

Columbia MA Math Camp

Differential Calculus

Vinayak Iyer

August 26, 2019

Table of Contents

Single Variable

Multivariable Derivative

Second Derivatives and Taylor Series Expansions

The Implicit Function Theorem

The fundamental concept in calculus is that of the **derivative**, which is simply a rate of change. Consider a function $f : \mathbb{R} \rightarrow \mathbb{R}$. The quantity

$$\frac{f(x_0 + h) - f(x_0)}{h}$$

tells us the average rate of change of f between x_0 and $x_0 + h$.

The big idea with derivatives is simply that we let h go to 0.

Definition 1.1

Let $A \subset \mathbb{R}$. A function $f : A \rightarrow \mathbb{R}$ is said to be differentiable at x_0 iff the limit :

$$f'(x_0) \equiv \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists.

We define the **derivative of f at x_0** as $f'(x_0)$

Derivative as an approximation

There's another interpretation of the derivative: it is the best linear approximation of a function.

- Suppose you wanted to approximate $f(x)$ by a linear function $g(x) = mx + b$ around the point x_0
- A good approximation should have the following properties:
 - The functions should agree at x_0 : $g(x_0) = f(x_0)$. So $g(x_0 + h) = f(x_0) + mh$
 - The per-unit error should be small near x_0 :

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - g(x_0 + h)}{h} = 0$$

These properties combined imply $m = f'(x_0)$

Common functions :

- $\frac{d}{dx} c = 0$
- $\frac{d}{dx} x^n = nx^{n-1}$
- $\frac{d}{dx} e^x = e^x$
- $\frac{d}{dx} \log x = \frac{1}{x}$

Combining derivatives :

- $\frac{d}{dx} (f(x) + g(x)) = f'(x) + g'(x)$
- $\frac{d}{dx} (\alpha f(x)) = \alpha f'(x)$
- $\frac{d}{dx} (f(x)g(x)) = f'(x)g(x) + f(x)g'(x)$ (product rule)
- $\frac{d}{dx} (f(g(x))) = f'(g(x))g'(x)$ (chain rule)

Chain Rule example

- Consider a consumer whose utility u is only directly dependent on consumption c :
 $u = u(c)$
- However, consumption depends on the consumer's wealth: $c = c(w)$
- Therefore $u = u(c(w))$. We can capture the dependencies in a graph



- The chain rule tells us:

$$\frac{du}{dw} = \frac{du}{dc} \frac{dc}{dw}$$

Table of Contents

Single Variable

Multivariable Derivative

Second Derivatives and Taylor Series Expansions

The Implicit Function Theorem

- We will be working with functions from \mathbb{R}^n to \mathbb{R} a lot
- For functions of two variables, one common tool we'll use is **level curves**. This is the graph of the equation $f(x, y) = c$
- Example: let $a \in \mathbb{R}^2$ and consider the function $f(x, y) = a_1x + a_2y$. The graph of this function is called a **hyperplane**

Notion of Linear Approximation

In the single-variable setting, $f'(x)$ was the term in the **best linear approximation** of f

$$f(y) \approx f(x) + f'(y)(y - x)$$

where “best” meant the relative error goes to 0. That is, $f'(x)$ is the value m such that

$$\lim_{y \rightarrow x} \frac{f(y) - (f(x) + m(y - x))}{y - x} = 0$$

Notion of Linear Approximation (cont.)

If $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, then a *linear approximation* of f is:

$$f(\mathbf{y}) = f(\mathbf{x}) + \mathbf{A}(\mathbf{y} - \mathbf{x})$$

where \mathbf{A} is a $m \times n$ matrix. Thus we will define the **derivative of f at \mathbf{x}** or the **Jacobian of f at \mathbf{x}** as the matrix A such that

$$\lim_{\mathbf{y} \rightarrow \mathbf{x}} \frac{f(\mathbf{y}) - (f(\mathbf{x}) + \mathbf{A}(\mathbf{y} - \mathbf{x}))}{\|\mathbf{y} - \mathbf{x}\|} \rightarrow 0$$

Let us introduce the notion of a very useful concept of a **partial derivative** :

Definition 2.1

For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, it's **partial derivative** of the i th coordinate wrt to the j th argument at $x \in \mathbb{R}^n$ is :

$$\frac{\partial f_i}{\partial x_j}(x) := \frac{d}{dt} f_i(x + te_j)|_{t=0}$$

What's the partial derivative of $f(x) = x_1^2 + x_1x_2$?

