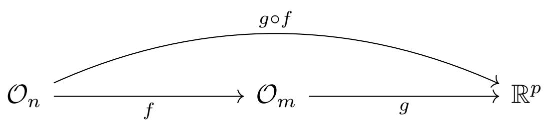
Lecture 11

11.1 Chain Rule

We will begin by recalling some results from the previous lecture.



If we are given two differentiable function $f : \mathcal{O}_n \to \mathcal{O}_m$ and $g : \mathcal{O}_m \to \mathbb{R}^p$, then $g \circ f$ is also differentiable. We also derived how to compute $D_{g\circ f}$ by **chain rule** as following,

$$D_{g\circ f}(a) = D_g(f(a)).D_f(a)$$

Now, comparing the $(i, j)^{\text{th}}$ element, we get,

$$\frac{\partial (g \circ f)_i(a)}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial g_i(b)}{\partial y_k} \cdot \frac{\partial f_k(a)}{\partial x_j}$$

where $b = f(a)$. This can be rewritten in a slightly more suggestive form by introducing new variables,

$$\begin{aligned} y_k &= f_k(x_1, \dots, x_n) \\ z_i &= g_i(y_1, \dots, y_m) \end{aligned}$$

Then, since $(g \circ f)_i = g_i \circ f$, the equation above can be written as,

$$\frac{\partial z_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k} \cdot \frac{\partial y_k}{\partial x_j}$$

This form of the chain rule is reminiscent of the one-variable chain rule.

Example 11.1.1

Let, $f(x, y, z) = xy^2z$ and $x = t, y = e^t, z = 1 + t$, we want to calculate $\frac{\mathrm{d}f}{\mathrm{d}t}$ in two ways.

First, we can write f as a function of t,

$$\begin{aligned} f(x, y, z) &= t(e^t)^2(1 + t) \\ &= (t + t^2)e^{2t} \end{aligned}$$

Hence, we have,

$$\begin{aligned} \frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\big(t + t^2\big)e^{2t} \\ &= (1 + 2t)e^{2t} + 2(t + t^2)e^{2t} \end{aligned}$$

$$= (2t^2 + 4t + 1)e^{2t}$$

Alternatively, if we apply the chain rule, we obtain,

$$\begin{aligned}
\frac{\mathrm{d}f}{\mathrm{d}t} &= \frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial f}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t} \\
&= y^2 z \cdot 1 + 2xyz \cdot e^t + xy^2 \cdot 1 \\
&= e^{2t}(1+t) + 2t(1+t)e^t.e^t + te^{2t} \\
&= e^{2t}(1 + t + 2t + 2t^2 + t) \\
&= (2t^2 + 4t + 1)e^{2t}
\end{aligned}$$

As we can see, both methods yield the same result!

11.2 Laplacian

The Laplacian operator plays a fundamental role in analyzing the behavior of functions and fields in multidimensional spaces. It quantifies the overall rate of change and spatial variations of a function, providing valuable insights into its properties and behavior.

Definition 11.2.1 ▶ Laplacian

$f : \mathcal{O}_n \to \mathbb{R}$ be a function. Then the Laplacian of f is defined as,

$$\Delta f = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}$$

Observe that,

$$\begin{aligned}
\Delta f &= \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} \\
&= \left\langle \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right\rangle \cdot \left\langle \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right\rangle \\
&= \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} f
\end{aligned}$$

Hence, Laplacian can be written as, $\Delta f = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} f = \nabla^2 f$.

