Math Refresher

Basic Calculus

Introduction

- This is a refresher on basic calculus. It is not meant to be a substitute for a full course on calculus but rather a quick review of the fundamental concepts and techniques that will be used this semester.

Limits

The limit of a function $f(x)$ as x approaches a is the value that $f(x)$ approaches as x gets closer and closer to a. We write this as:

$$\lim_{x \to a} f(x) = L$$

Here, L is the limit of the function $f(x)$ as x approaches a.

For example, consider the function $f(x) = x^2$. The limit of $f(x)$ as x approaches 2 is 4:

$$\lim_{x \to 2} x^2 = 4$$

Limits to Derivatives

Limits can also be used to define derivatives. The derivative of a function $f(x)$ is the slope of the function at a given point. The derivative of $f(x)$ at $x = a$ is written as $f'(a)$. The derivative is defined as:

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

In other words, the derivative is the slope of the function at a particular point a. This can be approximated numerically by choosing a very small value for h.

For example, consider the function $f(x) = x^2$. The derivative of $f(x)$ at $x = a$ is:

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

If other methods fail, one can always rely on numerical differentiation.

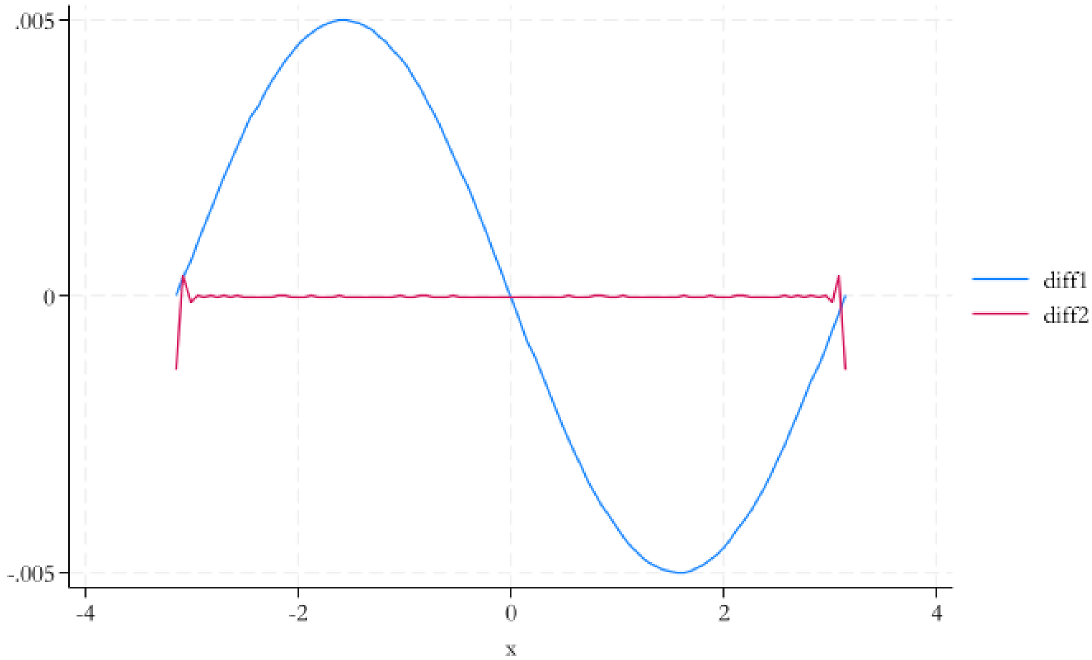
Stata and Numerical Differentiation

`Stata` can be used to calculate numerical derivatives. `mata` (matrix algebra language) has powerful rutines for numerical differentiation. `Stata` also has some capabilities, and you can always do it manually.

```
clear
range x -_pi _pi 100
gen y = sin(x)
gen dydx = (sin(x+0.01) - sin(x)) / 0.01
dydx y x, gen(dydx2)
gen dydx3 = cos(x)
gen diff1 = (dydx - dydx3)
gen diff2 = (dydx2 - dydx3)
line diff1 diff2 x
```

```
<IPython.core.display.HTML object>
```

```
Number of observations (_N) was 0, now 100.
```



Derivatives of Common Functions

For most common functions, the derivative can be calculated using the following rules:

- The derivative of a constant is zero.
- The derivative of x^n is nx^{n-1}.
- The derivative of $\ln(x)$ is $\frac{1}{x}$.
- The derivative of e^x is e^x.
- The derivative of a^x is $a^x \ln a$.

There are other rules for derivatives, but these are the ones that will be used most often.

