Mathematics for Economists

Mitri Kitti

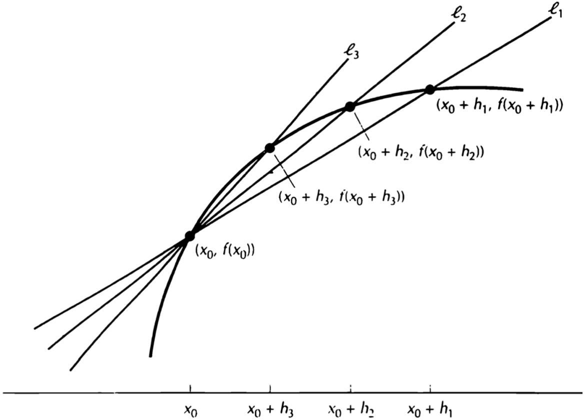
Aalto University

Calculus of Several Variables

Derivative as the Slope of the Tangent Line

- Geometrical interpretation of derivatives: The derivative of f at x_0 is the limit (as n goes to infinity) of the slope

$$\frac{f(x_0 + h_n) - f(x_0)}{h_n}.$$



Using Derivative in Approximating a Function

- Consider a function $f : \mathbb{R} \to \mathbb{R}$

- Suppose that at a point x_0 of the domain we change x from x_0 to $x_0 + \Delta x$ and we want to measure the corresponding change in f
- The exact change in f is $\Delta y = f(x_0 + \Delta x) - f(x_0)$
- But we can also *approximate* this change by using derivatives
- First define the differential of f as $df = f'(x)dx$
- Then notice that for a small change Δx, the derivative f' satisfies

$$f'(x_0) \approx \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x},$$

which can be rewritten as

$$\Delta x \, f'(x_0) \approx f(x_0 + \Delta x) - f(x_0) = \Delta y.$$

Using Derivative in Approximating a Function

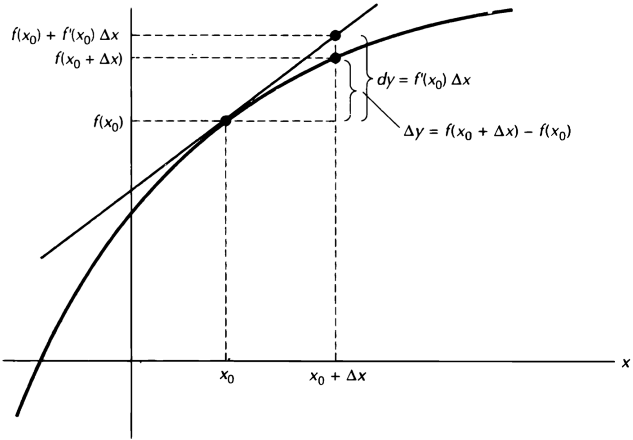
- In words, the actual change Δy is approximately equal to the derivative of f at x_0 multiplied by the change in x

- By the definition of differential, and by setting $dx = \Delta x$, we can conclude that, for a small change Δx, the actual change Δy is approximately equal to the differential of f evaluated at x_0, namely

$$\Delta y \approx df = f'(x_0)dx$$

Using Derivative in Approximating a Function

- Linear approximation via derivatives



Example: The Marginal Rate of Substitution (MRS)

- Let $U : \mathbb{R}^2_+ \longrightarrow \mathbb{R}$ be a differentiable utility function
- The *total differential* of U at a given point (x_0, y_0) in which $x_0, y_0 > 0$ is given by

$$dU = \frac{\partial U}{\partial x}(x_0, y_0)dx + \frac{\partial U}{\partial y}(x_0, y_0)dy \tag{1}$$

- In words, the total differential gives us an *approximate* measure of how much the value of U around (x_0, y_0) changes when we change both x from x_0 to $x_0 + dx$ and y from y_0 to $y_0 + dy$

Example: The Marginal Rate of Substitution (MRS)