Properties :

- If $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, then $f'(x)$ is an $m \times n$ matrix also known as the **Jacobian**.
 - If f is real-valued, the column vector $f'(x)^T$ is called the **gradient** of f or sometimes denoted as $\nabla f(x)$
- The derivative is **linear**. If $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\alpha \in \mathbb{R}$:
 - $(f + g)'(x) = f'(x) + g'(x)$
 - $(\alpha f)'(x) = \alpha f'(x)$
- If f is differentiable at x , f is continuous at x

The Chain Rule

- Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $g : \mathbb{R}^m \rightarrow \mathbb{R}^k$.
- Define $h : \mathbb{R}^n \rightarrow \mathbb{R}^k$ by $h(x) = g(f(x))$.
- If f is *differentiable* at x and g is *differentiable* at $f(x)$, then

$$h'(x) = g'(f(x))f'(x)$$

- (Heuristic proof) two linear approximations:

$$\begin{aligned} h(y) &= g(f(y)) \\ &\approx g[f(x) + f'(x)(y - x)] \\ &\approx g(f(x)) + g'(f(x))f'(x)(y - x) \\ &= h(x) + \underbrace{g'(f(x))f'(x)}_{h'(x)}(y - x) \end{aligned}$$

- Full proof in FMEA, page 96

What does $f'(x)$ look like?

- Consider the partial derivative wrt x_i i.e. :

$$\frac{\partial f(x)}{\partial x_i} = i\text{-th column of } f'(x)$$

- This allows us to calculate the Jacobian

$$f'(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

The Chain Rule: Interpretation

Returning to the chain rule, consider the following example:

- Utility u depends on consumption c and hours worked h
- However c and h depend on the going wage w . Define $x(w) : \mathbb{R} \rightarrow \mathbb{R}^2$ by $x(w) = (c(w), h(w))$.
- Define $v(w) = u(x(w))$. The chain rule says $v'(w) = u'(x(w))x'(w)$:

$$\begin{aligned} v'(w) &= \begin{pmatrix} \frac{\partial u}{\partial c} & \frac{\partial u}{\partial h} \end{pmatrix} \begin{pmatrix} \frac{\partial c}{\partial w} \\ \frac{\partial h}{\partial w} \end{pmatrix} \\ &= \frac{\partial u}{\partial c} \frac{\partial c}{\partial w} + \frac{\partial u}{\partial h} \frac{\partial h}{\partial w} \end{aligned}$$

Some common derivatives

Being comfortable taking vector derivatives in one step can save you a **lot** of algebra (especially in econometrics). You should know these identities by heart :

- Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $f(x) = Ax$ where A is an $m \times n$ matrix:

$$f'(x) = A$$

- Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with $f(x) = x'Ax$ where A is an $n \times n$ matrix:

$$f'(x) = x'(A + A')$$

If A is symmetric, $f'(x) = 2x'A$

- If $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ and $h(x) = f(x)g(x)$, then

$$h'(x) = f'(x)g(x) + f(x)g'(x)$$

Some technical concerns

- We have seen that if f is differentiable, its partials exist and the Jacobian is just the matrix of partial derivatives.
- What if we only know that the partials exist? Is that enough for differentiability?
- Sadly, the answer is no (idea: function could behave nicely along the axes, but misbehave along other directions)
- However, if the partials are also continuous, then the derivative exists. Almost every function we work with in economics will have continuous partial derivatives

Table of Contents

Single Variable

Multivariable Derivative

Second Derivatives and Taylor Series Expansions

The Implicit Function Theorem

Higher Derivatives: Single Variable

For a function $f : \mathbb{R} \rightarrow \mathbb{R}$, the **second derivative** of f at x is the derivative of f' at x

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

The second derivative measures the *change in the slope per unit change in x* :

- $f''(x) > 0$ means the derivative is (locally) increasing in x
- $f''(x) < 0$ means the derivative is (locally) decreasing in x

Can keep going to third, fourth derivatives, etc. (not commonly used)

Taylor Series: Single Variable

Suppose you want to approximate f by a polynomial around the point x_0

$$h(x) \equiv a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n$$

Two intuitive criteria for a “good” approximation are:

- $h(x_0) = f(x_0)$, which implies $a_0 = f(x_0)$
- The first n derivatives of h should match those of f at x_0

Differentiating repeatedly gives $h^k(x_0) = k!a_k$. Thus the **Taylor series expansion of order n of f around x_0** is

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \dots + \frac{1}{n!}f^n(x_0)(x - x_0)^n$$

Proposition 3.1

Let f be $k + 1$ times differentiable on $[a, x]$. Then

$$f(x) = f(a) + f'(a)(x - a) + \dots + \frac{f^k(a)}{k!}(x - a)^k + \frac{f^{k+1}(\zeta)}{(k + 1)!}(x - a)^{k+1}$$

for some $\zeta \in (a, x)$.

- Taylor series are useful tools for many proofs in econometrics
- Approximation methods are used frequently in economics to help simplify nonlinear equations
- Accuracy of the approximation depends on distance from x_0

[Show picture]

Second Derivations: Multiple Variables

- For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, we define second derivatives similarly, as the derivative of ∇f .
- Evidently, the second derivative of f is an $n \times n$ matrix, called the **Hessian** of f at x
- The form of the Hessian is

$$H(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial}{\partial x_n} \frac{\partial f}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} \frac{\partial f}{\partial x_n} & \cdots & \frac{\partial}{\partial x_n} \frac{\partial f}{\partial x_n} \end{pmatrix}$$

Worried about remembering the order of differentiation for the Hessian? In most cases, there's no need:

Theorem 3.1

(Schwarz) If f is twice continuously differentiable, the Hessian matrix is symmetric.

