Unconstrained Optimisation

Intro Math for Economists (PEARL, Spring 2019)

David Murakami[1]

2 April 2019

[1]Keio University, Graduate School of Economics
Email: dh.murakami@keio.jp

Mathematical Prerequisites

- In this section we will focus on unconstrained optimisation.
- In plain English, this means finding the maximum or minimum value of something when we do NOT face a resource constraint. As you can probably guess, we often don't deal with unconstrained optimisation in economics – most of the time we have to maximise or minimise things when we face a resource constraint.
- As such, this section will be pretty close to high school maths, and a great chance for us to review functions and calculus.

Functions

Definition

A function is a rule for transforming an object into another object. The object you start with is called the input, and comes from some set called the domain. What you get back is called the output; it comes from some set called the codomain.

A very simple example:

$$f(x) = x^2 \tag{1}$$

Functions

Where f in equation (1) is a function which transforms any number into its square. Note we assume that both the domain and codomain are the set of real numbers ($-\infty$ to ∞), typically denoted as $\mathbb{R}$. Thus squaring any real number will give you a real number back.

$$f(2) = 2^2 = 4$$
$$f(-4) = -4^2 = 16$$
$$f(100) = 100^2 = 10,000$$
$$f(-100) = -100^2 = 10,000$$

Functions

Of course, we don't have to have to call functions 'f'. We can define it in any way we want. We can use 'g' as in $g(x)$, 'u' as in $u(x)$, or 'y' as in $y(x)$. The point is that we are saying that the functions g, u, and y all depend on inputs, x. As you get more familiar with functional forms, you can even ignore the '(x)' part and simply write functions of some general form such as:

$$y = mx + c \tag{2}$$

Which basically says that inputs m, x, and c produce some kind of output, y. It's also wore noting hat functions of the form in (2) are called 'linear functions'.

Functions

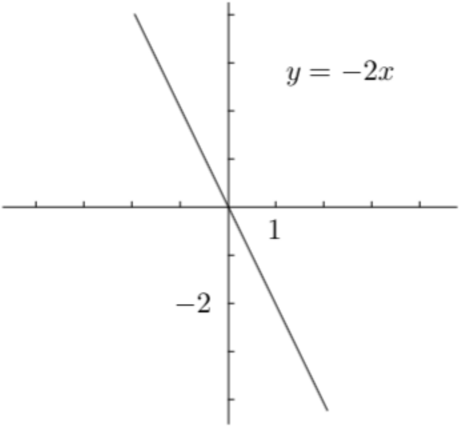
Definition

Functions of the form $y = mx + c$ are called linear. The slope of the linear function is given by m, and is referred to as the gradient.

Linear functions can be estimated and plotted by simply identifying two points along the function. This is because as the function's name implies, a single line connects all points of a linear function.

Functions

Figure: **Plot of a Linear Function**



Functions

- Figure 1 shows a cartesian plane representation for a function of the form in equation (2). The two variables, x and y are represented by the two axis.
- It is quite easy to extend the cartesian plane for functions of three variables (3D diagrams).
- Other functional forms that we will discover a lot in economics are polynomials.

Functions

Definition

Polynomial functions are functions which feature variables that are raised to nonnegative integer powers.

- A typical example of a polynomial function is one such as:

$$y = ax^2 + bx + c \tag{3}$$

Functions

- Equation (3) is commonly referred to as a 'quadratic function'.
- We won't have enough time to explore the quadratic function in further detail, but you will most likely run into it in future studies – as such I recommend reading the reference textbooks on this topic. More examples of polynomial functions are shown in Figure 2.
- Important topics not covered in the lecture include continuity, limits, and global/local maximums and minimums.

