MA 109 Week 6 Supplementary slides

Mayukh Mukherjee

Department of Mathematics

January 7, 2022

1. The Chain Rule

2. General (and better) notation for the chain rule

3. Problems involving the gradient

Theorem (Chain Rules)

Let $D \subset \mathbb{R}^2$ and let (x_0, y_0) be an interior point of D, and let $f : D \to \mathbb{R}$ be differentiable at (x_0, y_0).

(i) Let $E \subset \mathbb{R}$ be such that $f(D) \subset E$ and $z_0 := f(x_0, y_0)$ is an interior point of E. If $g : E \to \mathbb{R}$ is differentiable at $f(x_0, y_0)$, then the function $h := g \circ f : D \to \mathbb{R}$ is differentiable at (x_0, y_0),

$$h_x(x_0, y_0) = g'(z_0)\, f_x(x_0, y_0) \ \text{ and } \ h_y(x_0, y_0) = g'(z_0)\, f_y(x_0, y_0).$$

(ii) Let $E \subset \mathbb{R}$, and let t_0 be an interior point of E. If $x, y : E \to \mathbb{R}$ are differentiable at t_0, and if $(x(t), y(t)) \in D$ for all $t \in E$ and $(x(t_0), y(t_0)) := (x_0, y_0)$, then the function $\phi : E \to \mathbb{R}$ defined by $\phi(t) := f(x(t), y(t))$ for $t \in E$ is differentiable at t_0, and

$$\phi'(t_0) = f_x(x_0, y_0)\, x'(t_0) + f_y(x_0, y_0)\, y'(t_0).$$

Theorem (Chain Rules: continued)

(iii) Let $E \subset \mathbb{R}^2$, and let (u_0, v_0) be an interior point of E. If $x, y : E \to \mathbb{R}$ are differentiable at (u_0, v_0), and if $(x(u,v), y(u,v)) \in D$ for all $(u,v) \in E$ and $(x(u_0, v_0), y(u_0, v_0)) := (x_0, y_0)$, then the function $F : E \to \mathbb{R}$ defined by $F(u,v) := f(x(u,v), y(u,v))$ for $(u,v) \in E$ is differentiable at (u_0, v_0),

$$F_u(u_0, v_0) = f_x(x_0, y_0)\, x_u(u_0, v_0) + f_y(x_0, y_0)\, y_u(u_0, v_0),$$

and

$$F_v(u_0, v_0) = f_x(x_0, y_0)\, x_v(u_0, v_0) + f_y(x_0, y_0)\, y_v(u_0, v_0).$$

It is often helpful to write the identities given by the Chain Rules in an informal but suggestive notation as follows.

(i) If $z = f(x,y)$ and $w = g(z)$, then w is a function of (x,y), and

$$\frac{\partial w}{\partial x} = \frac{dw}{dz}\frac{\partial z}{\partial x}, \quad \frac{\partial w}{\partial y} = \frac{dw}{dz}\frac{\partial z}{\partial y}.$$

(ii) If $z = f(x, y)$ and if $x = x(t)$, $y = y(t)$, then z is a function of t, and

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

(iii) If $z = f(x, y)$ and if $x = x(u, v)$, $y = y(u, v)$, then z is a function of u and v, and

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} \quad \text{and} \quad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}.$$

It should be noted that the identities in (i), (ii), and (iii) above are valid when the concerned (partial) derivatives are evaluated at appropriate points and when, for instance, each of the (partial) derivatives exists in a neighbourhood of the relevant point and is continuous at that point.

