Chapter 7
Rules of Differentiation
& Taylor Series

Isaac Newton and Gottfried Leibniz

7.1 Review: Derivative and Derivative Rules

• Review: Definition of derivative.

\[ f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{dy}{dx} \]

• Applying this definition, we review the 9 rules of differentiation.

First, the constant rule:

The derivative of a constant function is zero for all values of \( x \).

\[ y = f(x) = k \quad \Rightarrow \quad \frac{dy}{dx} = \frac{d}{dx} k = 0 \]

Define \( \frac{dy}{dx} = f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \)

If \( f(x) = k \) then \( f(x + \Delta x) = k \)

\[ \lim_{\Delta x \to 0} \frac{k - k}{\Delta x} = \lim_{\Delta x \to 0} 0 = 0 \]
7.1 A Review of the 9 Rules of Differentiation

1) \( \frac{d}{dx} k = 0 \) (Constant function)

2) \( \frac{d}{dx} x^n = nx^{n-1} \) (Power function)

3) \( \frac{d}{dx} [f(x) \pm g(x)] = f'(x) \pm g'(x) \) (Sum - difference )

4) \( \frac{d}{dx} [f(x)g(x)] = g(x)f'(x) + f(x)g'(x) \) (Product rule)

5a) \( \frac{d}{dx} cx = c \) (Constant and product rules )
\[ \frac{d}{dx} cx = (x)(0) + (c)(x^0) = c \]

5b) \( \frac{d}{dx} cx^n = cnx^{n-1} \) (Constant, product, & power rules)
\[ \frac{d}{dx} cx^n = (x^n)(0) + (c)(nx^{n-1}) = cnx^{n-1} \]

6) \( \frac{d}{dx} \left[ \frac{f(x)}{g(x)} \right] = \frac{g(x)f'(x) - f(x)g'(x)}{g^2(x)} \) (Quotient rule)

7 - 8: Let \( z = f(g(x)) \)

7) \( \frac{df(g(x))}{dx} = \frac{dz}{dy} \frac{dy}{dx} = f'(y)g'(x) \) (Univariate Chain rule)

8) \( \frac{df(g(x_1,...,x_n))}{dx_1} \bigg|_{x_2,...,x_n=0} = \frac{dz}{dy} \frac{dy}{d\bar{x}_1} \) (Multivariate Chain rule)

9) Let \( y = f(x), \ dy/dx = f'(x), \) and \( x = f^{-1}(y) \)
where \( y \) is a strictly monotonic function of \( x \)
\[ \frac{dx}{dy} = f^{-1}'(y) = \frac{1}{dy/dx} \] (Inverse - function rule)
### 7.1.1 Power-Function Rule

\[ f'(x) = \lim_{{\Delta x \to 0}} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{{\Delta x \to 0}} \frac{(x+\Delta x)^n - x^n}{\Delta x} = \lim_{{\Delta x \to 0}} \frac{(x^n + nx^{n-1}\Delta x + (n-1)x^{n-2}\Delta x^2 + \cdots + nx\Delta x^{n-1} + \Delta x^n) - x^n}{\Delta x} \]

\[ = \lim_{{\Delta x \to 0}} \frac{x^n}{\Delta x} + nx^{n-1} + (n-1)x^{n-2}\Delta x + \cdots + nx\Delta x^{n-2} + \Delta x^{n-1} - x^n/\Delta x \]

\[ = nx^{n-1} \]

\[ \lim_{{\Delta x \to 0}} \frac{(x+\Delta x)^n - x^n}{\Delta x} = nx^{n-1} \]

**Example.** Let Total Revenue (R) be:

\[ R = 15Q - Q^2 \implies \frac{dR}{dQ} = MR = 15 - 2Q. \]

As Q increases R increases (as long as Q > 7.5).

### 7.1.2 Exponential-Function Rule

\[ f'(x) = \lim_{{\Delta x \to 0}} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{{\Delta x \to 0}} \frac{e^{\kappa(x+\Delta x)} - e^{\kappa x}}{\Delta x} = e^{\kappa x} \lim_{{\Delta x \to 0}} \frac{(e^{\kappa \Delta x} - 1)}{\Delta x} \]

Definition of \( e \): \( e \) unique positive number for which \( \lim_{{h \to 0}} \frac{(e^h - 1)}{h} = 1 \)

Let \( h = \kappa \Delta x \). Then, \( \lim_{{h \to 0}} \frac{(e^h - 1)}{h} = \kappa \)

Thus, \( \lim_{{\Delta x \to 0}} \frac{e^{\kappa(x+\Delta x)} - e^{\kappa x}}{\Delta x} = \kappa e^{\kappa x} \)

*Example: Exponential Growth*

\[ y = f(t) = e^{0.5t} \quad \frac{d}{dt} e^{0.5t} = 0.5e^{0.5t} \]
7.1.2 Exponential-Function Rule: Joke

- A mathematician went insane and believed that he was the differentiation operator. His friends had him placed in a mental hospital until he got better. All day he would go around frightening the other patients by staring at them and saying "I differentiate you!"
- One day he met a new patient; and true to form he stared at him and said "I differentiate you!", but for once, his victim's expression didn't change.
- Surprised, the mathematician collected all his energy, stared fiercely at the new patient and said loudly "I differentiate you!", but still the other man had no reaction. Finally, in frustration, the mathematician screamed out "I DIFFERENTIATE YOU!"
- The new patient calmly looked up and said, "You can differentiate me all you like: I'm e^x."

7.2.1 Sum or difference rule

3) \[ \frac{d}{dx} [f(x) \pm g(x)] = f'(x) \pm g'(x) \]

The derivative of a sum (or difference) of two functions is the sum (or difference) of the derivatives of the two functions.

