0, 1, 2, 3, ...) b: $b \in \mathbb{N}$. We need the *product rule* and *mathematical induction* for this proof though, so let's discuss those first.

Product rule

The product rule states that when $f(x) = g(x) * h(x)$, $f'(x) = g'(x) * h(x) + g(x) * h'(x)$. So when e.g. $g(x) = 2x$ and $h(x) = 3x^2$, $f(x) = 2x * 3x^2$ and

$f'(x) = g'(x) * h(x) + g(x) * h'(x) = 2 * 3x^2 + 2x * 6x = 6x^2 + 12x^2 = 18x^2$. We can show the product rule is correct by determining what $f'(x)$ should be using the original definition of the derivative:

$$f'(x) = \lim_{d \to 0} \frac{f(x+d) - f(x)}{d} = \frac{g(x+d) * h(x+d) - g(x) * h(x)}{d}$$

Since we want to write $f'(x)$ as $g'(x) * h(x) + g(x) * h'(x)$, let's rewrite the divisor to include the terms $g(x + d) - g(x)$ and $h(x + d) - h(x)$:

$$f'(x) = \lim_{d \to 0} \frac{h(x)(g(x+d) - g(x)) + g(x+d)(h(x+d) - h(x))}{d}$$

and note that indeed, $h(x)(g(x + d) - g(x)) + g(x + d)(h(x + d) - h(x)) =$

$$h(x)g(x+d) - g(x)h(x) + g(x+d)h(x+d) - h(x)g(x+d) =$$

$g(x + d)h(x + d) - g(x)h(x)$, which was our original divisor.

Simplifying $f'(x) = \lim_{d\to 0} \frac{h(x)(g(x+d)-g(x))+g(x+d)(h(x+d)-h(x))}{d}$, we get

$$f'(x) = \lim_{d\to 0} \frac{h(x)(g(x+d)-g(x))}{d} + \lim_{d\to 0} \frac{g(x+d)(h(x+d)-h(x))}{d}$$

Since h(x) doesn't contain d, we can take it outside the first limit term. We can also rewrite the second term:

$$f'(x) = h(x) * \lim_{d\to 0} \frac{g(x+d)-g(x)}{d} + \lim_{d\to 0} g(x+d) * \lim_{d\to 0} \frac{h(x+d)-h(x)}{d}$$

When d approaches 0, $\lim_{d\to 0} g(x+d)$ becomes g(x). Furthermore, by definition,

$\lim_{d\to 0} \frac{g(x+d)-g(x)}{d} = g'(x)$ and $\lim_{d\to 0} \frac{h(x+d)-h(x)}{d} = h'(x)$,

so we now have $f'(x) = h(x) * g'(x) + g(x) * h'(x) = g'(x) * h(x) + g(x) * h'(x)$,

which is the product rule!

Mathematical induction

Mathematical induction is a method for proving something is true for all natural numbers N. For example, say we want to proof that for every natural number $n \in N$, $S(n) = \frac{n(n+1)}{2}$, where S(n) is simply $0 + 1 + 2 + \ldots + n$. We can do this by first showing the condition holds for n = 0. That's Step 1, and yes, it does: $\frac{0(0+1)}{2} = 0 = S(0)$. Then, we show that *if* the condition holds for *some* n, it also holds for $n + 1$. That's Step 2. So for this step we *assume* $S(n) = \frac{n(n+1)}{2}$, and need to *show* that $S(n+1) = \frac{(n+1)((n+1)+1)}{2}$. That holds as well: *if* $S(n) = \frac{n(n+1)}{2}$, *then* $S(n+1) = 0 + 1 + 2 + \ldots + n + (n+1) = S(n) + n + 1$. Since for this step we assumed $S(n) = \frac{n(n+1)}{2}$, we have $S(n+1) = \frac{n(n+1)}{2} + n + 1 = \frac{n(n+1)}{2} + \frac{2(n+1)}{2}$. So $S(n+1) = \frac{n(n+1)+2(n+1)}{2} = \frac{(n+1)(n+2)}{2} = \frac{(n+1)((n+1)+1)}{2}$.