Functions

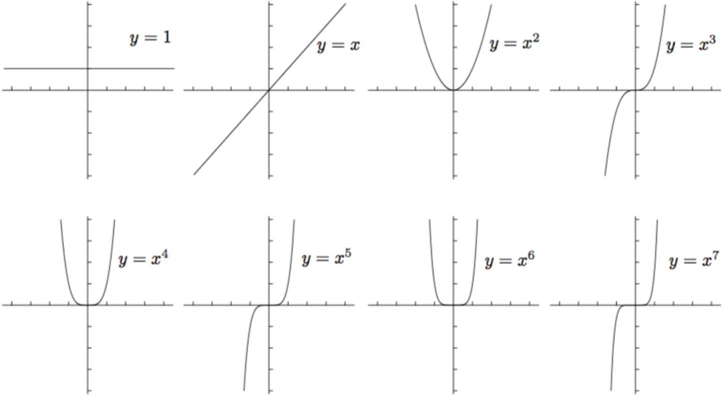


Figure: **Examples of Polynomial Functions**

Differentiation

- Having discussed functions, we need to quickly move onto differentiation – and this is where we will encounter calculus.
- The entire point of differentiation is to find out how fast/slow a function is changing.
- Consider an example where we wish to track the velocity of car travelling a certain distance in a certain amount of time.
- How would you calculate speed or velocity of the car?

$$\text{speed} = \frac{\text{distance}}{\text{time}} \tag{4}$$

Differentiation

- But in reality the above formula only gives you **average** speed.
- What if the car was travelling at different speeds throughout the journey?
- We need to consider displacement of the car – how far did the car travel in a certain amount of time? obviously this is related to how fast it was travelling at different points during the journey. We know

$$\text{displacement} = \text{final position - initial position}$$

Differentiation

- Refocus: how do we measure the velocity of the car at a given instant?
- Need to take average velocity of the car over smaller and smaller time periods.
- Let t be the moment in time we want to find out how fast the car is travelling, and suppose $t+1$ is a short time later than t.
- Write $v_{t \leftrightarrow t+1}$ to mean the average velocity of the car during the time interval beginning at time t and ending at time $t+1$.
- Now imagine squeezing t and $t+1$ closer and closer. From seconds, to milliseconds, to nanoseconds, and to picoseconds...
- This is the basic idea of limits.

Differentiation

- In fact, we can formalise this limit as:

$$\text{instantaneous velocity at time } t = \lim_{t+1 \to t} v_{t \leftrightarrow t+1}$$

- Then, suppose we have a function that tells us the displacement of the car – where the car is on its journey. Let $f(t)$ denote the function which tells us the car's position at time t. Thus:

$$v_{t+1 \to t} = \frac{f(t+1) - f(t)}{(t+1) - t} \tag{5}$$

- But obviously we can't say that $t = t + 1$, otherwise we end up with $\frac{0}{0}$, an indeterminate form.

Differentiation

- Need some algebra tricks.
- Define $h = (t+1) - t$. h must be very small since t and $t+1$ are so so close to each other.
- As $t+1 \to t$, indeed $h \to 0$.
- Let's put h into (5).
- Also, we can say that $t+1 = t+h$. This gives us:

$$\text{instantaneous velocity at time } t = \lim_{h \to 0} \frac{f(t+h) - f(t)}{h} \quad (6)$$

Differentiation

- Let look at an example. Suppose

$$f(t) = 15t^2 + 7 \tag{7}$$

- This function implies that $f(0) = 7$.
- Using (6) we have

$$\begin{aligned} &\lim_{h\to 0} \frac{f(t+h) - f(t)}{h} \\ &= \lim_{h\to 0} \frac{(15(t+h)^2 + 7) - (15t^2 + 7)}{h} \end{aligned}$$