Laplacian in Polar Coordinate

Let, f be a twice differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$. We can express $f(x, y)$ in polar coordinates as a function of (r, θ) by substituting, $x = r\cos\theta$ and $y = r\sin\theta$. Now, observe the following partial derivatives,

$$\begin{aligned}
\frac{\partial x}{\partial r} &= \cos\theta, \quad \frac{\partial x}{\partial \theta} = -r\sin\theta \\
\frac{\partial y}{\partial r} &= \sin\theta, \quad \frac{\partial y}{\partial \theta} = r\cos\theta
\end{aligned}$$

We want to express f_{xx} and f_{yy} in terms of partial derivatives of f in polar coordinates. Notice that,

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r}$$

$$\text{i.e., } \frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta$$

Differentiating once more with respect to r, we have,

$$\begin{aligned}
\frac{\partial^2 f}{\partial r^2} &= \frac{\partial}{\partial r}\left[\frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta\right] \\
&= \cos\theta\left[\frac{\partial}{\partial r}\frac{\partial f}{\partial x}\right] + \sin\theta\left[\frac{\partial}{\partial r}\frac{\partial f}{\partial y}\right] \\
&= \cos\theta\left[\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y\partial x}\frac{\partial y}{\partial r}\right] + \sin\theta\left[\frac{\partial^2 f}{\partial x\partial y}\frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial r}\right] \\
&= \cos\theta\left[\frac{\partial^2 f}{\partial x^2}\cos\theta + \frac{\partial^2 f}{\partial y\partial x}\sin\theta\right] + \sin\theta\left[\frac{\partial^2 f}{\partial x\partial y}\cos\theta + \frac{\partial^2 f}{\partial y^2}\sin\theta\right] \\
&= \cos\theta\left[\cos\theta f_{xx} + \sin\theta f_{xy}\right] + \sin\theta\left[\cos\theta f_{xy} + \sin\theta f_y\right]
\end{aligned}$$

Hence, we get,

$$\boxed{\frac{\partial^2 f}{\partial r^2} = \cos^2\theta f_{xx} + \sin^2\theta f_y + \sin 2\theta f_{xy}}$$

Similarly, we can find the expression for $\frac{\partial^2 f}{\partial\theta^2}$,

$$\boxed{\frac{\partial^2 f}{\partial\theta^2} = -r\left(\cos\theta f_x + \sin\theta f_y\right) + \left(r^2\sin^2\theta f_{xx} + r^2\cos^2\theta f_y - r^2\sin 2\theta f_{xy}\right)}$$

Combining the above two result we can write,

$$\boxed{\Delta f = f_{xx} + f_{yy} = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\cdot\frac{\partial f}{\partial r} + \frac{1}{r^2}\cdot\frac{\partial^2 f}{\partial\theta^2}}$$

Example 11.2.1 (Writing Laplacian in New coordinate)

Let, $z = z(u, v)$ where,

$$u(x, y) = x^2 y \text{ and } v(x, y) = 3x + 2y$$

We want to express the Laplacian with respect to u and v. Starting with the given coordinates,

$$\frac{\partial u}{\partial x} = 2xy, \quad \frac{\partial u}{\partial y} = x^2$$
$$\frac{\partial v}{\partial x} = 3, \quad \frac{\partial v}{\partial y} = 2$$

We can find $\frac{\partial z}{\partial x}$ using the chain rule,

$$\begin{aligned}
\frac{\partial z}{\partial x} &= \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x} \\
\implies \frac{\partial z}{\partial x} &= 2xy\frac{\partial z}{\partial u} + 3\frac{\partial z}{\partial v}
\end{aligned}$$

Differentiating once more with respect to x, we have,

$$\begin{aligned}
\frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x}\left[2xy\frac{\partial z}{\partial u} + 3\frac{\partial z}{\partial v}\right] \\
&= 2y\frac{\partial z}{\partial u} + 2xy\frac{\partial}{\partial x}\left[\frac{\partial z}{\partial u}\right] + 3\frac{\partial}{\partial x}\left[\frac{\partial z}{\partial v}\right]
\end{aligned}$$