So we now know that our condition holds for n = 0 and that *if* it holds for some n, it must also hold for $n + 1$. But then it holds for all natural numbers! Does our condition hold for n = 3? Yes! It holds for n = 0 by Step 1, so it holds for n = 1 by Step 2; but then, since it holds for n = 1, it also holds for n = 2, again by Step 2. Applying Step 2 one more time gives that the condition $S(n) = \frac{n(n+1)}{2}$ holds for n = 3 as well. And we can apply this process to every natural number!

Proof of the power rule for natural numbers

Using the product rule and mathematical induction, we can show that the power rule (for $f(x) = ax^b$, $f'(x) = abx^{b-1}$) works for all $b \in N$.

Step 1 is to show this is true for $b = 0$. Yes: then $f(x) = ax^b = ax^0 = a$, and $abx^{b-1} = a * 0x^{0-1} = 0 = f'(x)$. (Since $f(x)$ is constant (a), its derivative $f'(x)$ is indeed 0.

Step 2 is to show that if $f'(x) = abx^{b-1}$ for some $b \in N$ and $f(x) = ax^b$, then for $b + 1$ and $g(x) = ax^{b+1}$, $g'(x) = a(b + 1)x^b$.

We can write $g(x) = ax^{b+1}$ as $g(x) = x * ax^b$. Then, define $h(x) = x$. Then $g(x) = h(x) * f(x)$, and then the product rule says $g'(x) = h'(x) * f(x) + h(x) * f'(x)$. But by the assumption of Step 2, $f'(x) = abx^{b-1}$. Furthermore, $h'(x) = 1$. So $g'(x) = 1 * ax^b + x * abx^{b-1} = ax^b + abxx^{b-1} = ax^b + abx^b = a(b + 1)x^b$, which is what we wanted to proof!

So we have shown the power rule works for $b \in N$. We could extend this proof to e.g. cover negative integers for b as well. But I'd like to use a different method of proof, that proofs the power rule works for $b \in R$. For this, we first need to know the *chain rule*, the *constant multiple rule*, *Euler's number* and how to take the derivative of the natural logarithm.

Chain rule

Define $f(x) = (3x)^2$. (Note this is distinct from $3x^2$.) We want to determine its derivative. We could say $f(x) = (3x)^2 = 9x^2$, which would make $f'(x) = 18x$. This is true, but let's take the opportunity to study the *chain rule*. Define $g(x) = 3x$ and $h(x) = x^2$. We can then write $f(x)$ as $h(g(x))$. Then:

$$f'(x) = \lim_{d \to 0} \frac{f(x+d) - f(x)}{d} = \lim_{d \to 0} \frac{h(g(x+d)) - h(g(x))}{d}$$

Multiplying by $\frac{g(x+d)-g(x)}{g(x+d)-g(x)}$ which equals 1 and is allowed if $g(x + d) \neq g(x)$ (otherwise we are dividing by 0), gives:

$$f'(x) = \lim_{d \to 0} \frac{(h(g(x+d)) - h(g(x))) * (g(x+d) - g(x))}{(g(x+d) - g(x)) * d}$$

or $f'(x) = \lim_{d\to 0} \frac{h(g(x+d))-h(g(x))}{g(x+d)-g(x)} * \lim_{d\to 0} \frac{g(x+d)-g(x)}{d}$

Note that $\lim_{d\to 0} \frac{h(g(x+d))-h(g(x))}{g(x+d)-g(x)} = h'(g(x))$, and $\lim_{d\to 0} \frac{g(x+d)-g(x)}{d} = g'(x)$. So we now have $f'(x) = h'(g(x)) * g'(x)$. That's the *chain rule*, and it holds whenever we can write a function $f(x)$ as $f(x) = h(g(x))$. Originally, we said $f(x) = (3x)^2$, with $g(x) = 3x$ and $h(x) = x^2$. Then $h'(g(x)) = 2g(x) = 2 * 3x = 6x$ and $g'(x) = 3$. According to the chain rule, then, $f'(x) = 6x * 3 = 18x$, which is also what we got by applying the power rule to $f(x) = (3x)^2 = 9x^2$.