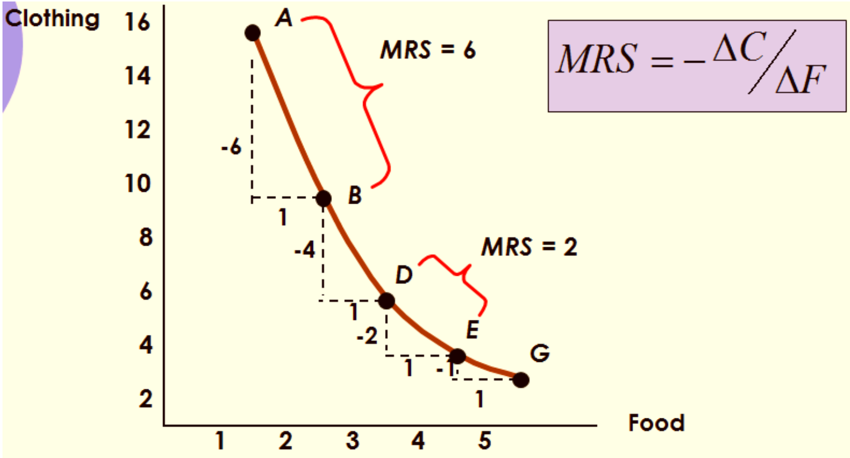
- If we require that $dU = 0$ (so that total utility is left unchanged), we can rewrite (1) as

$$\frac{dy}{dx} = -\frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}}(x_0, y_0),$$

  which is the Marginal Rate of Substitution at (x_0, y_0)

- Geometrically, the MRS denotes the slope at (x_0, y_0) of the indifference curve passing through that point

Example: The Marginal Rate of Substitution (MRS)



Total Differential

- In general, for a differentiable function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, the total differential at a point $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$ of the domain is

$$df = \frac{\partial f}{\partial x_1}(\mathbf{x}^*)dx_1 + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x}^*)dx_n$$

- Just like in the MRS example, the total differential provides an approximate measure of how much f changes in a neighborhood of x^* when the n variables $x_1, \ldots, x_n$ are changed by the amounts $dx_1, \ldots, dx_n$

Formal Definition of Derivative

Definition

A function $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is differentiable at $\mathbf{x} \in \mathbb{R}^n$ if there is a linear function $Df(\mathbf{x}) : \mathbb{R}^n \mapsto \mathbb{R}^m$ such that

$$\lim_{h \to \mathbf{0}} \frac{\|f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - [Df(\mathbf{x})]\mathbf{h}\|}{\|\mathbf{h}\|} = 0$$

- $Df(\mathbf{x})$ is the derivative of f at $\mathbf{x}$
- note 1: in the definition of the above limit, $Df(\mathbf{x})$ is independent on how $\mathbf{h}$ goes to zero
- note 2: a linear function is represented by a matrix! Hence $Df(\mathbf{x})$
- Differentiation/derivation is about making linear approximations

Jacobian Matrix

If f is differentiable at $\mathbf{x}$, then all partial derivatives exist at x and the (i,j)-component of $Df(\mathbf{x})$ is $\partial f_i(\mathbf{x})/\partial x_j$, we write

$$Df(x) = \begin{pmatrix} \partial f_1(\mathbf{x})/\partial x_1 & \partial f_1(\mathbf{x})/\partial x_2 & \cdots & \partial f_1(\mathbf{x})/\partial x_n \\ \partial f_2(\mathbf{x})/\partial x_1 & \partial f_2(\mathbf{x})/\partial x_2 & \cdots & \partial f_2(\mathbf{x})/\partial x_n \\ \vdots & \vdots & \ldots & \vdots \\ \partial f_m(\mathbf{x})/\partial x_1 & \partial f_m(\mathbf{x})/\partial x_2 & \cdots & \partial f_m(\mathbf{x})/\partial x_n \end{pmatrix}$$

and call this matrix as the Jacobian matrix

- The linear approximation of f at $\mathbf{x}^*$ is $f(\mathbf{x}) \approx f(\mathbf{x}^*) + Df(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)$