Example: \[ C = Q^3 - 4Q^2 + 10Q + 75 \]
\[ \frac{dC}{dQ} = \frac{d}{dQ}Q^3 - \frac{d}{dQ}4Q^2 + \frac{d}{dQ}10Q + \frac{d}{dQ}75 \]
\[ \frac{dC}{dQ} = 3Q^2 - 8Q + 10 + 0 \]
7.2.2 Product rule

4) \[ \frac{d}{dx} [f(x)g(x)] = g(x)f'(x) + f(x)g'(x) \]

The derivative of the product of two functions is equal to the second function times the derivative of the first plus the first function times the derivative of the second.

Example: Marginal Revenue

\[ R = PQ \quad P = 15 - Q \quad R = (15 - Q)Q \]

\[ \frac{dR}{dQ} = Q \frac{dP}{dQ} + P \frac{dQ}{dQ} = Q (-1) + (15 - Q)(1) = 15 - 2Q \]

Check: \[ R = 15Q - Q^2 \Rightarrow \frac{dR}{dQ} = 15 - 2Q \]

7.2.4 Quotient rule

6) \[ \frac{d}{dx} \left[ \frac{f(x)}{g(x)} \right] = \frac{g(x)f''(x) - f'(x)g'(x)}{g(x)^2} \]

Example:

\[ TC = C(Q) \quad \text{Total cost} \]
\[ AC = \frac{C(Q)}{Q} \quad \text{Average cost} \]

\[ \frac{d}{dQ} \frac{C(Q)}{Q} = \frac{Q \cdot C''(Q) - C(Q) \cdot 1}{Q^2} = \frac{1}{Q} \left[ C''(Q) - \frac{C(Q)}{Q} \right] = \frac{1}{Q} [MC - AC] \]

if \( \frac{d}{dQ} \frac{C(Q)}{Q} = 0 \), then \( AC = MC \) (Average Cost = Marginal Cost)
7.3.1 Chain rule

This is a case of two or more differentiable functions, in which each has a distinct independent variable, where \( z = f(g(x)) \). That is, \( z = f(y) \), i.e., \( z \) is a function of variable \( y \) and \( y = g(x) \), i.e., \( y \) is a function of variable \( x \).

7) \[
\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}
\]
\[
\frac{df(y)}{dx} = \frac{df(y)}{dy} \frac{dy}{dx} = f'(y)g'(x)
\]

Example: \( R = f(Q) \) (revenue) & \( Q = g(L) \) (output)

\[
\frac{dR}{dL} = \frac{dR}{dQ} \frac{dQ}{dL} = f'(Q)g'(L) = MR \cdot MPP_L = MRP_L
\]
7.3.1 Chain rule and its relation to total differential

Find $dz/dx_1$, where $z = f(y)$ and $y = g(x_1, x_2)$.

Algorithm: Substitute the total differential of $y$ into that of $z$ and divide through by $dx_1$ assuming $dx_2 = 0$

1) $dz = \frac{dz}{dy} dy$
2) $dy = \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2$
3) $dz = \frac{dz}{dy} \left( \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2 \right)$
4) $\frac{dz}{dx_1} \bigg|_{dx_2=0} = \frac{dz}{dy} \frac{\partial y}{\partial x_1}$

Note: A monotonic function is one in which a given value of $x$ yields a unique value of $y$, and given a value of $y$ will yield a unique value of $x$ (a one-to-one mapping). These types of functions have a defined inverse.

7.3.2 Inverse function rule

- Let $y=f(x)$ be a differentiable strictly monotonic function.

$$dx/dy = \frac{1}{dy/dx}$$

Note: A monotonic function is one in which a given value of $x$ yields a unique value of $y$, and given a value of $y$ will yield a unique value of $x$ (a one-to-one mapping). These types of functions have a defined inverse.

- Example:

  $Q = f(P)$
  $P = f^{-1}(Q)$

  $\frac{dQ_s}{dP} = b_l$
  $P = -b_0/b_l + (1/b_l) Q_s$ (where $b_l > 0$)
  $dP/dQ_s = 1/b_l = 1/[dQ_s/dP]$
7.3.2 Inverse-function rule

- This property of one-to-one mapping is unique to the class of functions known as monotonic functions.
- Recall the definition of a function
  - function: one y for each x
  - monotonic function: one x for each y (inverse function)
- if \( x_1 > x_2 \) \( \Rightarrow f(x_1) > f(x_2) \) monotonically increasing
  - \( Q_s = b_0 + b_1P \) supply function (where \( b_1 > 0 \))
  - \( P = -b_0/b_1 + (1/b_1)Q_s \) inverse supply function
- if \( x_1 > x_2 \) \( \Rightarrow f(x_1) < f(x_2) \) monotonically decreasing
  - \( Q_d = a_0 - a_1P \) demand function (where \( a_1 > 0 \))
  - \( P = a_0/a_1 - (1/a_1)Q_d \) inverse demand function

7.4 Second and Higher Derivatives

**Derivative of a derivative**

- Given \( y = f(x) \)
- The first derivative \( f'(x) \) or \( dy/dx \) is itself a function of \( x \); it should be differentiable with respect to \( x \); provided that it is continuous and smooth.
- The result of this differentiation is known as the second derivative of the function \( f \) and is denoted as \( f''(x) \) or \( d^2y/dx^2 \).