Differentiation

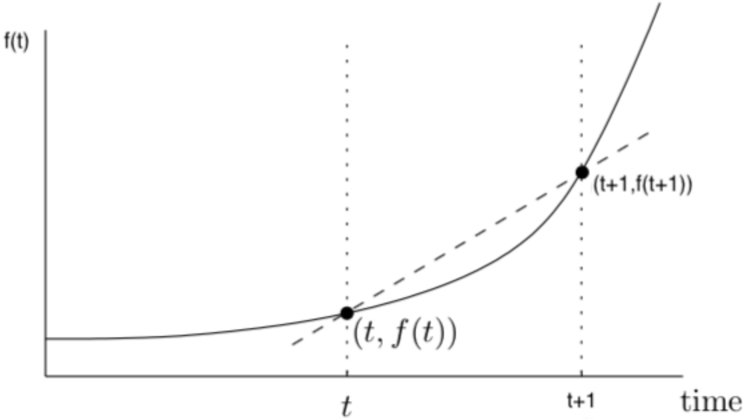
- Expand $(t+h)^2 = t^2 + 2th + h^2$ and simplify to get:

$$\begin{aligned}\lim_{h\to 0}\frac{15t^2 + 30th + 15h^2 + 7 - 15t^2 - 7}{h} &= \lim_{h\to 0}\frac{30th + 15h^2}{h}\\ &= \lim_{h\to 0}(30t + 15h)\end{aligned}$$

- Now just say $h = 0$ to see that instantaneous velocity at time t is $30t$.
- For every time unit of t, the car gets faster and faster at the constant rate of 30.

Differentiation

Figure: **Graphical Interpretation of Velocity**



Differentiation

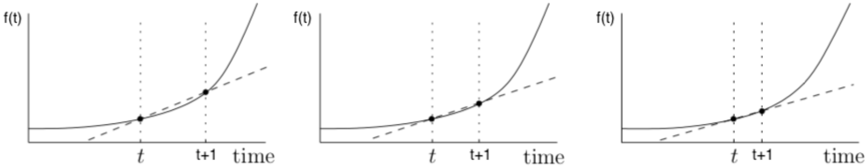
- Slope of the dashed line in Figure 3 is:

$$\frac{\text{rise}}{\text{run}} = \frac{f(t+1) - f(t)}{(t+1) - t}$$

- Exactly the same as average velocity $v_{t \leftrightarrow t+1}$ from the previous section.
- Now we want to find a similar graphical representation for instantaneous velocity.
- Need to take the limit as $t+1$ goes closer to t.

Differentiation

Figure: **Instantaneous Velocity Plots for Smaller Values of $t + 1$**

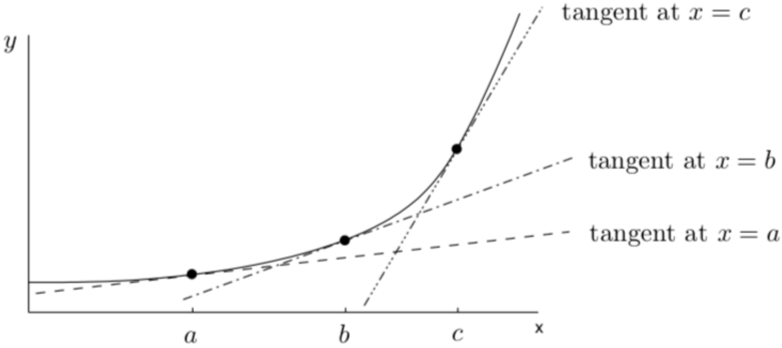


Differentiation

- Lines seem to get closer to the tangent line at the point $(t, f(t))$.
- Instantaneous velocity is the limit of the slopes of these lines as $t + 1 \to t$.
- Thus, we can say that the instantaneous velocity is exactly equal to the slope of the tangent line through $(t, f(t))$.
- If you can follow the logic thus far, then congratulations, you've basically understood the main point of calculus.
- Let's just focus on tangent lines a bit more for a bit...

Differentiation

Figure: **Tangent Lines**



Differentiation

- Clearly the lines have different slopes.
- Slope of the tangent lines depend on the value of x.
- i.e. Slope of the tangent line through $(x, f(x))$ is itself a function of x.
- This function is called the **derivative** or **differential** of f and is often written as f'.
- We say that we have differentiated the function f with respect to its variable x to get the function f'.

