Derivatives and optimization

AI503
Shan Shan

Lecture 8

Correction from last time

Definition (Critical Points)

Points where $\nabla f = \mathbf{0}$ or undefined.

Find and analyze any critical points of $f(x, y) = -\sqrt{x^2 + y^2}$.

The partial derivatives are undefined when $x = y = 0$ (i.e., at the origin). Hence $(0, 0)$ is a critical point. However, we cannot apply the second derivative test, because we cannot compute the second derivatives at $(0, 0)$.

Chain rule

Theorem

If f, g, and h are differentiable and if $z = f(x, y)$, and $x = g(t)$, and $y = h(t)$, then

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

Chain rule in a diagram

The following diagram keeps track of how a change in t propagates through the chain of composed functions.

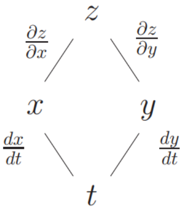

Figure: There are two paths from t to z, one through x and one through y. For each path, we multiply together the derivatives along the path. Then, to calculate dz/dt, we add the contributions from the two paths.

Questions

Suppose that $z = f(x, y) = x \sin y$, where $x = t^2$ and $y = 2t + 1$. Let $z = g(t)$. Compute $g'(t)$ directly and using the chain rule.

Questions

If f, g, h are differentiable and if $z = f(x, y)$, with $x = g(u, v)$ and $y = h(u, v)$. What is $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$?

Chain rule in multivariable

Theorem (Multivariable Chain Rule)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable, and let each x_i be a differentiable function of $t_1, \ldots, t_m$:

$$x_i = x_i(t_1, \ldots, t_m), \quad i = 1, \ldots, n.$$

Define

$$z = f(x_1, \ldots, x_n).$$

Then z is differentiable with respect to t_j, and

$$\frac{\partial z}{\partial t_j} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial t_j}, \quad j = 1, \ldots, m.$$

Differentiability of Multivariable Functions

Definition (Differentiable)

A function $f : \mathbb{R}^n \to \mathbb{R}$ is **differentiable** at the point $\mathbf{a} = (a_1, \ldots, a_n)$ if there exists a linear function

$$L(\mathbf{x}) = f(\mathbf{a}) + \sum_{i=1}^{n} m_i(x_i - a_i)$$

such that, if the error is defined by $f(\mathbf{x}) = L(\mathbf{x}) + E(\mathbf{x})$, and $\mathbf{h} = \mathbf{x} - \mathbf{a}$, then the **relative error** satisfies

$$\lim_{\mathbf{h} \to 0} \frac{E(\mathbf{a} + \mathbf{h})}{\|\mathbf{h}\|} = 0,$$

where $\|\mathbf{h}\| = \sqrt{h_1^2 + \cdots + h_n^2}$ is the Euclidean norm of $\mathbf{h}$.

- The function f is **differentiable on a region** $R \subset \mathbb{R}^n$ if it is differentiable at each point of R.
- The function L is called the local linearization of f near $\mathbf{a}$.
- Intuitively, differentiable means f is well-approximated by a linear map near $\mathbf{x}_0$.

If f is differentiable, then all partial derivatives exist there.

Show that if f is differentiable with the local linearization

$$L(\mathbf{x}) = f(\mathbf{a}) + \sum_{i=1}^{n} m_i(x_i - a_i),$$

then

$$m_i = \frac{\partial f}{\partial x_i}(\mathbf{a}).$$

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\mathbf{a} = (a_1, \dots, a_n)$. By definition,

$$\lim_{\mathbf{h} \to 0} \frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - \sum_{i=1}^{n} m_i h_i}{\|\mathbf{h}\|} = 0,$$

where $\mathbf{h} = (h_1, \dots, h_n)$.
Let $h_i > 0$. Taking $\mathbf{h} = (0, \dots, h_i, \dots, 0)$ (only the i-th component nonzero), we get

$$\lim_{h_i \to 0} \frac{f(a_1, \dots, a_i + h_i, \dots, a_n) - f(\mathbf{a}) - m_i h_i}{|h_i|} = 0.$$

Hence,

$$\lim_{h_i \to 0} \frac{f(a_1, \dots, a_i + h_i, \dots, a_n) - f(\mathbf{a}) - m_i h_i}{h_i} = \frac{\partial f}{\partial x_i}(\mathbf{a}) - m_i = 0.$$

Similar result holds for $h_i < 0$.

Therefore, if any of the partial derivatives fail to exist, then the function cannot be differentiable

Consider the function $f(x, y) = \sqrt{x^2 + y^2}$. Is f differentiable at the origin?

Existence of both partial derivatives at a point is not sufficient to guarantee differentiability

Consider the function $f(x, y) = x^{\frac{1}{3}} y^{\frac{1}{3}}$. Compute the partial derivatives at $(x, y) = (0, 0)$. Is f differentiable at $(0, 0)$?