- The second derivative can be differentiated with respect to \( x \) again to produce a third derivative:

\[ f'''(x) \text{ and so on to } f^{(n)}(x) \text{ or } d^n y/dx^n \]

- This process can be continued to produce an \( n \)-th derivative.
7.4 Example: 1st, 2nd and 3rd derivatives

1) \( R = f(Q) = 1200Q - 2Q^2 \)  
   primitive function

2) \( f'(Q) = 1200 - 4Q \)
   1st derivative

3) \( f''(Q) = -4 \)
   2nd derivative

4) \( f'''(Q) = 0 \)
   3rd derivative

Graphically:

1) \( R = 1200Q - 2Q^2 \)

2) \( MR = 1200 - 4Q \)
   \( 1200 - 4Q = 0 \)
   \( Q = 300 \)

3) \( MR' = -4 \)

7.4 Example: Higher Order Derivatives

1) \( y = f(x) = 4x^4 - x^3 + 17x^2 + 3x - 1 \)  
   primitive function

2) \( f'(x) = 16x^3 - 3x^2 + 34x + 3 \)
   1st derivative

3) \( f''(x) = 48x^2 - 6x + 34 \)
   2nd derivative

4) \( f'''(x) = 96x - 6 \)
   3rd derivative

5) \( f^{(4)}(x) = 96 \)
   4th derivative

6) \( f^{(5)}(x) = 0 \)
   5th derivative
7.4 Interpretation of the second derivative

- $f'(x)$ measures the rate of change of a function
  - e.g., whether the slope is increasing or decreasing

- $f''(x)$ measures the rate of change in the rate of change of a function
  - e.g., whether the slope is increasing or decreasing at an increasing or decreasing rate
  - how the curve tends to bend itself

- Utility functions are increasing in consumption $f'(x) > 0$. But they differ by the rate of change in $f'(x) > 0$; that is, they differ on $f''(x)$.
  - $f''(x) > 0$, increasing $f'(x) > 0$
  - $f''(x) = 0$, constant $f'(x) > 0$
  - $f''(x) < 0$, decreasing $f'(x) > 0$ (usual assumption)

7.4 Strict concavity and convexity

- Strictly concave: if we pick any pair of points M and N on its curve and joint them by a straight line, the line segment MN must lie entirely below the curve, except at points MN.

- A strictly concave curve can never contain a linear segment anywhere (if it does it’s just concave, not strictly concave).

- Test: if $f''(x)$ is negative for all $x$, then strictly concave.

- Strictly convex: if we pick any pair of points M and N on its curve and joint them by a straight line, the line segment MN must lie entirely above the curve, except at points MN.

- A strictly convex curve can never contain a linear segment anywhere (if it does it’s just convex, not strictly convex).

- Test: if $f''(x)$ is positive for all $x$, then strictly convex.
7.4 Concavity and Convexity: 😞 & 😊

- If $f''(x) < 0$ for all $x$, then strictly concave. => a global maxima

- If $f''(x) > 0$ for all $x$, then strictly convex. => a global minima

- Concave functions have valuable properties: critical points are global maxima, & the weighted sum of concave functions is also concave. A popular choice to describe an average utility and production functions.

Example: AP = Arrow-Pratt risk aversion measure = $-U''(w)/U'(w)$

Let $U(w) = \beta \ln(w)$ \quad ($\beta > 0$)

$U'(w) = \beta/w > 0$

$U''(w) = -\beta/w^2 < 0$

$AP = 1/w$ => as $w$ (wealth) increases, risk aversion decreases.
Figure 7.5 Logarithmic Utility Function

\[ U(x) = \beta \ln(x) \]

\[ U'(x) = \frac{\beta}{x} \]

\[ U''(x) = -\frac{\beta}{x^2} \]

Figure 7.7 Utility Functions for Risk-Averse and Risk-Loving Individuals

(a) Risk-Averse Individual

(b) Risk-Loving Individual
7.5 Series

- **Definition**: Series, Partial Sums and Convergence

Let \( \{a_n\} \) be an infinite sequence.

1. The formal expression \( \sum a_n \) is called an (infinite) series.
2. For \( N = 1, 2, 3, \ldots \), the expression \( S_n = \sum_{k=1}^{n} a_k \) is called the N-th partial sum of the series.
3. If \( \lim S_n \) exists and is finite, the series is said to converge.
4. If \( \lim S_n \) does not exist or is infinite, the series is said to diverge.

**Example**: \( \sum (1/2)^n = 1/2 + 1/4 + 1/8 + 1/16 + \ldots \) (an infinite series)

The 3rd, and 4th partial sums are, respectively: 0.875, and 0.9375.

The n-th partial sum for this series is defined as
\[
S_n = 1/2 + 1/2^2 + 1/2^3 + \ldots + 1/2^n
\]

Divide \( S_n \) by 2 and subtract it from the original one, we get:
\[
S_n - 1/2 S_n = 1/2 - 1/2^{n+1} \Rightarrow S_n = 2 \left(1/2 - 1/2^{n+1}\right)
\]

Then, \( \lim S_n = 1 \) (the infinite series converges to 1)

7.5 Series: Convergence

- A series may contain positive and negative terms, many of them may cancel out when added together. Hence, there are different modes of convergence: one mode for series with positive terms, and another mode for series whose terms may be negative and positive.

- **Definition**: Absolute and Conditional Convergence

A series \( \sum a_n \) converges absolutely if the sum of the absolute values \( \sum |a_n| \) converges.
A series converges conditionally, if it converges, but not absolutely.

**Example**: \( \sum (-1)^n = -1 + 1 - 1 + 1 \ldots \) => no absolute convergence

Conditional convergence? Consider the sequence of partial sums:
\[
S_n = -1 + 1 - 1 + 1 \ldots \quad \text{if } n \text{ is odd, and}
\]
\[
S_n = -1 + 1 - 1 + 1 \ldots \quad \text{if } n \text{ is even.}
\]

Then, \( S_n = -1 \) if \( n \) is odd and 0 if \( n \) is even. The series is divergent.
7.5 Series: Rearrangement

- Conditionally convergent sequences are rather difficult to work with. Several operations do not work for such series. For example, the commutative law. Since \( a + b = b + a \) for any two real numbers \( a \) and \( b \), positive or negative, one would expect also that changing the order of summation in a series should have little effect on the outcome.