Differentiation

- Congratulations, you've just done calculus (well, at least one side of calculus – more on this later)!
- To summarise what we said neatly:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{8}$$

- Which works so long as f is differentiable at x.
- Let's work through a very simple example:

$$f(x) = x^2 \tag{9}$$

Differentiation

Let's find $f'(x)$ of (9):

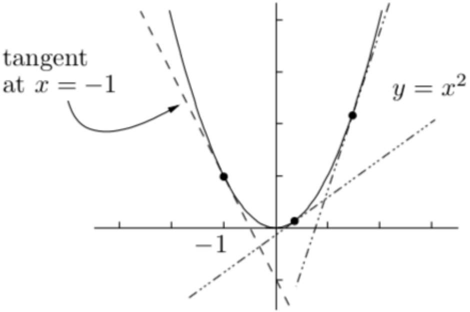
$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} \\ &= \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} = \lim_{h \to 0} \frac{2xh + h^2}{h} \\ &= \lim_{h \to 0} (2x + h) = 2x \end{aligned}$$

So the derivative of $f(x) = x^2$ is $f'(x) = 2x$.

Differentiation

This means that the slope of the tangent to the function $y = x^2$ at the point (x, x^2) is precisely $2x$. Here is the cartesian plane representation:

Figure: **Tangent Lines of $y = x^2$**



Differentiation

- In equation (8) we evaluated the quantity $f(x+h)$.
- The amount h represented how much x changed, so let's simply replace it with Δx:

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \tag{10}$$

- Since we know that $f(x)$ is simply y, we can state the following: $\Delta x = x_{\text{new}} - x$ and $\Delta y = y_{\text{new}} - y$.
- Algebraically this is equivalent to saying that $x_{\text{new}} = x + \Delta x$. Thus:

$$\Delta y = y_{\text{new}} - y = f(x_{\text{new}}) - f(x) = f(x+\Delta x) - f(x) \tag{11}$$

Differentiation

- But (11) is simply the numerator (top expression) of (10)! So we can rewrite (10) as:

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} \tag{12}$$

- This should look familiar to those that took calculus in year 12.
- Suppose we want to be even more precise than (12). Rather looking at the limit, I want to look at changes so small that they actually give a precise slope of the tangent line:

$$f'(x) = \frac{dy}{dx} \tag{13}$$

Differentiation

- We have gone through a lot of notation, so let's summarise:

$$f'(x) = \frac{dy}{dx} = \frac{d(x^2)}{dx} = \frac{d}{dx}(x^2) = 2x$$

- Example: Suppose we look at our car example again. $f(t) = 15t^2 + 7$. We know that $f'(t) = 30t$. This is the same as saying that $\frac{df(t)}{dt} = 30t$.

Differentiation

Let's find the derivative of the linear function $y = mx + c$:

$$
\begin{aligned}
f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{(m(x+h)+c) - (mx+c)}{h} \\
&= \lim_{h \to 0} \frac{mh}{h} = \lim_{h \to 0} m = m
\end{aligned}
$$

So the derivative of a linear function is always m.

Differentiation

- That's essentially it for the concept of differentiation and derivatives.
- That's roughly 50% of all of calculus. We won't look at the anti-differentiation – aka integration – but if you're interested, consider the fundamental theorem of calculus:

$$\int_a^b \frac{df(x)}{dx} dx = f(b) - f(a)$$

- From here we will be skipping a lot of the mechanics of differentiation – we simply do not have enough time to go through these.
- What I want to focus on is how differentiation can be useful for economics.