In summary,

- If a function is differentiable at a point, then both partial derivatives exist there.
- Having both partial derivatives at a point does not guarantee that a function is differentiable there.

Continuity and differentiability

- If a function is differentiable at a point, then it is continuous there.
- Having both partial derivatives at a point does not guarantee that a function is continuous there.
- But, continuity of partial derivatives implies differentiability

Theorem

If the partial derivatives, f_x and f_y, of a function f exist and are continuous on a small disk centered at the point $\mathbf{a}$, then f is differentiable at $\mathbf{a}$.

Question

Show that the function $f(x, y) = ln(x^2 + y^2)$ is differentiable everywhere in its domain.

A word on differential

Sometimes (e.g., in differential geometry course) you will see something called a differential. This is simply an inner product of the gradient vector with another vector.

$$df_{\mathbf{x}_0}(\mathbf{h}) = \nabla f(\mathbf{x}_0) \cdot \mathbf{h}.$$

Equivalently, a differential is a linear combination of partial derivatives.

Definition of Definite Integral of Single Variable Function

Definition

Let $f(x)$ be a function on $[a, b]$. Then the **definite integral** of f from a to b is defined as

$$\int_a^b f(x)\,dx = \lim_{n\to\infty}\sum_{k=1}^{n} f(x_k^*)\Delta x$$

if the limit exists, where $\Delta x = \frac{b-a}{n}$, $x_k = a + k\Delta x$, and $x_k^* \in [x_{k-1}, x_k]$. If the limit exists, we say that f is **integrable** on $[a, b]$.

- The definite integral represents the area under the graph of f from a to b.
- If f is continuous on $[a, b]$, or has only finitely many jump discontinuities, then f is integrable on $[a, b]$.

Use the definition of the definite integral to compute

$$\int_a^b c\,dx \quad \text{for constants } a < b \text{ and } c.$$

Check your answer geometrically.

Use the definition of definite integral to simplify

$$\int_a^b cf(x)\,dx \quad \text{for constant } c.$$

Use the definition of definite integral to show that

$$\int_a^b [f(x) + g(x)]\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx.$$

Evaluate

$$\int_0^3 2x \, dx$$

as a limit of Riemann sums.

Fundamental Theorem of Calculus (Evaluation Theorem)

Theorem

If f is continuous on $[a, b]$, then

$$\int_a^b f(x)\,dx = F(b) - F(a),$$

where F is any antiderivative of f, i.e., $F'(x) = f(x)$.

Proof of FTC

- Divide $[a, b]$ into n subintervals with endpoints

$$a = x_0 < x_1 < \cdots < x_n = b.$$

The distance between subsequent x_k's is $\Delta x = \frac{b-a}{n}$.

- Let F be an antiderivative of f. We see that

$$\begin{aligned} F(b) - F(a) &= \sum_{k=1}^{n} [F(x_k) - F(x_{k-1})] \\ &= \sum_{k=1}^{n} F'(c_k)\Delta x \quad \text{(Mean Value Theorem)} \\ &= \sum_{k=1}^{n} f(c_k)\Delta x \end{aligned}$$

for some $c_k \in [x_{k-1}, x_k]$.

Proof of FTC

- Since F is differentiable, it is continuous. Taking the limit as $n \to \infty$ gives
$$F(b) - F(a) = \int_a^b f(x)\,dx.$$

Problem: Using FTC

Compute

$$\int_0^3 2x \, dx$$

using the fundamental theorem of calculus.

Second Fundamental Theorem of Calculus

Theorem

Let f be continuous on an interval. Then, for x and a in that interval,

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x)$$

Proof of Second FTC

Suppose $g(x) = \int_a^x f(t)dt$. Then

$$\frac{g(x+h) - g(x)}{h} = \frac{\int_x^{x+h} f(t)dt}{h}.$$

Since f is continuous on $[x, x+h]$, by the Extreme Value Theorem, f attains minimum m and maximum M on $[x, x+h]$, so

$$mh \leq \int_x^{x+h} f(t)dt \leq Mh.$$

Dividing by h gives

$$m \leq \frac{\int_x^{x+h} f(t)dt}{h} \leq M.$$

As $h \to 0$, $m \to f(x)$ and $M \to f(x)$, so by the Squeeze Theorem,

$$\frac{d}{dx}\left[\int_a^x f(t)dt\right] = f(x).$$

Problems: Second FTC

Find

$$\frac{d}{dx}\int_2^x \cos t \, dt.$$

Problems: Second FTC

Let $g(x) = \int_1^x \sqrt{1+t^2}\, dt$.

- Find $g'(x)$
- Find $g(x^3)$
- Compute $\frac{d}{dx} g(x^3)$.