Gradient Vector

If $f : \mathbb{R}^n \mapsto \mathbb{R}$ is differentiable at $\mathbf{x}$, then its Jacobian matrix is

$$Df(\mathbf{x}) = \begin{pmatrix} \partial f(\mathbf{x})/\partial x_1 & \partial f(\mathbf{x})/\partial x_2 & \cdots & \partial f(\mathbf{x})/\partial x_n \end{pmatrix}$$

- The linear approximation of f at $\mathbf{x}^*$ is

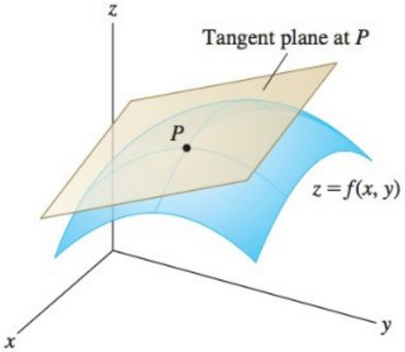
$$f(\mathbf{x}) \approx f(\mathbf{x}^*) + Df(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) = f(\mathbf{x}^*) + \sum_{i=1}^{n} \frac{\partial f(\mathbf{x}^*)}{\partial x_i}(x_i - x_i^*)$$

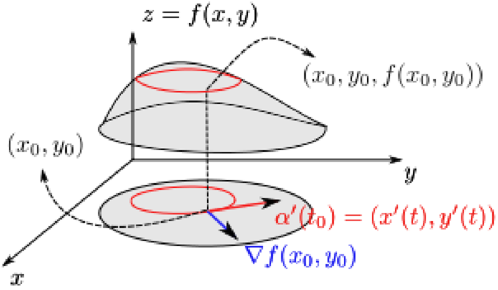
- It is convenient to write the sum as an inner product of a vector of partial derivatives and $\mathbf{x} - \mathbf{x}^*$
- The column vector of partial derivatives is called as the gradient of f:

$$\nabla f(\mathbf{x}^*) = \begin{pmatrix} \partial f(\mathbf{x})/\partial x_1 \\ \partial f(\mathbf{x})/\partial x_2 \\ \vdots \\ \partial f(\mathbf{x})/\partial x_n \end{pmatrix}$$

Gradient Vector

The normal of tangential plane of the graph of the function is $(\nabla f(\mathbf{x}), -1)$





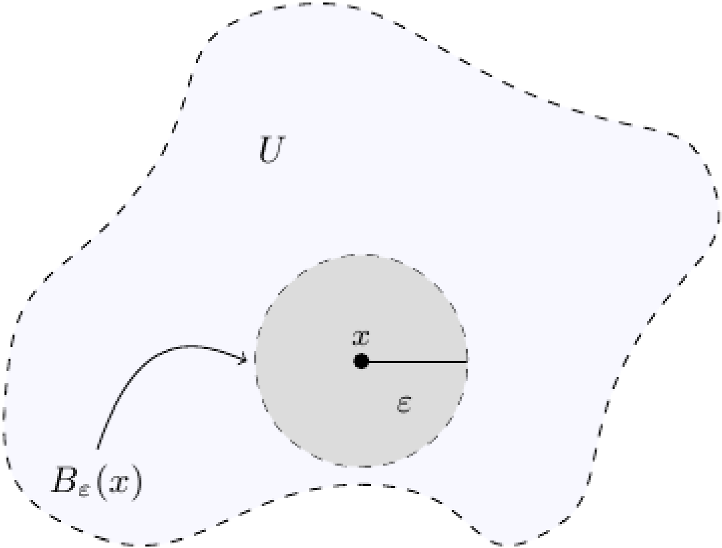
Failure of Differentiability

- Up to now, we have been implicitly assuming that derivatives always exist. However, this is not always the case. E.g., $f(x) = |x|$
- In order to define differentiable functions (i.e. functions whose derivatives exist), we need to introduce the concepts of open and closed sets