Differentiation

- Some differentiation rules:

$$\text{if } y = \frac{1}{x}, \quad \text{then, } \frac{dy}{dx} = -\frac{1}{x^2},$$

$$\text{if } y = \sqrt{x}, \quad \text{then, } \frac{dy}{dx} = -\frac{1}{\sqrt[2]{x}}$$

Differentiation

- Power Rule:

$$\text{if } y = x^a, \quad \text{then, } \frac{dy}{dx} = ax^{a-1}$$

where a is some constant. Furthermore:

$$\text{if } y = a, \quad \text{then, } \frac{dy}{dx} = 0$$

$$\text{if } y = x, \quad \text{then, } \frac{dy}{dx} = 1$$

Differentiation

- Log Rule:

$$\text{if } y = \log v(x), \quad \text{then, } \frac{dy}{dx} = \frac{v'(x)}{v(x)}$$

- Exponential Rule:

$$\text{if } y = e^{v(x)}, \quad \text{then, } \frac{dy}{dx} = v'(x)e^{v(x)}$$

Differentiation

- Product Rule:

$$\text{if } y = u(x)v(x), \quad \text{then, } \frac{dy}{dx} = u'(x)v(x) + u(x)v'(x)$$

or

$$\text{if } y = uv, \quad \text{then, } \frac{dy}{dx} = v\frac{du}{dx} + u\frac{dv}{dx}$$

Differentiation

- Quotient Rule:

$$\text{if } y = \frac{u(x)}{v(x)}, \quad \text{then, } \frac{dy}{dx} = \frac{u'(x)v(x) - u(x)v'(x)}{(v(x))^2}$$

or

$$\text{if } y = \frac{u}{v}, \quad \text{then, } \frac{dy}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}$$

Differentiation

- Chain Rule:

$$\text{if } y = f(g(x)), \quad \text{then, } \frac{dy}{dx} = f'(g(x))g'(x)$$

or

$$\frac{dy}{dx} = \frac{dy}{dx}\frac{du}{dx}$$

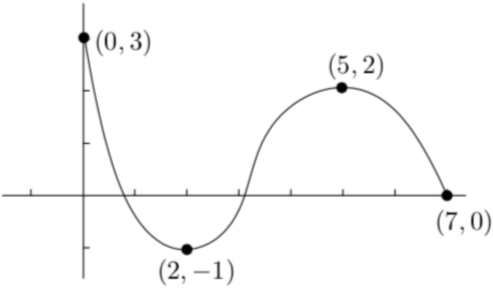
where y is a function of u, and u is a function of x.

Maxima and Minima

Definition

If we say that $x = a$ is an extremum of a function f, this means that f ha a maximum or minimum at $x = a$.

Figure: **Extrema**



Maxima and Minima

Definition

Suppose that f is defined on (a, b) and c is in (a, b). If c is a local maximum or minimum of f, then c must be a critical point for f. That is, either $f'(c) = 0$ or $f'(c)$ does not exist.

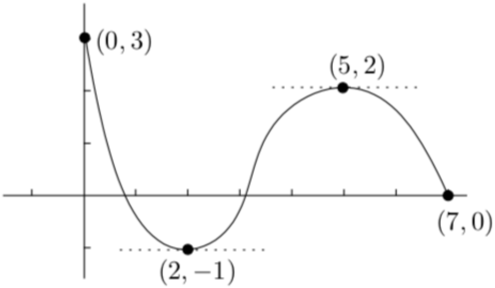
- Think about this definition for a moment. $f'(c) = 0$ is saying that the slope of a tangent line to a function is zero. What does this mean?
- It should be quite clear how or why maximum and minimums have applications in economics. Think of a firm's profit function, a consumer's welfare, policy functions, and so on.

Maxima and Minima

Theorem

If $x = c$ is a local maximum or minimum of $f(x)$ then $x = c$ is a critical point.