$$
\begin{aligned}
&= 2y\frac{\partial z}{\partial u} + 2xy\left(\frac{\partial}{\partial u}\left[\frac{\partial z}{\partial u}\right]\frac{\partial u}{\partial x} + \frac{\partial}{\partial v}\left[\frac{\partial z}{\partial u}\right]\frac{\partial v}{\partial x}\right) + 3\left(\frac{\partial}{\partial u}\left[\frac{\partial z}{\partial v}\right]\frac{\partial u}{\partial x} + \frac{\partial}{\partial v}\left[\frac{\partial z}{\partial v}\right]\frac{\partial v}{\partial x}\right) \\
&= 2y\frac{\partial z}{\partial u} + 2xy\left(\frac{\partial^2 z}{\partial u^2}\frac{\partial u}{\partial x} + \frac{\partial^2 z}{\partial v \partial u}\frac{\partial v}{\partial x}\right) + 3\left(\frac{\partial^2 z}{\partial u \partial v}\frac{\partial u}{\partial x} + \frac{\partial^2 z}{\partial v^2}\frac{\partial v}{\partial x}\right) \\
&= 2y\frac{\partial z}{\partial u} + 2xy\left(2xy\frac{\partial^2 z}{\partial u^2} + 3\frac{\partial^2 z}{\partial v \partial u}\right) + 3\left(\frac{1}{2xy}\cdot\frac{\partial^2 z}{\partial u \partial v} + 3\frac{\partial^2 z}{\partial v^2}\right)
\end{aligned}
$$

Hence, we get,

$$
\boxed{\frac{\partial^2 z}{\partial x^2} = 2yz_u + 4x^2y^2 z_{uu} + 6xyz_{uv} + 6xyz_{vu} + 9z_{vv}}
$$

Exercise. Find z_{yy}, z_{yx}, z_{xy} and check if $z_{xy} = z_{yx}$.

11.3 Extrema of a function

Finding the extrema of a function is crucial in calculus, allowing us to identify maximum and minimum points and to relate the structure of functions. We will now extend this concept to functions of several variables.

Definition 11.3.1 ▶ Extrema

Let, a is an interior point of $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ be a function.

- f attains a **local maximum** at a if there exists an open neighborhood $\mathcal{O}_n$ of a such that, $f(a) \geq f(x)\ \forall\ x \in \mathcal{O}_n$.
- Similarly, f attains a **local minimum** at a if there exists an open neighborhood $\mathcal{O}_n$ of a such that, $f(a) \leq f(x)\ \forall\ x \in \mathcal{O}_n$.

Any point at which f attains a local(global) maxima (or minima) is called extremum point of that function. In plural, it is called **Extrema**.

Definition 11.3.2 ▶ Critical Point or Stationary Point

Let, $f : S(\subseteq \mathbb{R}^n) \to \mathbb{R}$ be a function and $a \in \mathcal{O}_n \subseteq S$. We say that a is a **critical point** or **stationary point**. If

$$
(\boldsymbol{\nabla} f)(a) = 0
$$

Or, equivalently all the partial derivatives $\frac{\partial f}{\partial x_i}$ are zero.

Theorem 11.3.1

Let, $f : \mathcal{O}_n \to \mathbb{R}$ is differentiable at $a \in \mathcal{O}_n$. If a is a local extremum, then

$$
(\boldsymbol{\nabla} f)(a) = 0
$$

Proof. Fix $i \in \{1, 2, \ldots, n\}$. We want to show $\frac{\partial f}{\partial x_i} = 0$. For this set, $\phi_i : (a_i - \epsilon, a_i + \epsilon) \to \mathbb{R}$ defined by

$$
\phi_i(t) = f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n)
$$

Notice that, $\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = f_{x_i}(a)$. Since a is local extremum of f, we can say that a_i is a local extremum of ϕ_i. So, $\frac{\mathrm{d}\phi_i}{\mathrm{d}t} = 0$, which means, $\frac{\partial f(a)}{\partial x_i} = 0$. We can do this for all i and hence, $(\boldsymbol{\nabla} f)(a) = 0$. ☐

Question. When we did calculation for local extremum for the functions f with one variable, we used to evaluate the stationary points by calculating, $f'(x) = 0$. Then we used to check the second derivative in order to know whether the stationary point is local minima or maxima or saddle point. For multivariate case also, we need 2nd order derivative to know the behavior of the stationary point. Now what could be 2nd order total derivative?

Answer. For this purpose we will introduce **Hessian Matrix** in next class.