Taylor Series for Multivariable Functions

- We will consider second-order expansions, as the notation is cumbersome beyond that point (and expansions beyond two orders are rare in economics).
- Let's take $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Suppose we want to approximate f around (x_1^*, x_2^*) with a second-order polynomial

$$\begin{aligned} h(x_1, x_2) &= a_0 + a_1(x_1 - x_1^*) + a_2(x_2 - x_2^*) \\ &\quad + a_{11}(x_1 - x_1^*)^2 + a_{12}(x_1 - x_1^*)(x_2 - x_2^*) + a_{22}(x_2 - x_2^*)^2 \end{aligned}$$

- We again require that $h(x_0, y_0) = f(x_0, y_0)$ and that all first and second order derivatives of h match f

Taylor Series for Multivariable Functions (cont.)

- Differentiating and matching terms gives

$$\begin{aligned}h(x_1, x_2) &= f(x_1^*, x_2^*) + \frac{\partial f(x_1^*, x_2^*)}{\partial x_1}(x_1 - x_1^*) + \frac{\partial f(x_1^*, x_2^*)}{\partial x_2}(x_2 - x_2^*) \\&+ \frac{1}{2} \frac{\partial^2 f(x_1^*, x_2^*)}{\partial x_1^2}(x_1 - x_1^*)^2 + \frac{\partial^2 f(x_1^*, x_2^*)}{\partial x_1 \partial x_2}(x_1 - x_1^*)(x_2 - x_2^*) \\&+ \frac{1}{2} \frac{\partial^2 f(x_1^*, x_2^*)}{\partial x_2^2}(x_2 - x_2^*)^2\end{aligned}$$

- Much cleaner to write in matrix form:

$$h(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2}(x - x^*)^T H(x^*)(x - x^*)$$

- This formula holds for functions with more than 2 variables, and is known as the **second order Taylor series expansion of f around x^***

Table of Contents

Single Variable

Multivariable Derivative

Second Derivatives and Taylor Series Expansions

The Implicit Function Theorem

- Many economic analyses introduce equations of the form $f(x, y) = 0$, where x is a vector of “exogenous” variables and y a vector of “endogenous” variables
- We are frequently interested in understanding the impact of x on y , namely $y'(x)$
- However, the equations are complicated and it may not be possible to solve explicitly for $y(x)$ in order to take derivatives
- The implicit function theorem gives us a way to do such comparative statics even in the absence of a closed form solution

- Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$
- Assume for every x there is a unique y that satisfies $f(x, y) = 0$. Write $y = y(x)$
- Differentiate the expression $f(x, y(x)) = 0$ with respect to x and apply the chain rule:

$$f_x(x, y(x)) + f_y(x, y(x))y'(x) = 0$$

- So long as $f_y(x, y(x)) \neq 0$, we can solve for $y'(x)$:

$$y'(x) = -\frac{f_x(x, y(x))}{f_y(x, y(x))}$$

- Let $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$
- Assume for every x there is a unique y that satisfies $f(x, y) = 0$. Write $y = y(x)$
- Now differentiate the expression $f(x, y(x)) = 0$ with respect to x and apply the chain rule:

$$f_x(x, y(x)) + f_y(x, y(x))y'(x) = 0$$

- So long as $f_y(x, y(x))$ is invertible, we have

$$y'(x) = - \left(f_y(x, y(x))^{-1} \right) f_x(x, y(x))$$

- So long as f is continuously differentiable and $\det(f_y(x, y(x))) \neq 0$, the above formula is correct. See FMEA page 84 for a full statement.

IFT: Supply and Demand Example

- Let $\theta \in \mathbb{R}^n$ be a vector of variables that affect supply and demand.
- Market clearing implies

$$Q^s(\theta, p) = Q^d(\theta, p)$$

for all θ

- To put it into our format:

$$\underbrace{Q^s(\theta, p) - Q^d(\theta, p)}_{f(\theta, p): \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}} = 0$$

- This implicitly defines p as a function of θ . Differentiating gives:

$$Q_\theta^s(\theta, p(\theta)) + Q_p^s(\theta, p(\theta))p'(\theta) = Q_\theta^d(\theta, p(\theta)) + Q_p^d(\theta, p(\theta))p'(\theta)$$

- Solving for $p'(\theta)$ gives:

$$p'(\theta) = \frac{Q_{\theta}^d - Q_{\theta}^s}{Q_p^s - Q_p^d}$$

- The denominator is positive
- Therefore the sign of $p'(\theta)$ depends on the sign of $Q_{\theta}^d - Q_{\theta}^s$
- If demand reacts more strongly than supply to changes in θ , price increases