Lecture 7: Calculus

Robert Garrard

Taylor Series

There is a class of functions, called **analytic functions**, which have a very special property.

In a neighbourhood of some point, the function is described by a convergent power series.

Examples of analytic functions include all polynomials, the exponential and logarithm functions, trigonometric functions, and power functions.

Let $f : \mathbb{R} \to \mathbb{R}$ be an infinitely differentiable function. The **Taylor series** about some point $a \in \mathbb{R}$ is the power series

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n$$

Taylor Series

Example

Determine the Taylor series about $a = 0$ for the function

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) = e^x$$

Since the derivative of the exponential function is itself, $f^{(n)}(x) = e^x$. The nth derivative about $a = 0$ is therefore $e^0 = 1$. The Taylor series for e^x about zero is

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}...$$

Taylor Series

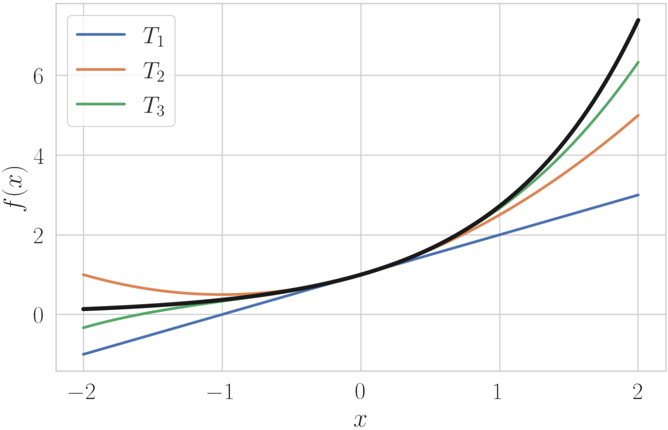


Figure: 3 Taylor polynomials of e^x

Taylor Series

Exercise 7.1 Find the Taylor expansion of $f(x) = \log(1 + x)$ about $x = 0$.

Taylor Series

Theorem

Let $f : \mathbb{R} \to \mathbb{R}$ be k times differentiable at the point $a \in \mathbb{R}$. Then there exists a function $R_k : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = f(a) + f^{'}(a)\,(x-a) + \frac{f^{''}(a)}{2}\,(x-a)^2 + \cdots + \frac{f^{(k)}}{k!}\,(x-a)^k + R_k(x)$$

such that

$$lim_{x \to a} R_k(x)/x^k = 0$$

Extrema

Consider a function $f : U \to \mathbb{R}$, where U is an interval in $\mathbb{R}$. A point $x_0 \in U$ is a

Local Maximum if there exists an interval $V \subset U$ containing x_0 such that $\forall a \in V \quad f(x_0) \geq f(a)$.

Local Minimum if there exists an interval $V \subset U$ containing x_0 such that $\forall a \in V \quad f(x_0) \leq f(a)$.

Global Maximum if $\forall a \in U \quad f(x_0) \geq f(a)$.

Global Minimum if $\forall a \in U \quad f(x_0) \leq f(a)$.

First order conditions

Proposition

If $f : U \to \mathbb{R}$ is a differentiable function and $x_0 \in U$ is a local maximiser/minimiser of f, then

$$f^{'}(x_0) = 0$$

First order conditions

Proof: Assume that $x_0 \in U$ is a local maximiser of f. Then there exists some interval V containing x_0 such that $f(x_0 + h) - f(x_0) \leq 0$ for all $x_0 + h \in V$. Consider the first order Taylor polynomial for f about x_0

$$f(x_0 + h) - f(x_0) = f^{'}(x_0) \cdot h + R_1(h)$$

where $R_1(h)/h \to 0$ as $h \to 0$.

Assume that $f^{'}(x_0) \neq 0$.Then we can find a sub-interval of $S \subset V$ containing x_0 such that

$$\left| \frac{R_1(h)}{h} \right| < |f^{'}(x_0)|$$

or

$$|f^{'}(x_0)h| - |R_1(h)| > 0$$

for all $x_0 + h \in S$.

First order conditions

Choose an h such that $f^{'}(x_0)h > 0$. Then by the above inequality

$$f(x_0 + h) - f(x_0) = f^{'}(x_0)h + R_1(h) > 0$$

Which contradicts x_0 being a local maximiser. So if x_0 is a local maximiser, $f^{'}(x_0) = 0$.
An identical argument applies for local minimisers. $\blacksquare$

First order conditiotns

This is known as a **first order necessary condition**.

We'd like some extra condition which guarantees that a point satisfying the first order condition will be a local max or min.

For this we have a **second order sufficient condition**.