Figure: **Maxima and Minima**



Maxima and Minima (Single Variable Functions)

- Let's try find the maximum or minima of the following functions:

$$y = 5x^2$$

$$y = -6x^2 + 20x + 5$$

$$y = 5 \log x$$

$$y = x^3 + 10x^2$$

Unconstrained Optimisation: Functions of Many Variables

- While finding the derivative and the optimal point for functions of single variables ($y = f(x)$) is relatively easy, most functions that we will encounter in economics are that of many variables.
- That is, there are variables other than x that determine an output function such as y. i.e. $y = f(x_1, x_2, ..., x_n)$.
- Notable examples include the Cobb Douglas function:

$$y = Ax_1^{\alpha}x_2^{\beta}, \quad A, \alpha, \beta \text{ are constants}$$

and the constant elasticity of substitution (CES) production function:

$$y = (ax_1^{-\beta} + bx_2^{-\beta})^{-1/\beta}, \quad a, b > 0$$

Partial Differentiation

- Clearly our simple rules for single variable differentiation won't work for functions of many variables.
- However, we can extend our simple rules so that we can make them work.
- Recall that that the derivative for a function of a single variable is:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \tag{14}$$

- Now let $y = f(\mathbf{x}) = f(x_1, ..., x_n)$, then the partial derivative of y w.r.t. to x_i $(i = 1, ..., n)$ is:

$$\frac{\partial y}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, ..., x_i + \Delta x_i, ..., x_{in}) - f(x_1, ..., x_i, ..., x_n)}{\Delta x_i} \tag{15}$$

Partial Differentiation

- Comparing (14) with (15), we see that the definitions are essentially the same, so long as we treat all variables except x_i as constants.

Definition

The partial derivative of $y = f(\mathbf{x})$ taken w.r.t x_i is the derivative obtained when treating variables $x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n$ as constant.

- Notations for partial derivatives are:

$$\frac{\partial y}{\partial x_i}, \frac{\partial f}{\partial x_i}, f_{x_i}, f_i$$

Partial Differentiation

- Some Chain Rule examples. Suppose

$$C = C(Y^D(Y(G, T))) \tag{16}$$

C is a consumption of Y^D, Y^D is a function of Y, which is in turn a function of G and T. Then

$$\frac{\partial C}{\partial T} = \frac{dC}{dY^D}\frac{dY^D}{dY}\frac{\partial Y}{\partial T}$$

- Consider:

$$L = L(r(G, T), Y(G, T)) \tag{17}$$

$$\frac{\partial L}{\partial G} = \frac{\partial L}{\partial r}\frac{\partial r}{\partial G} + \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial G}$$

Partial Differentiation

Some more examples:

$$y = 3x_1^2 x_2 + 8x_2^{-5} x_1^3 \tag{18}$$
$$\frac{\partial y}{\partial x_1} = 6x_1 x_2 + 24x_2^{-5} x_1^2$$
$$\frac{\partial y}{\partial x_2} = 3x_1^2 - 40x_2^{-6} x_1^3$$

$$y = e^{x_1^3 + 2x_2^{-7}} \tag{19}$$
$$f_1 = 3x_1^2 e^{x_1^3 + 2x_2^{-7}}$$
$$f_2 = -14x_2^{-8} e^{x_1^3 + 2x_2^{-7}}$$

Partial Differentiation

- What does this all mean? Why do we need partial derivatives?
- Marginal analysis in economics essentially involves looking at rates of change in the dependent variable when there is a small change in one of the independent variables.
- This is the gist of unconstrained optimisation.
- i.e. Marginal analysis in economics involves partial derivatives.
- Consider for example the Cobb Douglas function $Q = Ax_1^{\alpha}x_2^{\beta}$, and suppose that this is a production function. Then the **marginal product** of input 1 is

$$\frac{\partial Q}{\partial x_1} = \alpha A x_1^{\alpha-1}x_2^{\beta}$$

Partial Differentiation

- Suppose $U(x_1, x_2) = \log x_1 + \log x_2$is a consumer's utility function then the **marginal utility** of good 2 is

$$\frac{\partial U}{\partial x_2} = \frac{1}{x_2}$$

- Neoclassical economists with their marginal analysis were actually reinventing differentiation.