Open Sets

- Given a point/vector x in $\mathbb{R}^n$ and a real number $\epsilon > 0$, the **open ϵ-ball** around x is the set
$$B_\epsilon(x) = \{y \in \mathbb{R}^n : ||y - x|| < \epsilon\}$$
- In $\mathbb{R}$, open balls are intervals (a, b), with $a < b$
- A set $S \subseteq \mathbb{R}$ is **open** if for any $x \in S$ there exists an open ϵ-ball around x completely contained in S. That is,
$$x \in S \implies \text{ there exists an } \epsilon > 0 \text{ such that } B_\epsilon(x) \subseteq S.$$

Open Sets



Open Sets

- Examples of open sets:
  - open intervals $(a, b) \subset \mathbb{R}$, with $a < b$
  - open ϵ-balls
  - $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$
  - ...
  - $\mathbb{R}^n$
- A couple of properties:
  - Any union of open sets is an open set
  - The intersection of finitely many open sets is an open set

Closed Sets

- A set $S \subseteq \mathbb{R}^n$ is **closed** if, whenever $\{x_n\}_{n=1}^{\infty}$ is a convergent sequence completely contained in S, its limit is also contained in S.
- Equivalently, a set $S \subseteq \mathbb{R}^n$ is closed if and only if its complement $S^c = \mathbb{R}^n \backslash S$ is open

Closed Sets

- Examples of closed sets:
  - closed intervals $[a, b] \subset \mathbb{R}$, with $a \leq b$
  - $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$
  - ...
  - $\mathbb{R}^n$
- A couple of properties:
  - Any intersection of closed sets is a closed set
  - The union of finitely many closed sets is a closed set

Continuous Differentiability

- A function $f : \mathbb{R}^n \to \mathbb{R}$ is **continuously differentiable** (or C^1) on an open set $U \subseteq \mathbb{R}^n$ if and only if for each i, the derivative $\frac{\partial f}{\partial x_i}(y)$
  - exists for all $y \in U$ and
  - is continuous in y.
- Remark: Every differentiable function must be continuous but not every continuous function is differentiable

Chain Rule

- Let $h : \mathbb{R} \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}$ be two continuously differentiable functions. Let f be the composite function $g \circ h$, so that $f(x) = g(h(x)) = g(h_1(x), \ldots, h_n(x))$
- How to compute the derivative of f with respect to x?
- **Chain rule:** The derivative of f at x_0 is

$$\frac{df}{dx}(x_0) = \frac{\partial g}{\partial h_1}(h(x_0))\frac{dh_1}{dx}(x_0) + \cdots + \frac{\partial g}{\partial h_n}(h(x_0))\frac{dh_n}{dx}(x_0)$$

- In more compact form,

$$\frac{df}{dx} = \frac{\partial g}{\partial h_1}\frac{dh_1}{dx} + \cdots + \frac{\partial g}{\partial h_n}\frac{dh_n}{dx}$$

Chain Rule: Example

- **Example:** Consider the Cobb-Douglas production function $Q = 4K^{\frac{3}{4}}L^{\frac{1}{4}}$. Suppose that the inputs K and L vary with time t via the expressions

$$K(t) = 10t^2 \text{ and } L(t) = 6t^2.$$

- What is the rate of change (i.e. the derivative) of output Q with respect to t when $t = 10$?