- **Theorem**: Convergence and Rearrangement
  A series \( \sum_n a_n \) be an absolutely convergent series. Then any rearrangement of terms in that series results in a new series that is also absolutely convergent to the same limit.

Let \( \sum_n a_n \) be a conditionally convergent series. Then, for any real number \( c \) there is a rearrangement of the series such that the new resulting series will converge to \( c \).

7.5 Series: Absolute Convergent Series

- Absolutely convergent series behave just as expected.

- **Theorem**: Algebra of Absolute Convergent Series
  Let \( \sum_n a_n \) and \( \sum_n b_n \) be two absolutely convergent series. Then:
  1. The sum of the two series is again absolutely convergent. Its limit is the sum of the limit of the two series.
  2. The difference of the two series is again absolutely convergent. Its limit is the difference of the limit of the two series.
  3. The product of the two series is again absolutely convergent. Its limit is the product of the limit of the two series (*Cauchy Product*).
7.5 Series: Convergence Tests

- There are many tests for convergence or divergence of series. Here are the most popular in economics.

- **Divergence Test**
If the series \( \sum a_n \) converges, then \( \{a_n\} \) converges to 0. Equivalently:
If \( \{a_n\} \) does not converge to 0, then the series \( \sum a_n \) can not converge.

- **Limit Comparison Test**
Suppose \( \sum a_n \) and \( \sum b_n \) are two infinite series. Suppose also that 
\[ r = \lim |a_n/b_n| \exists \text{ and } 0 < r < \infty. \]
Then \( \sum a_n \) converges absolutely iff \( \sum b_n \) converges absolutely.

7.5 Series: Convergence Tests

- **p Series Test**
The series \( \sum (1/n^p) \) is called a **p Series**.
  if \( p > 1 \) the **p-series converges**
  if \( p \leq 1 \) the **p-series diverges**.

- **Alternating Series Test**
A series of the form \( \sum (-1)^n b_n \) with \( b_n \geq 0 \) is called **alternating series**. If \( \{b_n\} \) is decreasing and converges to 0, then the sum converges.

- **Geometric Series Test**
Let \( a \in \mathbb{R} \). The series \( \sum a^n \) is called **geometric series**. Then,
  if \( |a| < 1 \) the geometric series converges
  if \( |a| \geq 1 \) the geometric series diverges.
7.5 Series: Power Series

• **Definition:** Power Series

A function series of the form
\[ \sum_{n=0}^{\infty} a_n (x - c)^n = a_0 + a_1(x-c) + a_2(x-c)^2 + \ldots \]
is called a power series centered at \( c \).

That is, a power series is an infinite series of functions where each
term consists of a coefficient \( a_n \) and a power \((x-c)^n\).

**Examples:**
- \( \sum_{n=0}^{\infty} (n+1) x^n = 1 + 2x + 3x^2 + 4x^3 + \ldots \)
- \( \sum_{n=0}^{\infty} (-1)^n \frac{1}{2^n} x^n = 1 - \frac{1}{2} x + \frac{1}{4} x^2 - \frac{1}{8} x^3 + \frac{1}{16} x^4 + \ldots \)
- \( \sum_{n=0}^{\infty} (3n/2^n) (x-2)^n = \frac{3}{2} (x-2) + 6/4 (x-2)^2 + 9/8 (x-2)^3 + \ldots \)

**Properties:**

- The power series converges at its center, i.e. for \( x = c \)
- There exists an \( r \) such that the series converges absolutely and
  uniformly for all \( |x - c| \leq p \), where \( p < r \), and diverges for all \( |x - c| > r \).

\( r \) is called the **radius of convergence** for the power series and is given by:
\[ r = \lim \sup |a_n / a_{n+1}| \]

**Note:** It is possible for \( r \) to be zero --i.e., the power series converges
only for \( x = c \)-- or to be \( \infty \) -i.e., the series converges for all \( x \).

**Example:** \( \sum_{n=0}^{\infty} (3n/2^n) (x-2)^n \)

\[ r = \lim \sup |a_n / a_{n+1}| = \lim \sup \left| \frac{3n/2^n}{(3(n+1)/2^{n+1})} \right| \]
\[ = \lim \sup \left| \frac{n}{n+1} \times 2 \right| = 2 \]

\( \Rightarrow \) the series converges absolutely and uniformly on any subinterval
of \( |x - 2| < 2 \).
7.5 Series: Power Series

- Polynomials are relatively simple functions: they can be added, subtracted, and multiplied (but not divided), and, again, we get a polynomial. Differentiation and integration are particularly simple and yield again polynomials.

- We know a lot about polynomials (e.g. they can have at most \( n \) zeros) and we feel pretty comfortable with them.

- Power series share many of these properties. Since we can add, subtract, and multiply absolutely convergent series, we can add, subtract, and multiply (think Cauchy product) power series, as long as they have overlapping regions of convergence.

- Differentiating and integrating works as expected. Important result: Power series are infinitely often (lim sup) differentiable.

7.6 Taylor Series

- The Taylor series is a representation of a (infinitely differentiable) function as an infinite sum of terms calculated from the values of its derivatives at a single point, \( x_0 \). It may be regarded as the limit of the Taylor polynomials.

**Definition: Taylor Series**

Suppose \( f \) is an infinitely often differentiable function on a set \( D \) and \( \epsilon \in D \). Then, the series

\[
T_f(x, \epsilon) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\epsilon)}{n!} (x - \epsilon)^n
\]

is called the (formal) Taylor series of \( f \) centered at, or around, \( \epsilon \).