Second order conditions

Proposition

Let $f : U \to \mathbb{R}$ be a twice differentiable function and $x_0 \in U$ be a point such that $f^{'}(x_0) = 0$. Then if

$$\begin{array}{ll} f^{''}(x_0) < 0 & x_0 \text{ is a local maximiser} \\ f^{''}(x_0) > 0 & x_0 \text{ is a local minimiser} \end{array}$$

Observe that the sufficient condition requires *strict* concavity/convexity.

The second derivative being *equal* to zero at the critical point is not sufficient for a max/min.

Global Extrema

These first and second order conditions give us the tools to find *local* maxima and minima.

However, when solving an optimisation problem, it's usually the *global* maxima and minima we care about.

Finding global maxima/minima can be quite tedious, often involving checking and comparing several different potential solutions (assuming a solution exists).

If the objective function is well behaved, our first and second order conditions (for a local optimum) will be sufficient for a global solution.

Global Extrema

Proposition

Let $f : U \to \mathbb{R}$ be a $\mathcal{C}^2$ function such that $f^{''}(x) < 0$ on all of U. If f has a critical point on $x_0 \in U$, then x_0 is a global maximiser. Respectively, if $f^{''}(x) > 0$, x_0 will be a global minimiser.

Exercise 7.2 Solve the following optimization problems:

1.
$$\max_{x \in \mathbb{R}} \quad -x^2 + 10x + 5$$

2.
$$\max_{x \in (0,1)} \quad \log x$$

Global Extrema

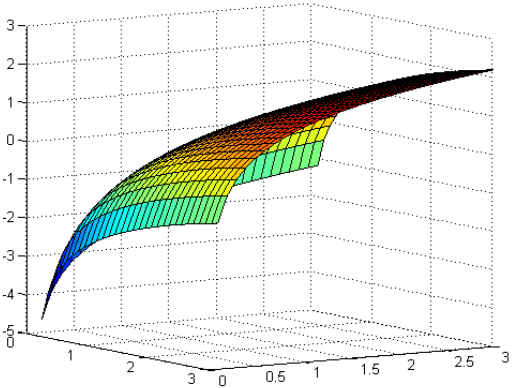
Exercise 7.3 An agent lives for two periods. At the beginning of time, the agent is endowed with wealth w. The agent consumes out of wealth in each period, with period utility represented by $U(c) = \log c$. Second period utility is discounted by a factor $\beta \in (0, 1)$. The price of the consumption good in each period is normalised to one unit of wealth. The agent chooses consumption for each period to maximise lifetime utility.

Set up and solve the agent's problem, interpreting the first order condition in words.

Multivariable Calculus

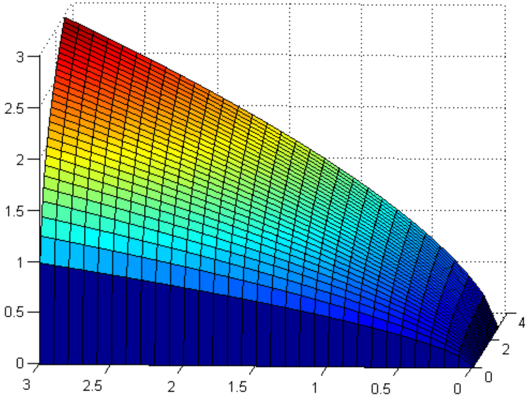
Now we will be considering functions mapping $\mathbb{R}^n$ into $\mathbb{R}$. You should already be familiar with some functions from $\mathbb{R}^2$ to $\mathbb{R}$.

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad f(c_1, c_2) = \log c_1 + \log c_2$$



Multivariable Calculus

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad f(K, L) = K^{\frac{1}{3}} L^{\frac{2}{3}}$$



Level Sets

A **level set** of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the set

$$\mathcal{L} = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = c\}$$

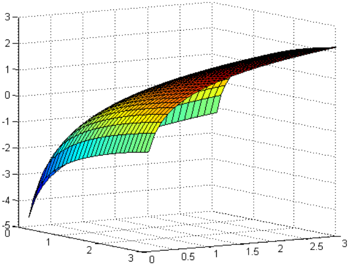
for some $c \in \mathbb{R}$.

Level sets for functions of two variables are often called **level curves** or **contours**.

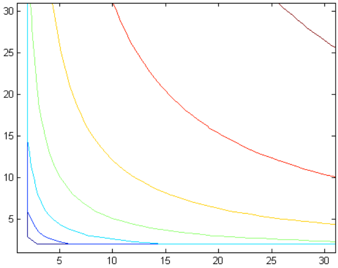
Familiar level sets from economics are the "indifference curves" of a utility function, the "isoquants" of a production function, and the "iso-profit lines" of a profit function.

Level Sets

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad f(c_1, c_2) = \log c_1 + \log c_2$$



(a) Surface



(b) Contours

Partial Derivatives

Derivatives tell us about the rate of change of a function. However, for functions of several variables, the rate of change may be different depending on what direction we wish to move.