Before, we temporarily assumed $g(x+d) \neq g(x)$. What if $g(x+d) = g(x)$? Well, then $h(g(x+d)) = h(g(x))$, and $h(g(x+d)) - h(g(x)) = 0$. Then $f'(x) = \lim_{d \to 0} \frac{0}{d} = 0$, and $g'(x) = \lim_{d \to 0} \frac{g(x+d)-g(x)}{d} = \lim_{d \to 0} \frac{0}{d} = 0$. So the chain rule would still apply, as $h'(g(x)) * g'(x) = h'(g(x)) * 0 = 0 = f'(x)$.

Constant multiple rule

If $f(x) = a * g(x)$, $f'(x) = a * g'(x)$. This might make intuitive sense, but it also follows from the chain rule: define $h(x) = ax$ and $f(x) = h(g(x))$. Then $f'(x) = h'(g(x)) * g'(x) = a * g'(x)$, which is the *constant multiple rule*. Indeed, this same rule also follows from the product rule: if $f(x) = a * g(x)$, define $h(x) = a$. Then $f(x) = h(x) * g(x)$ and $f'(x) = h'(x) * g(x) + h(x) * g'(x) = 0 * g(x) + a * g'(x) = a * g'(x)$.

Euler's number and the natural logarithm

You might know that Euler's number e, which is chosen so that if $f(x) = e^x$, $f'(x) = e^x$. You may also remember the natural logarithm ln, where $e^{\ln x} = x$. What's the derivative of $f(x) = \ln x$? We can find it with the chain rule! Define $g(x) = e^{f(x)} = e^{\ln x}$ and $h(x) = e^x$. Then $g(x) = h(f(x))$, and applying the chain rule gives $g'(x) = h'(f(x)) * f'(x) = e^{\ln x} * f'(x)$. But also, $g(x) = e^{\ln x} = x$, so $g'(x) = 1$. So we learn $g'(x) = e^{\ln x} * f'(x) = 1$, or $x * f'(x) = 1$, and so $f'(x) = \frac{1}{x}$. So for $f(x) = \ln x$, $f'(x) = \frac{1}{x}$.

General proof of the power rule

Now we're ready to proof the power rule (for $f(x) = ax^b$, $f'(x) = abx^{b-1}$) works for $b \in R$. Let's rewrite x as $e^{\ln x}$. Then $f(x) = a(e^{\ln x})^b = ae^{b \ln x}$. Define $g(x) = ae^x$ and $h(x) = b \ln x$. Then $f(x) = g(h(x))$, and via the chain rule (and the constant multiple rule)
$f'(x) = g'(h(x)) * h'(x) = ae^{b \ln x} * b * \frac{1}{x} = ae^{b \ln x} * \frac{b}{x}$. Remember $ae^{b \ln x} = a(e^{\ln x})^b = ax^b$, so
$f'(x) = ax^b * \frac{b}{x} = \frac{abx^b}{x} = abx^{b-1}$, which is what we need to proof the power rule for $b \in R$.

Local maxima, local minima and second derivatives

As you might know, polynomials like $f(x) = -x^2$ can have local maxima (or peaks, where the graph first goes up and then goes down) and local minima (or minima, where the graph first goes down and then goes up). When a graph goes up, the derivative is positive; when it goes down, the derivative is negative. In the peak, the derivative must be 0! It's similar for valleys - the derivative is 0there, too.

That means we can find local maxima and local minima by setting the derivative to 0! For $f(x) = -x^2$

, $f'(x) = -2x$. $f'(x) = 0$ gives $-2x = 0$ and thus $x = 0$. Therefore, there must be a local maximum or minimum at $f(0)$. Which is it? Well, note that in a local maximum, the derivative must be decreasing (through 0): otherwise, the graph wouldn't first go up and then go down. But if the derivative is decreasing, the *derivative of the derivative, called the second derivative* (written $f''(x)$)*, must be negative*! Conversely, in a local minimum, the second derivative must be positive. For $f(x) = -x^2$, $f'(x) = -2x$ and $f''(x) = -2 < 0$. So $f(x) = 0$ is a local maximum!