If \( \epsilon = 0 \), the series is also called MacLaurin Series.
7.6 Taylor Series: Remarks

- A Taylor series is associated with a given function \( f \). A power series contains (in principle) arbitrary coefficients \( a_n \). Therefore, every Taylor series is a power series but not every power series is a Taylor series.
- \( T_f(x, c) \) converges trivially for \( x = c \), but it may or may not converge anywhere else. In other words, the “\( r \)” of \( T_f(x, c) \) is not necessarily greater than zero.
- Even if \( T_f(x, c) \) converges, it may or may not converge to \( f \).

Example: A Taylor Series that does not converge to its function

\[
f(x) = \exp\left(-\frac{1}{x^2}\right)
\]

- If \( x \neq 0 \)
- If \( x = 0 \)

• The function is infinitely often differentiable, with \( f'(0)=0 \). \( T_g(x, 0) \) around \( c=0 \) has radius of convergence infinity.
• \( T_g(x, 0) \) around \( c = 0 \) does not converge to the original function \( (T_g(x, 0) =0 \) for all \( x \)).

7.6 Maclaurin Series: Power Series Derivation

\[
f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots + a_nx^n
\]

\( f(x) \) is the primitive function

\[
\begin{align*}
f'(x) &= a_1 + 2a_2x + 3a_3x^2 + \ldots + na_nx^{n-1} \\
n &= 1 \\
f''(x) &= 2a_2 + 6a_3x + \ldots + n(n-1)a_nx^{n-2} \\
n &= 2 \\
f'''(x) &= 6a_3 + \ldots + n(n-1)(n-2)a_nx^{n-3} \\
n &= 3 \\
&\vdots
\end{align*}
\]

\( f^n(x) = n(n-1)(n-2)(n-3)\ldots(3)(2)(1)a_n \)

Evaluating each function at \( c = 0 \), simplifying & solving for the coefficient \( t \)

\[
\begin{align*}
f(0) &= a_0 \\
f'(0) &= a_1 \\
f''(0) &= 2a_2 \\
f'''(0) &= 6a_3 \\
&\vdots
\end{align*}
\]

\[
\begin{align*}
a_0 &= f(0)/0! \\
a_1 &= f'(0)/1! \\
a_2 &= f''(0)/2! \\
a_3 &= f'''(0)/3! \\
&\vdots
\end{align*}
\]

\[
\begin{align*}
f^n(0) &= n(n-1)(n-2)(n-3)\ldots(3)(2)(1)a_n \\
&\rightarrow a_n = f^n(0)/n!
\end{align*}
\]

Substituting the value of the coefficient \( t \) into the primitive function

\[
f(x) \approx \frac{f(0)}{0!}x^0 + \frac{f'(0)}{1!}x^1 + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^n(0)}{n!}x^n
\]
7.6 Taylor Series: Taylor’s Theorem

Suppose $f \in C^{n+1}([a, b])$ - i.e., $f$ is $(n+1)$-times continuously differentiable on $[a, b]$. Then, for $c \in [a, b]$ we have:

$$f(x) = \frac{f(c)}{0!}(x-c)^0 + \frac{f'(c)}{1!}(x-c)^1 + \frac{f''(c)}{2!}(x-c)^2 + \ldots + \frac{f^{(n+1)}(c)}{n!}(x-c)^n + R$$

where $R_{n+1}(x) = \frac{1}{n!} \int_{c}^{x} f^{(n+1)}(p)(x-p)^n dp$

In particular, the $T_f(x, c)$ for an infinitely often differentiable function $f$ converges to $f$ iff the remainder $R_{n+1}(x)$ converges to 0 as $n \to \infty$.

- We can show that a function really has a Taylor series by checking to that the remainder goes to zero. Lagrange found an easier expression:

$$R_{n+1}(x) = \frac{f^{(n+1)}(p)}{(n+1)!}(x-c)^{n+1}$$

for some $p$ between $x$ and $c$.

7.6 Taylor Series: Taylor’s Theorem

- Implications:
  - A function that is $(n+1)$-times continuously differentiable can be approximated by a polynomial of degree $n$.
  - If $f$ is a function that is $(n+1)$-times continuously differentiable and $f^{(n+1)}(\infty) = 0$ for all $x$ then $f$ is necessarily a polynomial of degree $n$.
  - If a function $f$ has a Taylor series centered at $c$ then the series converges in the largest interval $(c-r, c+r)$ where $f$ is differentiable.

- In practice, a function is approximated by its Taylor series using a small $n$, say $n=2$:

$$f(x) = \frac{f(c)}{0!}(x-c)^0 + \frac{f'(c)}{1!}(x-c)^1 + \frac{f''(c)}{2!}(x-c)^2$$

with $R_2(x) = \frac{f^{(3)}(p)}{3!}(x-c)^3$

- The error (& the approximation) depends on the curvature of $f$. 