Partial Differentiation: Second Order Partial Derivatives

- Clearly the partial derivatives themselves are functions of the x's. But we can take the partial derivatives of the partial derivatives to yield the second order partial derivatives:

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} = \frac{\partial}{\partial x_1}\left(\frac{\partial f(\mathbf{x})}{\partial x_1}\right) = f_{11}$$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} = \frac{\partial}{\partial x_2}\left(\frac{\partial f(\mathbf{x})}{\partial x_1}\right) = f_{12}$$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_1 x_2} = \frac{\partial}{\partial x_1}\left(\frac{\partial f(\mathbf{x})}{\partial x_2}\right) = f_{21}$$

Partial Differentiation: Second Order Partial Derivatives

Example: Consider the Cobb-Douglas function $Q = 4K^{3/4}L^{1/4}$. Then

$$\frac{\partial Q}{\partial K} = 3K^{-1/4}L^{1/4} = Q_K,$$
$$\frac{\partial Q}{\partial L} = K^{3/4}L^{-3/4} = Q_L.$$

$$\frac{\partial^2 Q}{\partial K^2} = \frac{\partial}{\partial K}\left(\frac{\partial Q}{\partial K}\right) = -\frac{3}{4}K^{-5/4}L^{1/4} = Q_{KK},$$
$$\frac{\partial^2 Q}{\partial L \partial K} = \frac{\partial}{\partial L}\left(\frac{\partial Q}{\partial K}\right) = \frac{3}{4}K^{-1/4}L^{-3/4} = Q_{KL},$$
$$\frac{\partial^2 Q}{\partial K \partial L} = \frac{\partial}{\partial K}\left(\frac{\partial Q}{\partial L}\right) = \frac{3}{4}K^{-1/4}L^{-3/4} = Q_{LK},$$
$$\frac{\partial^2 Q}{\partial L^2} = \frac{\partial}{\partial L}\left(\frac{\partial Q}{\partial L}\right) = -\frac{3}{4}K^{3/4}L^{-7/4} = Q_{LL}.$$

Partial Differentiation: Second Order Partial Derivatives

- Clearly the partial derivatives themselves are functions of the x's. But we can take the partial derivatives of the partial derivatives to yield the second order partial derivatives:

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} = \frac{\partial}{\partial x_1}\left(\frac{\partial f(\mathbf{x})}{\partial x_1}\right) = f_{11}$$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} = \frac{\partial}{\partial x_2}\left(\frac{\partial f(\mathbf{x})}{\partial x_1}\right) = f_{12}$$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_1 x_2} = \frac{\partial}{\partial x_1}\left(\frac{\partial f(\mathbf{x})}{\partial x_2}\right) = f_{21}$$

Convex and Concave Functions

- We technically already covered this. But let's quickly formalise some definitions.

Definitions

Let $y = f(\mathbf{x})$ be a function of many variables. Then $f(\mathbf{x})$ is a **convex** function on a convex set S in $\mathbb{R}^n$ if
$\lambda f(\mathbf{u}) + (1-\lambda)f(v) \geq f(\lambda\mathbf{u} + (1-\lambda)\mathbf{v})$, $0 \leq \lambda \leq 1 \; \forall \; \mathbf{u}, \mathbf{v}$
belonging to S.
The function $y = f(\mathbf{x})$ is **concave** on a set S contained in $\mathbb{R}^n$ if
$\lambda f(\mathbf{u}) + (1-\lambda)f(v) \leq f(\lambda\mathbf{u} + (1-\lambda)\mathbf{v})$, $0 \leq \lambda \leq 1 \; \forall \; \mathbf{u}, \mathbf{v}$
belonging to S

Comments

- That should be most of unconstrained optimisation.
- However, for the sake of time, we skipped a lot of material. I strongly recommend you to read up if you're uncertain on a few topics. In particular, we totally skipped discussion on set theory, hyperplanes, Hessian matrices, and matrix algebra.
- These concepts should be covered in future classes however. But, it's a good idea to get a head start now.