Now consider $g(x) = x^3 + x^2$. We have $g'(x) = 3x^2 + 2x$, and $g'(x) = 0$ gives $3x^2 + 2x = 0$ or $x(3x + 2) = 0$. Then $x = 0 \lor 3x + 2 = 0$ and so $x = 0 \lor x = -\frac{2}{3}$. We have a local maximum or minimum in $x = 0$ and a local maximum or minimum in $x = -\frac{2}{3}$. $g''(x) = 6x + 2$, $g''(0) = 6 * 0 + 2 = 2 > 0$ and $g''(-\frac{2}{3}) = 6 * -\frac{2}{3} + 2 = -2 < 0$. Therefore, we have a local minimum in $x = 0$ and a local maximum in $x = -\frac{2}{3}$.

Multivariable functions

Multivariable functions are functions with, well, more than one variable. Take for example $f(x, y) = 2x + 3y$. For $x = 4$ and $y = 5$, we have $f(4, 5) = 2 * 4 + 3 * 5 = 23$. Or we can take $g(x, y, z) = x + 5y + 2z$, with $g(1, 1, 1) = 1 + 5 + 2 = 8$.

Partial derivatives

A *partial derivative* of a multivariable function is determined by treating all but one variable like constants and taking the derivative with respect to the one variable left. For example, for $f(x, y) = 2x + 3y$, we can derive with respect to x: $f'_x(x, y) = 2$ and with respect to y: $f'_y = 3$. More generally, for any 2-variable function $f(x, y)$, $f'_x(x, y) = \lim_{d\to 0} \frac{f(x+d,y)-f(x,y)}{d}$ and $f'_y(x, y) = \lim_{d\to 0} \frac{f(x,y+d)-f(x,y)}{d}$. For $f(x, y) = 2x + 3y$, this means $f'_x(x, y) = \lim_{d\to 0} \frac{f(x+d,y)-f(x,y)}{d} = \lim_{d\to 0} \frac{(2(x+d)+3y)-(2x+3y)}{d} = \lim_{d\to 0} \frac{(2x+2d+3y)-2x-3y}{d}$, which indeed simplifies to $f'_x(x, y) = \lim_{d\to 0} \frac{2d}{d} = \lim_{d\to 0} 2 = 2$.

The Calculus of Nash Equilibria

Now that we know a bit about derivatives, it's time to use them to find *dominant strategies* and *Nash equilibria*. It helps if the reader is familiar with Nash equilibria already.

Prisoner's dilemma

The payoff matrix of the Prisoner's dilemma can be as follows:

		Prisoner 2			
		Cooperate		Defect	
Prisoner 1	Cooperate		\$20		\$30
		\$20		\$0	
	Defect		\$0		\$10
		\$30		\$10	

We can see that the payoff for Prisoner 1 depends on her own action (Cooperate/Defect) but also on the action of Prisoner 2. Therefore, the payoff function for Prisoner 1 is a multivariable function: $V1(a_1, a_2)$, where a_n is the action of Prisoner n (and $n \in \{1, 2\}$). Let's say $a_n = 0$ when the action of Prisoner n is Cooperate, and $a_n = 1$ for Defect. So $a_n \in \{0, 1\}$.

Then $V1(a_1, a_2) = 20 - 20a_2 + 10a_1$, and crucially, $V1'_{a1}(a_1, a_2) = 10$. So for Defect ($a_1 = 1$), Prisoner 1's payoff will be 10 higher than for Cooperate ($a_2 = 0$), as can be confirmed in the table. Note that a_2 doesn't show up in $V1'_{a_1}(a_1, a_2)$: Defect gives \$10 more for Prisoner 1 *regardless* of what Prisoner 2 does, which makes Defect a *dominant strategy*. Don't get me wrong: Prisoner 1's payoff certainly *does* depend on what Prisoner 2 does. The point is that no matter what Prisoner 2 does, Prisoner 1's payoff will be \$10 higher when she (Prisoner 1) defects - and that's what's reflected in $V1'_{a1}(a_1, a_2) = 10$.