37

38
7.6 Taylor Series: Taylor’s Theorem - Example

• Taylor series expansion of a quadratic polynomial around \( c = 1 \).

\[
\begin{align*}
&f(x) = 5 + 2x + x^2 \\
&f'(x) = 2 + 2x \\
&f''(x) = 2 \\
&f'''(x) = 0
\end{align*}
\]

Taylor’s series formula:

\[
f(x) \approx f(c) + \frac{f'(c)}{1!}(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \ldots + \frac{f^{(n)}(c)}{n!}(x-c)^n
\]

For \( n = 1 \),

\[
f(x) = f(c) + \frac{f'(c)}{1!}(x-c)
\]

&

\[
R_2(x) = \frac{f^{(2)}(p)}{2!}(x-c)^2
\]

• First-order Taylor series:

\[
f(x) \approx 8 + 4(x-1) + R = 4 + 4x + R_2
\]

with

\[
R_2 = \frac{2}{2!}(x-1)^2 = (x-1)^2
\]

7.6 Taylor Series: Taylor’s Theorem - Example

• Let’s check the approximation error (\( R_2 = (x-1)^2 \)):

\[
\begin{array}{cccc}
x_0 = 1 & f(1) = 8 & f(1) \approx 8 & R_2 = 0 \\
x_0 = 1.1 & f(1.1) = 8.41 & f(1.1) \approx 8.4 & 0.1^2 \\
x_0 = 1.2 & f(1.2) = 8.84 & f(1.2) \approx 8.8 & 0.2^2
\end{array}
\]

• For \( n = 2 \):

\[
f(x) = 8 + \frac{4}{1!}(x-1) + \frac{2}{2!}(x-x_0)^2 + R_3
\]

\[
= 8 + 4(x-1) + 1(x-1)^2 + R_3
\]

\[
= 4 + 4x + x^2 - 2x + 1 + R_3
\]

\[
= 5 + 2x + x^2 + R_3
\]

\[
R_3 = 0 \Rightarrow \text{perfect fit}
\]

Note: Polynomial can be approximated with great accuracy.
7.6 Maclaurin Series of $e^x$

Let’s do a Taylor series around $c=0$:

\[
\begin{align*}
f(x) &= e^x & \text{primitive function} & \Rightarrow f(0) = e^0 = 1 \\
f'(x) &= e^x & \text{1st derivative} & \Rightarrow f'(0) = e^0 = 1 \\
f''(x) &= e^x & \text{2nd derivative} & \Rightarrow f''(0) = e^0 = 1 \\
f'''(x) &= e^x & \text{3rd derivative} & \Rightarrow f'''(0) = e^0 = 1 \\
\vdots & & \vdots & \\
F^n(x) &= e^x & \text{n th derivative} & \Rightarrow f^n(0) = e^0 = 1
\end{align*}
\]

Substituting the value of the coefficients into the primitive function

\[
e^x = \frac{1}{0!}(x)^0 + \frac{1}{1!}(x)^1 + \frac{1}{2!}(x)^2 + \ldots + \frac{1}{n!}(x)^n + \ldots = \sum_{n=0}^{\infty} \frac{1}{n!} x^n
\]

7.6 Maclaurin Series of $\cos(x)$

Let’s do a Taylor series around $c=0$:

\[
\begin{align*}
f(x) &= \cos(x) & \text{primitive function} & \Rightarrow f(0) = \cos(0) = 1 \\
f'(x) &= -\sin(x) & \text{1st derivative} & \Rightarrow f'(0) = -\sin(0) = 0 \\
f''(x) &= -\cos(x) & \text{2nd derivative} & \Rightarrow f''(0) = -\cos(0) = -1 \\
f'''(x) &= \sin(x) & \text{3rd derivative} & \Rightarrow f'''(0) = \sin(0) = 0 \\
f^{(4)}(x) &= \cos(x) & \text{4th derivative} & \Rightarrow f^{(4)}(0) = \cos(0) = 1 \\
\vdots & & \vdots & \\
\end{align*}
\]

Substituting the value of the coefficients into the primitive function

\[
\cos(x) = 1 - \frac{1}{2!}(x)^2 + \frac{1}{4!}(x)^4 + \ldots = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n!} x^{2n}
\]

• Now, let’s check if the remainder $R_{2[n+1]}$ goes to 0 as $n \rightarrow \infty$:

\[
R_{2n+2}(x) = \left| \frac{f^{(2n+2)}(p)}{(2n+2)!} (x-0)^{2n+2} \right| = \left| \frac{\cos(p)}{(2n+2)!} x^{2n+2} \right| \leq \left| \frac{x^{2n+2}}{(2n+2)!} \right|
\]

and the last term is a converging series to 0, as $n \rightarrow \infty$. 


## 7.6 Maclaurin Series of \(\sin(x)\) & Euler’s formula

Similarly, we can a Taylor series for \(\sin(x)\):

\[
\sin(x) = x - \frac{1}{3!}(x^3) + \frac{1}{5!}(x^5) - \frac{1}{7!}(x^7) + \ldots + \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}
\]

- Now, let’s go back to the Taylor series of \(e^x\). Let’s look at \(e^{ix}\):

\[
e^{ix} = \sum_{n=0}^{\infty} \frac{1}{n!} (ix)^n = 1 + ix + \frac{1}{2!}(ix)^2 + \frac{1}{3!}(ix)^3 + \frac{1}{4!}(ix)^4 + \frac{1}{5!}(ix)^5 + \ldots + \frac{1}{n!}(ix)^n + \ldots
\]

\[
= 1 + ix - \frac{1}{2!}(x^2) - i\frac{1}{3!}(x^3) + \frac{1}{4!}(x^4) + i\frac{1}{5!}(x^5) + \ldots
\]

\[
= \cos(x) + i\sin(x)
\]

This last result is called Euler’s formula.

(It will re-appear when solving differential equations with complex roots.)