- By the chain rule,

$$\begin{aligned} \frac{dQ}{dt}(t) &= \frac{\partial Q}{\partial K}(K(t), L(t))\frac{dK}{dt}(t) + \frac{\partial Q}{\partial L}(K(t), L(t))\frac{dL}{dt}(t) \\ &= 3K^{-\frac{1}{4}}(t)L^{\frac{1}{4}}(t)\,20t + K^{\frac{3}{4}}(t)L^{-\frac{3}{4}}(t)\,12t \end{aligned}$$

Chain Rule: Example

- When $t = 10$, we have

$$\begin{aligned}\frac{dQ}{dt}(t) &= 3K^{-\frac{1}{4}}(t)L^{\frac{1}{4}}(t)\,20t + K^{\frac{3}{4}}(t)L^{-\frac{3}{4}}(t)\,12t \\ &= 3(1000)^{-\frac{1}{4}}(600)^{\frac{1}{4}}\,200 + (1000)^{\frac{3}{4}}(600)^{-\frac{3}{4}}\,120 \\ &\approx 704.\end{aligned}$$

Higher-order derivatives

- **Higher-order derivatives:** The partial derivative $\frac{\partial f}{\partial x_i}$ of a function f of n variables is itself a function of n variables
- Provided that partial derivatives are differentiable, we can form higher-order derivatives by taking partial derivatives of partial derivatives
- For example,

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$$

is the partial derivative w.r.t. x_j of the partial derivative of f w.r.t. x_i. This is called the $x_i x_j$-second order partial derivative of f

Second Order Derivatives

- The $x_i x_i$-second order derivative is often written as

$$\frac{\partial^2 f}{\partial x_i^2}$$

- For $i \neq j$, we usually write the $x_i x_j$-derivative (also called cross partial derivative) as

$$\frac{\partial^2 f}{\partial x_j \partial x_i}$$

Hessian Matrix

- For a function f of n variables, we have $n \times n$ second order derivatives. These derivatives are often arranged in a square matrix called the **Hessian** matrix of f and written as $D^2 f(\mathbf{x})$:

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

- If all the derivatives in the Hessian exist and are continuous, we say that f is twice continuously differentiable or C^2

First and Second Order Approximations

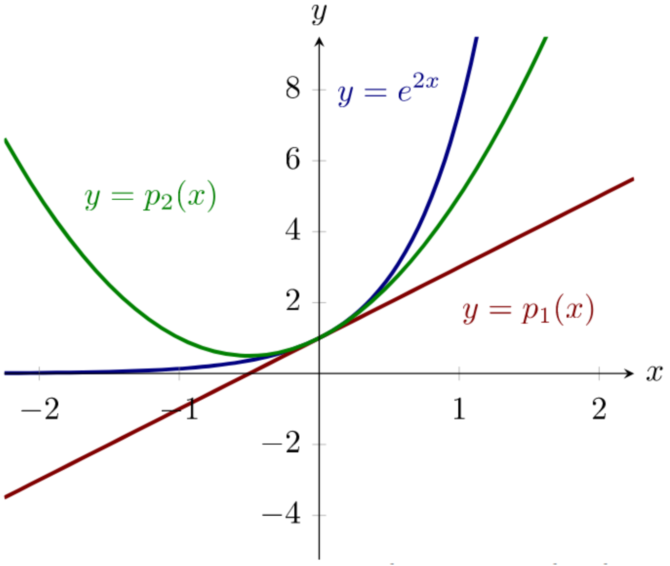
- Assume $f : \mathbb{R}^n \mapsto \mathbb{R}$
- The first order (or linear) approximation of f around $\mathbf{x}^*$ is

$$f(\mathbf{x}) \approx f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*)$$

- The second order (or quadratic) approximation is

$$f(\mathbf{x}) \approx f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) + (1/2)(\mathbf{x} - \mathbf{x}^*)^T D^2 f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)$$

First and Second Order Approximations: Example



Hessian Matrix

- An important property of second order partial derivatives is established by Young's theorem

- **Young's theorem:** If $f : \mathbb{R}^n \mapsto \mathbb{R}$ is twice continuously differentiable on an open set $A \subseteq \mathbb{R}^n$, then for all $\mathbf{x} \in A$ and for each pair of indices i, j, we have

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}.$$

- In other words, the Hessian matrix is symmetric and the order of differentiation does not matter

Hessian Matrix

- **Exercise:** Compute the Hessian matrix of the Cobb-Douglas production function $Q(K, L) = CK^a L^b$, with $C, a, b > 0$