Since the payoff matrix is symmetrical, $V2(a_1, a_2) = 20 - 20a_1 + 10a_2$ and $V2'_{a_2}(a_1, a_2) = 10$. Prisoner 2 therefore also has a dominant strategy: Defect. The Prisoner's dilemma, then, has a Nash equilibrium: when both prisoners defect. With the partial derivatives, we demonstrated that when both prisoners defect, no one prisoner can do better by changing her action to Cooperate. If e.g. Prisoner 1 were to do this, then a_1 would go from 1 to 0, and since $V1'_{a_1}(a_1, a_2) > 0$, that would lower $V1$ (*regardless* of a_2). By symmetry, the same is true for Prisoner 2.

Nonlinear payoff functions

In the Prisoner's dilemma, the payoffs of both players (prisoners) can be modelled by linear payoff functions. What if the payoffs are nonlinear?

Let's say $V1(a_1, a_2) = -a_1^2$ and $V2(a_1, a_2) = -a_2^2 + a_2$. Then $V1'_{a_1}(a_1, a_2) = -2a_1$ and $V2'_{a_2}(a_1, a_2) = -2a_2 + 1$. A Nash equilibrium is a point where no player can do better by doing another action *given* the action of the other player; therefore, $V1(a_1, a_2)$ should be maximized with respect to a_1 while keeping a_2 constant, whereas $V2(a_1, a_2)$ should be maximized with respect to a_2 while keeping a_1 constant. If $V1(a_1, a_2)$ has a peak value with respect to a_1, $V1'_{a_1}(a_1, a_2) = -2a_1$ must be 0 in that point. $V1'_{a_1}(a_1, a_2) = -2a_1 = 0$ gives $a_1 = \frac{-0}{2} = 0$. So $a_1 = 0$ could represent a peak, but also a valley, since $V1'_{a_1}(a_1, a_2)$ would be 0 in both. If $V1'_{a_1}(a_1, a_2) = -2a_1$, $V1''_{a_1}(a_1, a_2) = -2 < 0$. So $a_1 = 0$ represents a local maximum in $V1(a_1, a_2)$ (when a_2 is held constant)! Since $V1(a_1, a_2)$ is quadratic, we can be sure this local maximum is the global maximum too (so there are no values for a_1 for which $V1(a_1, a_2)$ is higher when a_2 is held constant).

$V2'_{a_2}(a_1, a_2) = -2a_2 + 1 = 0$ gives $2a_2 = 1$ and $a_2 = \frac{1}{2}$. $V2''_{a_2}(a_1, a_2) = -2 < 0$, so $a_2 = \frac{1}{2}$ again represents a local maximum. $V2(a_1, a_2)$ is quadratic, so this is a global maximum as well.

So $a_1 = 0$ represents a global maximum for $V1$ (for a constant a_2), and $a_2 = \frac{1}{2}$ represents a global maximum for $V2$ (for a constant a_1). That means $a_1 = 0$ is a dominant strategy for player 1, $a_2 = \frac{1}{2}$ is a dominant strategy for player 2 and we have a Nash equilibrium in ($a_1 = 0$, $a_2 = \frac{1}{2}$).

Making things a bit more complicated

Let's now define $V1(a_1, a_2) = -a_1^2 * a^2$ and $V2(a_1, a_2) = -(a_2 - 1)^2$. Then for $V1'_{a_1}(a_1, a_2) = -2a_2 * a_1 = 0$, we have $a_2 = 0 \lor a_1 = 0$. $V1''_{a_1}(a_1, a_2) = -2a_2$, which is negative when $a_2 > 0$.

For $V2'_{a_2} = -2a_2 + 2 = 0$ we have $a_2 = 1$. $V2''_{a_2} = -2 < 0$, so this is a local optimum - and also the global one, since $V2(a_1, a_2)$ is quadratic. For $a_2 = 1$, $V1'_{a_1} = -2 * 1 * a_1 = -2a_1$. Solving for 0 gives $a_1 = 0$ (which we found earlier as well). And since $a_2 = 1 > 0$ and therefore $V1''_{a_1}(a_1, a_2) < 0$, we now have a local maximum for $V1(a_1, a_2)$! For a constant a_2, $V1(a_1, a_2)$ is quadratic, so this is the global maximum as well. We found a Nash equilibrium: $(a_1 = 0, a_2 = 1)$.