So let's look at the simple case where we're only changing one variable at a time.

The derivative when we only change one variable, keeping all others constant, is called the **partial derivative**.

Partial Derivatives

Recall that the derivative of a function of one variable at some point x_0 is

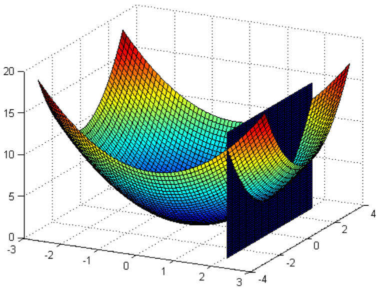
$$\frac{df}{dx}(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

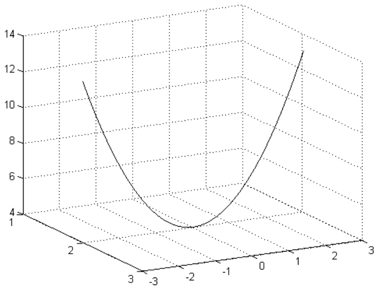
The partial derivative of a function $f(x_1, x_2, \ldots, x_n)$ with respect to x_i at a point $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ is defined similarly.

$$\frac{\partial f}{\partial x_i}(x_1^0, x_2^0, \ldots, x_n^0) = \lim_{h \to 0} \frac{f(x_1^0, \ldots, x_i^0 + h, \ldots x_n^0) - f(x_1^0, \ldots, x_i^0, \ldots, x_n^0)}{h}$$

if this limit exists.

Partial Derivatives





Chain Rules

Theorem

Let $f\left(x_1(t), \ldots, x_n(t)\right) : \mathbb{R}^n \to \mathbb{R}$ *be a differentiable function.*

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial x_1}\frac{\mathrm{d}x_1}{\mathrm{d}t} + \frac{\partial f}{\partial x_2}\frac{\mathrm{d}x_2}{\mathrm{d}t} + \cdots + \frac{\partial f}{\partial x_n}\frac{\mathrm{d}x_n}{\mathrm{d}t}$$

Theorem

Let $f\left(x_1(u,v), \ldots, x_n(u,v)\right) : \mathbb{R}^n \to \mathbb{R}$ *be a differentiable function*

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial u} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u}$$

Total Differential

It is often useful to refer to the **total differential**, which looks like the chain rule but using differentials rather than derivatives.

$$\mathrm{d}f = \frac{\partial f}{\partial x_1}\mathrm{d}x_1 + \cdots + \frac{\partial f}{\partial x_n}\mathrm{d}x_n$$

Total Differential

Example

For a function $f(x_1, x_2) : \mathbb{R}^2 \to \mathbb{R}$, find the slope of the tangent to a level curve of f (assuming $f_{x_2} \neq 0$).
Recall that a level curve sets the function to a constant level

$$f(x_1, x_2) = k$$

Now totally differentiate both sides

$$\begin{aligned} f_{x_1}\mathrm{d}x_1 + f_{x_2}\mathrm{d}x_2 &= 0 \\ f_{x_2}\mathrm{d}x_2 &= -f_{x_1}\mathrm{d}x_1 \\ \frac{\mathrm{d}x_2}{\mathrm{d}x_1} &= -\frac{f_{x_1}}{f_{x_2}} \end{aligned}$$

Gradient

The **gradient** is a vector containing all the first partial derivatives. The gradient vector is written interchangeably as $Df(\mathbf{x})$ or $\nabla f(\mathbf{x})$.

Let $f : \mathbb{R}^n \to \mathbb{R}$.

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

Exercise 7.4 Let $f : \mathbb{R}^3 \to \mathbb{R}$ and $f(x, y, z) = x^2 + xyz + y^2 - z^3$. Find the gradient.

Gradient

Proposition

Let $f : \mathbb{R}^2 \to \mathbb{R}$ and let $\mathbf{x}$ be a regular point of f, then the tangent to the level curve at $\mathbf{x}$ is orthogonal to the gradient at $\mathbf{x}$.

Exercise 7.5 *Proof:*

Directional Derivative

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $\mathcal{C}^1$ function on a ball about $\mathbf{x}$.

Let $\mathbf{v} \in \mathbb{R}^n$ be a unit vector.

The derivative of f at $\mathbf{x}$ in the direction $\mathbf{v}$ is

$$D_v f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{v}$$

Directional Derivative

Recall that

$$a \cdot b = ||a|| \, ||b|| \, \cos\theta$$

The rate of increase in some direction $\mathbf{v}$ at a point $\mathbf{x}$ is

$$D_{\mathbf{v}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{v} = ||\nabla f(\mathbf{x})|| \, ||\mathbf{v}|| \, \cos\theta$$

Recall that $-1 \leq \cos\theta \leq 1$. Since the length of $\mathbf{v}$ is fixed at unity, the derivative is maximised by choosing a $\mathbf{v}$ such that $\cos\theta = 1$.