## 7.6 Maclaurin Series of \(\log (1+x)\)

\[
f(x) = \log(1+x)
\]

\[
f'(x) = (1+x)^{-1}
\]

\[
f''(x) = -(1+x)^{-2}
\]

\[
f''(x) = 2(1+x)^{-3}
\]

\[
\vdots
\]

\[
f^{(n)}(x) = (-1)^{(n-1)}(n-1)! (1+x)^{-n}
\]

Evaluating each function at \(x_0 = 0\), simplifying & solving for the coefficient

\[
f(0) = 0 \quad \& \quad f'(0) = f''(0) = f'''(0) = \ldots = f^{(n)}(0) = 1
\]

Substituting the value of the coefficients into the primitive function

\[
\log(1+x) = \sum_{n=0}^{\infty} \frac{(-1)^{n-1}}{n!} x^n
\]
7.6 Maclaurin Series of log (1+x)

\[
\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots
\]

- A 1st order series expansion: \( \log(1+x) = x + O(x^2) \).
  **Notation:** \( O(x^2) \): R is bounded by \( Ax^2 \) as \( x \to 0 \) for some \( A < \infty \).

- Below, we plot \( \log(1+x) \) & its linear approximation \( x \). When \( x \) is small (say, \(|x| < 0.1\), typical of annual interest rates or monthly stock returns), the linear approximation works very well.

7.6 Taylor series: Approximations

- Taylor series work very well for polynomials; the exponential function \( e^x \) and the sine and cosine functions. (They are all examples of entire functions –i.e., \( f(x) \) equals its Taylor series everywhere).
- Taylor series do not always work well. For example, for the logarithm function, the Taylor series do not converge if \( x \) is far from \( x_0 \).
- Log approximation around 0:
7.6 Taylor Series: Newton Raphson Method

The Newton Raphson (NR) method is a procedure to iteratively find roots of functions, or a solution to a system of equations, where \( f(x) = 0 \). Very popular in numerical optimization, where \( f'(x) = 0 \).

To find the roots of a function \( f(x) \). We start with:

\[
\Delta y \approx \frac{\partial f(x)}{\partial x} \Delta x \quad \Rightarrow \quad \Delta x = \frac{\Delta y}{f'(x)}
\]

\[
\Rightarrow x_{k+1} = x_k + \frac{f(x_{k+1}) - f(x_k)}{f'(x)} = x_k - \frac{0 - f'(x_k)}{f'(x_k)} = x_k - \frac{f(x_k)}{f'(x_k)}
\]

We derived an iterative process to find the roots of a function \( f(x) \).

7.6 Taylor Series: NR Method

We can use the NR method to minimize a function.

Recall that \( f'(x^*) = 0 \) at a minimum or maximum, thus stationary points can be found by applying NR method to the derivative. The iteration becomes:

\[
x_{k+1} = x_k - \frac{f''(x_k)}{f'''(x_k)}
\]

We need \( f'''(x_k) \neq 0 \); otherwise the iterations are undefined. Usually, we add a step-size, \( \lambda_k \), in the updating step of \( x \):

\[
x_{k+1} = x_k - \lambda_k \frac{f'(x_k)}{f'''(x_k)}
\]
7.6 Taylor Series: NR – Quadratic Approximation

- When used for minimization, the NR method approximates $f(x)$ by its quadratic approximation near $x_k$.
- Expand $f(x)$ locally using a 2nd-order Taylor series:
  \[ f(x + \delta x) = f(x) + f'(x) \delta x + \frac{1}{2} f''(x) (\delta x)^2 + o(\delta x^2) \]
- Find the $\delta x$ which minimizes this local quadratic approximation:
  \[ \delta x = -\frac{f'(x)}{f''(x)} \]
- Update $x$: $x_{k+1} = x_k - \frac{f'(x)}{f''(x)}$.

7.6 Taylor Series: The Delta Method

- The **delta method** is used to obtain the asymptotic distribution of a non-linear function of a random variable (usually, an estimator). It uses a 1st-order Taylor series expansion and Slutsky’s theorem.
- Let $x_n$ be a RV, with plim $x_n = \theta$ and Var($x_n$) = $\sigma^2 < \infty$. We know the asymptotic distribution of $x_n$: $n^{1/2}(x_n - \theta)/\sigma \xrightarrow{d} N(0, 1)$
- But we want to know the asymptotic distribution of $g(x_n)$. (We assume $g(x_n)$ is continuously differentiable, independent of $n$.)

Steps:
1. Taylor series approximation around $\theta$:
   \[ g(x_n) = g(\theta) + g'(\theta) (x_n - \theta) + \text{higher order terms} \]
   We will assume the higher order terms vanish as $n$ grows.
7.6 Taylor Series: The Delta Method

(2) Use Slutsky theorem: \( \text{plim } g(x_n) = g(\theta) \)
\[ \text{plim } g'(x_n) = g'(\theta) \]

Then, as \( n \) grows,
\[ g(x_n) \approx g(\theta) + g'(\theta)(x_n - \theta) \]
\[ \Rightarrow n^{1/2}[g(x_n) - g(\theta)] \approx g'(\theta)[n^{1/2}(x_n - \theta)]. \]
\[ \Rightarrow n^{1/2}[g(x_n) - g(\theta)] / \sigma \approx g'(\theta)[n^{1/2}(x_n - \theta) / \sigma]. \]

If \( g(.) \) does not behave badly, the asymptotic distribution of \( (g(x_n) - g(\theta)) \) is given by that of \( [n^{1/2}(x_n - \theta)/\sigma] \), which is a standard normal. Then,
\[ n^{1/2}[g(x_n) - g(\theta)] \rightarrow N(0, [g'(\theta)]^2 \sigma^2). \]

After some work (“inversion”), we obtain:
\[ g(x_n) \rightarrow N(g(\theta), [g'(\theta)]^2 \sigma^2/n). \]

---

7.6 Taylor Series: The Delta Method – Example

Let \( x_n \rightarrow N(\theta, \sigma^2/n) \)

Q: \( g(x_n) = \delta / x_n \rightarrow ? \) \hspace{1cm} (\( \delta \) is a constant)

First, calculate the first two moments of \( g(x_n) \): \[
\begin{align*}
g(x_n) &= \delta / x_n \quad \Rightarrow \text{plim } g(x_n) = (\delta / \theta) \\
g'(x_n) &= -(\delta / x_n^2) \quad \Rightarrow \text{plim } g'(x_n) = -(\delta / \theta^2)
\end{align*}
\]

Recall delta method formula: \( g(x_n) \rightarrow N(g(\theta), [g'(\theta)]^2 \sigma^2/n). \)

Then,
\[ g(x_n) \rightarrow N(\delta / \theta, (\delta^2 / \theta^2) \sigma^2/n) \]
7.7 Geometric series

- Geometric series: Each term is obtained from the preceding one by multiplying it by \( x \), convergent if \(|x| < 1\).