Derivatives of Composite Functions

The derivative of a composite function $f(g(x))$ is given by the chain rule:

$$\frac{d}{dx} f(g(x)) = f'(g(x)) \cdot g'(x).$$

For example, consider the function $f(x) = \ln(x^2)$. The derivative of $f(x)$ is:

$$\begin{aligned} \frac{d}{dx} \ln(x^2) &= \frac{1}{x^2} \cdot \frac{d}{dx}(x^2) \\ &= \frac{1}{x^2} \cdot 2x \\ &= \frac{2}{x}. \end{aligned}$$

Derivatives of Sums and Products

The derivative of a sum of functions is the sum of the derivatives of the functions:

$$\frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx} f(x) + \frac{d}{dx} g(x).$$

The derivative of a product of functions is given by the product rule:

$$\frac{d}{dx}(f(x) \cdot g(x)) = f'(x) \cdot g(x) + f(x) \cdot g'(x).$$

The derivative of a quotient of functions is given by the quotient rule:

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x) \cdot g(x) - f(x) \cdot g'(x)}{g(x)^2}.$$

This is a special case of the product rule.

Maximization and Minimization

- Derivatives can be used to identify the maximum and minimum values of a function. Consider a function $f(x)$.
- To find the maximum (or minimum) value of $f(x)$, we take the derivative of $f(x)$ and set it equal to zero.
  - This is called the first-order condition.
  - **The idea** is that at the maximum (or minimum), the value of $f(x)$ shouldn't change anymore (it should be flat). Thus, the derivative of $f(x)$ should be zero.

For example, consider the function $f(x) = 5x^2 - 4x + 2$. The derivative of $f(x)$ is:

$$f'(x) = 10x - 4 = 0$$
$$x = \frac{4}{10} = 0.4.$$

So when x is equal to 0.4, the function $f(x)$ does not change anymore.

- This, however, is insufficient to determine whether the function is at a maximum or a minimum.

To determine this, we take the second derivative of $f(x)$, known as the second-order condition:

$$f''(x) = 10 > 0.$$

- Because the second derivative is positive, we know that $f(x)$ is at a minimum when $x = 0.4$.
  - If the second derivative were negative, we would know that $f(x)$ is at a maximum when $x = 0.4$.

Why is this the case?

- $f'(x)$ measures the changes in $f(x)$ along x. When $f'(x) = 0$, $f(x)$ is not changing anymore.
- $f''(x)$ measures the changes in $f'(x)$ (the changes in those changes).
  - Because it is positive, we know that $f'(x)$ is increasing. This means that at $x = 0.4$, the changes in $f(x)$ are going from negative to positive, indicating a minimum.

Optimization with Multiple Variables

When considering multiple variables, we also need to rely on the first- and second-order conditions to find minimum and maximum values. Consider a function $f(x, y)$. The first-order conditions are:

$$\frac{\partial}{\partial x} f(x, y) = 0,$$
$$\frac{\partial}{\partial y} f(x, y) = 0.$$

These conditions indicate that, in the direction of x and y, the function $f(x, y)$ is not changing anymore. Thus, we have a potential maximum or minimum. To identify a minimum, we need second-order conditions:

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

where H is the **Hessian** matrix.

- If $\text{Det}(H) > 0$ and $f_{xx} > 0$, then we have a minimum.
- If $\text{Det}(H) > 0$ and $f_{xx} < 0$, then we have a maximum.
- If $\text{Det}(H) < 0$, then we have a saddle point.
- If $\text{Det}(H) = 0$, the result is inconclusive.

Optimization with Constraints

When optimizing a function with constraints, we can use the method of Lagrange multipliers. Consider a function $f(x, y)$ subject to the constraint $g(x, y) = z$. The Lagrangian is:

$$L(x, y, \lambda) = f(x, y) + \lambda(z - g(x, y)).$$

- The Lagrangian is the function $f(x, y)$ plus the constraint $g(x, y)$ multiplied by a constant λ.
- The constant λ is called the Lagrange multiplier.
- The constraint is written as the difference between the constant z and the function $g(x, y)$.

- The Lagrangian is then optimized with respect to x, y, and λ.

These are the equivalent first-order conditions:

$$
\begin{aligned}
&\frac{\partial}{\partial x} L(x, y, \lambda) = 0, \\
&\frac{\partial}{\partial y} L(x, y, \lambda) = 0, \\
&\frac{\partial}{\partial \lambda} L(x, y, \lambda) = z - g(x, y) = 0.
\end{aligned}
$$

The last condition is the constraint, and it implies that the constraint must be satisfied. The second-order conditions are the same as before.