This occurs when the angle between $\mathbf{v}$ and the gradient is $\theta = 0$, or when $\mathbf{v}$ points in the same direction as $\nabla f(\mathbf{x})$.

The Hessian

Similar to how the gradient is a vector containing all the first order partials of a function, we may want to consider a matrix full of the second order partials. This is called the **Hessian** matrix and is written $D^2 f(\mathbf{x})$.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice differentiable function. The Hessian is

$$D^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Note that if f is a twice *continuously* differentiable function, the Hessian will be a symmetric matrix.

Taylor Series in $\mathbb{R}^n$

Theorem

Let F be a $\mathcal{C}^2$ function defined on an open ball in $\mathbb{R}^n$. For any point $\mathbf{a}$ in the ball there exists a $\mathcal{C}^2$ function $R_2(\mathbf{h};\mathbf{a})$ such that for any point $\mathbf{a}+\mathbf{h}$ in the ball

$$F(\mathbf{a}+\mathbf{h}) = F(\mathbf{a}) + DF(\mathbf{a})\mathbf{h} + \frac{1}{2}\mathbf{h}^T D^2F(\mathbf{a})\mathbf{h} + R_2(\mathbf{h};\mathbf{a})$$

where

$$\frac{R_2(\mathbf{h};\mathbf{a})}{||\mathbf{h}||^2} \to 0 \text{ as } \mathbf{h} \to 0$$

The key difference between Taylor expansions for single variable vs multivariable functions is that a multivariable expansion isn't just a function of the first k derivatives, but is also a function of the *cross partials*.

Steady State

As we'll see in the coming lectures, our optimization problems will yield a description of a dynamic system.

For example, the Solow growth model assumes a production function that uses capital, $F(k_t)$, a constant savings rate s, and a capital depreciation rate δ.

$$k_{t+1} = sF(k_t) + (1 - \delta)k_t$$

Some dynamic models admit a **steady state**, where the variables become time invariant.

$$\bar{k} = sF(\bar{k}) + (1 - \delta)\bar{k}$$

Log-linearization

In log-linearizing an equation, we want to transform it from being in levels to being in percentage deviation from the steady state

$$\hat{x}_t = \frac{x_t - \bar{x}}{\bar{x}}$$

For most equations, we won't be able to write it like this exactly. We'll need to write it approximately:

$$\begin{aligned} f(x_t) &\approx f(\bar{x}) + f'(\bar{x})(x_t - \bar{x}) \\ &= f(\bar{x}) + f'(\bar{x})\bar{x}\hat{x}_t \end{aligned}$$

Log-linearization

We can do the same Taylor expansion for multivariable functions:

$$
\begin{aligned}
f(x_t, y_t) &\approx f(\bar{x}, \bar{y}) + f_x(\bar{x}, \bar{y})(x_t - \bar{x}) + f_y(\bar{x}, \bar{y})(y_t - \bar{y}) \\
&= f(\bar{x}, \bar{y}) + f_x(\bar{x}, \bar{y})\bar{x}\hat{x}_t + f_y(\bar{x}, \bar{y})\bar{y}\hat{y}_t
\end{aligned}
$$

Log-linearization

Exercise 7.6 Log-linearize the following equations around their steady state:

1. $Y_t = C_t + I_t$
2. $k_{t+1} = sk_t^{\alpha} + (1-\delta)k_t$

Log-linearization

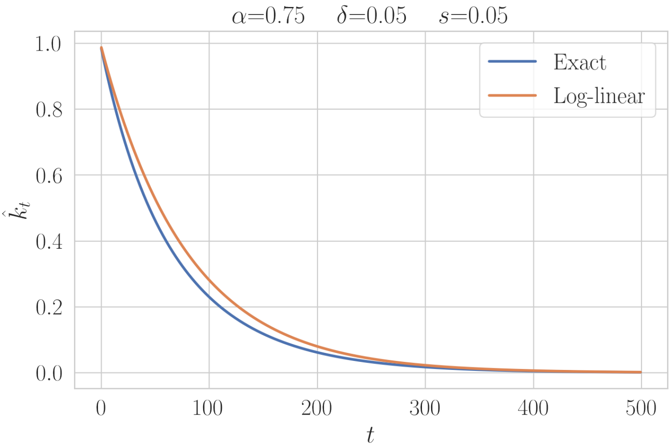


Figure: Solow model: Exact vs Log-linearized

Learning Outcomes

You should be able to:

- Expand a function into its Taylor series.
- Verify first order necessary and second order sufficient conditions for univariate functions.
- Take partial derivatives.
- Construct the gradient vector.
- Construct the Hessian.
- Log-linearize an equation.