Given \( f(x) = (1-x)^{-1} \). Find the first five terms of the Maclaurin series (\( n = 4 \)) around \( c = 0 \).

\[
\begin{align*}
\frac{f(x)}{1-x} &= f(c) + f'(c)(x-c) + \frac{f''(c)}{2!} (x-c)^2 + \frac{f'''(c)}{3!} (x-c)^3 + \frac{f^{(4)}(c)}{4!} (x-c)^4 \\
&= (1-x)^{-1} \quad \Rightarrow \quad f(c) = (1-0)^{-1} = 1 \\
f'(x) &= -1(1-x)^{-2} = (1-x)^{-2} \quad \Rightarrow \quad f'(c) = (1-0)^{-2} = 1 \\
f''(x) &= (-2)(1-x)^{-3} = 2(1-x)^{-3} \quad \Rightarrow \quad f''(c) = 2(1-0)^{-3} = 2 \\
f'''(x) &= (-3)(2)(1-x)^{-4} = 6(1-x)^{-4} \quad \Rightarrow \quad f'''(c) = 6(1-0)^{-4} = 6 \\
f^{(4)}(x) &= (-4)(3)(2)(1-x)^{-5} = 24(1-x)^{-5} \quad \Rightarrow \quad f^{(4)}(c) = 24(1-0)^{-5} = 24
\end{align*}
\]

\[
\begin{align*}
f(x) &= 1 + 1(x-0) + \frac{2}{2}(x-0)^2 + \frac{6}{6}(x-0)^3 + \frac{24}{24}(x-0)^4 + \ldots \\
f(x) &= 1 + x + x^2 + x^3 + x^4 + \ldots = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \Rightarrow \quad \sum_{n=0}^{\infty} ax^n = \frac{a}{1-x} \quad \Rightarrow \quad \sum_{n=0}^{\infty} ax^n = a \left( \frac{1}{1-x} \right)
\end{align*}
\]

7.7 Geometric series: Approximating \((1-a)^{-1}\)

- It is possible to accurately approximate some ratios, with \((1-x)\) term in the denominator, with a geometric series. Recall:

\[
y = (1-a)^{-1} \quad \text{where} \quad a = 0 \ldots 0.9
\]

\[
y = 1 + a + a^2 + a^3 + a^4 + \ldots
\]

- For \( n=4 \),
- \( a=0.1 \Rightarrow 1/(1-a) = 1.1111 \)
- \& \( 1+0.1+0.1^2+0.1^3+0.1^4 = 1.1111 \)
- \( a=0.9 \Rightarrow 1/(1-a) = 10 \)
- \& \( 1+0.9+0.9^2+0.9^3+0.9^4 = 4.0951 \)
### 7.7 Geometric series: Approximating $A^{-1}$

$$AA^{-1} = I \quad A'A^{-1} = A^0 \quad A^0 = I$$

Taylor series approximation

$$(1-x)^{-1} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + … + x^{n-1} \quad \text{if } |x| < 1 \text{ scalar algebra}$$

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n = I + A + A^2 + … + A^{n-1} \quad \text{by analogy}$$

- Application to the Leontief Model ($x = Ax + d \Rightarrow (I-A)x = d$)

$$A = \begin{bmatrix} .15 & 0.25 \\ 0.20 & 0.05 \end{bmatrix} \Rightarrow (I - A)^{-1} = \begin{bmatrix} 1.2541 & 0.3300 \\ 0.2640 & 1.1221 \end{bmatrix}$$

$$(I - A)^{-1} \approx \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.15 & 0.25 \\ 0.20 & 0.05 \end{bmatrix} + \begin{bmatrix} 0.15 & 0.25 \\ 0.20 & 0.05 \end{bmatrix}^2 = \begin{bmatrix} 1.1725 & 0.3125 \\ 0.2400 & 1.0525 \end{bmatrix}$$

with $n = 6$  

$$\begin{bmatrix} 1.2506 & 0.3290 \\ 0.2632 & 1.1214 \end{bmatrix}$$

### 7.7 Application: Geometric series & PV Models

A stock price ($P$) is equal to the discounted sum of all futures dividends. Assume dividends are constant ($d$) and the discount rate is $r$. Then:

$$P = \sum_{n=1}^{\infty} \frac{d}{(1+r)^n} = d \left( \frac{1}{1} + \frac{1}{(1+r)^2} + … + \frac{1}{(1+r)^n} + … \right)$$

Let $x = \frac{1}{1+r} \Rightarrow P = d \left( x + x^2 + … + x^n \right) = d \left( \frac{1}{1-x} - 1 \right)$

$$P = d \left( \frac{1}{1-x} - 1 \right) = d \left( \frac{1}{1+r-1} - 1 \right)$$

$$P = d \left( \frac{1+r}{r} - 1 \right) = d \left( \frac{1+r-r}{r} \right) = d \left( \frac{1}{r} \right) = \frac{d}{r}$$
Q: What is the first derivative of a cow?
A: Prime Rib!