The Calculus of Newcomb's Problem

In the previous post, we applied some calculus to game theoretic problems. Let's now look at the famous Newcomb's problem, and how we can use calculus to find a solution.

Newcomb's problem

From Functional Decision Theory: A New Theory of Instrumental Rationality:

> *An agent finds herself standing in front of a transparent box labeled "A" that contains $1,000, and an opaque box labeled "B" that contains either $1,000,000 or $0. A reliable predictor, who has made similar predictions in the past and been correct 99% of the time, claims to have placed $1,000,000 in box B iff she predicted that the agent would leave box A behind. The predictor has already made her prediction and left. Box B is now empty or full. Should the agent take both boxes ("two-boxing"), or only box B, leaving the transparent box containing $1,000 behind ("one-boxing")?*

First, we need to come up with a function modeling the amount of money the agent gets. There's only one action to do: call this a. Then $V(a)$ is the function for the amount of money earned for each action (one-boxing or two-boxing). However, as you might know, historically, thinkers have diverged on what $V(a)$ should be.

Causal Decision Theory

Causal Decision Theory (CDT) states that an agent should look only at the causal effects of her actions. In Newcomb's problem, this means acknowledging that the predictor already made her prediction (and either put $\$1,000,000$ in box B or not), and that the agent's action now can't causally influence what's in box B. So either there's $\$1,000,000$in box B or not. In both cases, two-boxing earns $\$1,000$ more (the content of box A) than one-boxing. Then $V(a) = B + 1,000a$, where B is a constant representing the amount of money in box B and $a \in \{0, 1\}$, where $a = 0$ is one-boxing and $a = 1$ is two-boxing. $V'(a) = 1,000 > 0$, and of course $V(a)$ returns the most value for the highest value of a: $a = 1$ (two-boxing).

This is all pretty straightforward, but the problem is that almost all agents using CDT ends up with only $\$1,000$, as the predictor predicts the agents will two-box with 0.99 accuracy and put nothing in box B. It's one-boxing that gets you the $\$1,000,000$. Enter logical decision theories, e.g. Functional Decision Theory.

Functional Decision Theory

Functional Decision Theory (FDT) reasons about the effects of the agents *decision procedure* (which produces an action) instead of the effects of her *actions*. The point is that a decision procedure can be implemented multiple times. In Newcomb's problem, it seems the predictor implements the decision procedure of the agent: she can run this *model* of the agent's decision procedure to see what action it produces, and use it to predict what the actual agent will do. In $V(a) = B + 1,000a$, this means that B and a are *dependent on the same decision procedure*! The decision procedure's implementation in the agent produces a, and the implementation in the predictor is used to determine what B should be. Let's write d for the decision procedure: $d \in \{0, 1\}$, where 0 means a one-boxing decision and 1 means a two-boxing decision. Then $a = d$, and $B = 0.99 * 1,000,000 - 0.98 * 1,000,000d = 990,000 - 980,000d$. After all: if $d = 0$, the agent decides on one-boxing and the predictor will have predicted that with 0.99accuracy, giving an expected value of $0.99 * \$1,000,000 = \$990,000$ in box B. Should the agent decide to two-box, the predictor will have predicted *that* with probability $0.99$and only put $\$1,000,000$ in box B if she mistakenly predicted a one-box action. Then the expected value of box B is $0.01 * \$1,000,000 = \$10,000$, which for $d = 1$ is represented by $B = 990,000 - 980,000d$. Great! We now have $V(d) = (990,000 - 980,000d) + 1000d = 990,000 - 979,000d$. $V'(d) = -979,000 < 0$. The lowest possible decision, then, wins: $d = 0$, which gives $V(0) = 990,000$, whereas $V(1) = 990,000 - 979,000 = 11,000$.

This outcome reflects the fact that it's one-boxers who almost always win $\$1,000,000$, whereas two-boxer rarely do. If they do, they get the $\$1,000,000$ and the $\$1,000$ of box A, for a total of $\$1,001,000$, but the probability of getting the $\$1,000,000$ is too low for this to matter enough.