LECTURE 7

Differentials and the Chain Rule

In this lecture we will elaborate on notion of gradient that we introduced when we discussed the differentiability of maps from $\mathbb{R}^n$ to $\mathbb{R}^m$.

DEFINITION 7.1. *The **differential** $\mathbf{D}f$ of a map $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ at the point of $\mathbf{x}$ is the following matrix of partial derivative*

$$\mathbf{D}f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

In the special case where f is a function from $\mathbb{R}^n$ to $\mathbb{R}$ (i.e., $m = 1$) the differential $\mathbf{D}f(\mathbf{x})$ coincides with the gradient $\nabla f(\mathbf{x})$ of f at $\mathbf{x}$.

EXAMPLE 7.2. Let $f : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \mapsto xy^2z$. Then

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}(x, y, z), \frac{\partial f}{\partial y}(x, y, z), \frac{\partial f}{\partial z}(x, y, z) \right) = \left(y^2 z, 2xyz, xy^2\right)$$

EXAMPLE 7.3. Let $f : \mathbb{R}^2 \to \mathbb{R}^3 : (x, y) \mapsto (xy, y^2, x^2 + y^2)$. Then

$$\mathbf{D}f(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} \end{pmatrix} = \begin{pmatrix} y & x \\ 0 & 2y \\ 2x & 2y \end{pmatrix}$$

THEOREM 7.4. *Let f and g be functions from $U \subset \mathbb{R}^n$ to $\mathbb{R}^m$ that are differentiable at $\mathbf{x}_0 \in U$. Then*

1. $\mathbf{D}\,(cf)\,(\mathbf{x}_0) = c\mathbf{D}f(\mathbf{x}_0)$ *if c is a constant.*
2. $\mathbf{D}(f + g)(\mathbf{x}_0) = \mathbf{D}f(\mathbf{x}_0) + \mathbf{D}g(\mathbf{x}_0)$. *(The addition on the left hand side is addition of functions, the addition on the right hand side is addition of matrices.)*

THEOREM 7.5. *If Let f and g be functions from $U \subset \mathbb{R}^n$ to $\mathbb{R}$ that are differentiable at $\mathbf{x}_0 \in U$. Then*

1. $\nabla(fg)(\mathbf{x}_0) = g(\mathbf{x}_0)\nabla f(\mathbf{x}_0) + f(\mathbf{x}_0)\nabla g(\mathbf{x}_0)$. *(Product Rule.)*
2. $\nabla(f/g)(\mathbf{x}_0) = g(\mathbf{x}_0)\nabla f(\mathbf{x}_0) - f(\mathbf{x}_0)\nabla g(\mathbf{x}_0)/\left[g(\mathbf{x}_0)\right]^2$ *(Quotient Rule.)*

1. The Chain Rule

Lets us recall the chain rule for functions $f$and g are each functions of single variable then

$$\frac{d}{dx}(g \circ f)(x) = \frac{dg}{df}\frac{df}{dx}$$

or more precisely, regarding f as a function sending x to $f(x)$, and $g(u)$ as a function sending u to $g(u)$:

$$\frac{d}{dx}(g \circ f)(x) = \left.\frac{dg}{du}\right|_{u=f(x)} \frac{df}{dx}(x)$$

The analog for this chain rule for functions of complex variables has a similar form when expressed in terms of the differentials defined above.

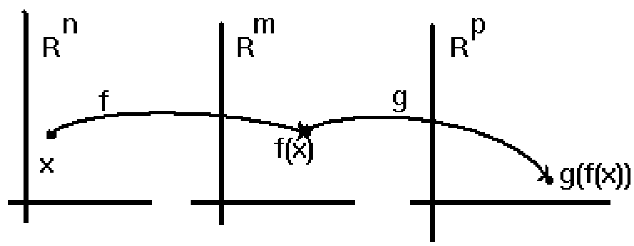
THEOREM 7.6. *Let $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ and $g : f(U) \subset \mathbb{R}^m \to \mathbb{R}^p$ be differentiable functions. Then the composed function*

$$g \circ f : \mathbb{R}^n \to \mathbb{R}^m \to \mathbb{R}^p$$

is differentiable and

$$[\mathbf{D}\,(g \circ f)]\,(\mathbf{x}) = [\mathbf{D}g\,(f(\mathbf{x}))]\,[\mathbf{D}f(\mathbf{x})]$$

where the product of the two differentials on the right hand side is the product of the $p \times m$ matrix $\mathbf{D}g\,(f(\mathbf{x}))$ by the $m \times n$ matrix $\mathbf{D}f(\mathbf{x})$.



EXAMPLE 7.7. Suppose $\gamma : \mathbb{R} \to \mathbb{R}^3 : t \mapsto (\gamma_x(t), \gamma_y(t), \gamma_z(t))$ and $f : \mathbb{R}^3 \to \mathbb{R} : (x, y, z) \mapsto f(x, y, z)$. Then $f \circ \gamma : \mathbb{R} \to \mathbb{R}$ is a function of a single variable and

$$\begin{aligned}
\frac{d(f \circ \gamma)}{dt} &= \mathbf{D}f(\gamma(t)\mathbf{D}\gamma(t) \\
&= \begin{pmatrix} \frac{\partial f}{\partial x}\left(\gamma(t)\right) & \frac{\partial f}{\partial y}\left(\gamma(t)\right) & \frac{\partial f}{\partial z}\left(\gamma(t)\right) \end{pmatrix} \begin{pmatrix} \frac{d\gamma_x}{dt}(t) \\ \frac{d\gamma_y}{dt}(t) \\ \frac{d\gamma_z}{dt}(t) \end{pmatrix} \\
&= \frac{\partial f}{\partial x}\left(\gamma(t)\right)\frac{d\gamma_x}{dt}(t) + \frac{\partial f}{\partial y}\left(\gamma(t)\right)\frac{d\gamma_y}{dt}(t) + \frac{\partial f}{\partial z}\left(\gamma(t)\right)\frac{d\gamma_z}{dt}(t) \\
&\approx \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}
\end{aligned}$$

EXAMPLE 7.8. Suppose $f : \mathbb{R}^3 \to \mathbb{R}^3 : (x, y, z) \mapsto (u, v, w)$ and $g : \mathbb{R}^3 \to \mathbb{R} : (u, v, w) \mapsto g(u, v, w)$. Then $g \circ f : \mathbb{R}^3 \to \mathbb{R}$ is a function of a three variables and

$$\begin{aligned}
\nabla(g \circ f) &= \mathbf{D}(g \circ f)(x, y, z) \\
&= \mathbf{D}g(f((x, y, z))\mathbf{D}f(x, y, z) \\
&= \begin{pmatrix} \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} & \frac{\partial g}{\partial w} \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\partial g}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial g}{\partial w}\frac{\partial w}{\partial x}, & \frac{\partial g}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial g}{\partial w}\frac{\partial w}{\partial y}, & \frac{\partial g}{\partial u}\frac{\partial u}{\partial z} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial z} + \frac{\partial g}{\partial w}\frac{\partial w}{\partial z} \end{pmatrix}
\end{aligned}$$