The conclusions of Chain Rules (i), (ii) and (iii) can be written in terms of the gradient $(\nabla f)(x_0, y_0)$ of f at (x_0, y_0) as follows:

$$
\begin{array}{lrcl}
\text{(i)} & (\nabla h)(x_0, y_0) & = & g'(z_0)(\nabla f)(x_0, y_0) \\[2ex]
\text{(ii)} & \phi'(t_0) & = & (\nabla f)(x_0, y_0) \cdot \big(x'(t_0), y'(t_0)\big) \\[2ex]
\text{(iii)} & (\nabla F)(u_0, v_0) & = & \Big((\nabla f)(x_0, y_0) \cdot \big(x_u(u_0, v_0), y_u(u_0, v_0)\big), \\
 & & & \quad (\nabla f)(x_0, y_0) \cdot \big(x_v(u_0, v_0), y_v(u_0, v_0)\big)\Big)
\end{array}
$$

The derivative for $f : U \to \mathbb{R}^n$

We now define the derivative for a function $f : U \to \mathbb{R}^n$, where U is a subset of $\mathbb{R}^m$.

The function f is said to be differentiable at a point x if there exists an $n \times m$ matrix $Df(x)$ such that

$$\lim_{\|h\| \to 0} \frac{\|f(x+h) - f(x) - Df(x) \cdot h\|}{\|h\|} = 0.$$

Here $x = (x_1, x_2, \ldots, x_m)$ and $h = (h_1, h_2, \ldots h_m)$ are vectors in $\mathbb{R}^m$.

The matrix $Df(x)$ is usually called the total derivative of f. It is also referred to as the Jacobian matrix. What are its entries?

From our experience in the 2×1 case we might guess (correctly!) that the entries will be the partial derivatives.

Here is the total derivative or the derivative matrix written out fully.

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_m}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_m}(x) \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \frac{\partial f_n}{\partial x_1}(x) & \frac{\partial f_n}{\partial x_2}(x) & \cdots & \frac{\partial f_n}{\partial x_m}(x) \end{pmatrix}$$

In the 2×2 case we get

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) \end{pmatrix}.$$

As before, the derivative may be viewed as a linear map, this time from $\mathbb{R}^m$ to $\mathbb{R}^n$ (or, in the case just above, from $\mathbb{R}^2$ to $\mathbb{R}^2$).

Rules for the total derivative

Just like in the one variable case, it is easy to prove that

$$D(f+g)(x) = Df(x) + Dg(x).$$

Somewhat harder, but only because the notation gets more cumbersome, is the Chain rule:

$$D(f \circ g)(x) = Df(g(x)) \circ Dg(x),$$

where $\circ$ on the right hand side denotes matrix multiplication. The theorem about the sufficient condition for differentiability holds in this greater generality - a function from $\mathbb{R}^m$ to $\mathbb{R}^n$ is differentiable at a point x_0 if all the partial derivatives $\frac{\partial f_i}{\partial x_j}$ $1 \leq i \leq n$, $1 \leq j \leq m$, are continuous in a neighborhood of x_0 (define a neighborhood of x_0 in $\mathbb{R}^m$!).

Review - problems involving the gradient

Exercise 1: Find the points on the hyperboloid $x^2 - y^2 + 2z^2 = 1$ where the normal line is parallel to the line that joins the points $(3, -1, 0)$ and $(5, 3, 6)$.

Solution: The hyperboloid is an implicitly defined surface. A normal vector at a point (x_0, y_0, z_0) on the hyperboloid is given by the gradient of the function $x^2 - y^2 + 2z^2$ at (x_0, y_0, z_0):

$$\nabla f(x_0, y_0, z_0) = (2x_0, -2y_0, 4z_0).$$

We require this vector to be parallel to the line joining the points $(3, -1, 0)$ and $(5, 3, 6)$. This line lies in the same direction as the vector $(5-3, 3+1, 6-0) = (2, 4, 6)$. Thus we need only solve the equations

$$2x_0 = 2, \quad -2y_0 = 4, \quad 4z_0 = 6,$$

which give $x_0 = 1$, $y_0 = -2$ and $z_0 = 3/2$. Thus, we need to find λ such that $\lambda(1, -2, 3/2)$ lies on the hyperboloid. Substituting in the equation yields $\lambda = \pm\sqrt{2/3}$.

Problems involving the gradient, continued

Exercise 2: Find the directions in which the directional derivative of $f(x, y) = x^2 + \sin xy$ at the point $(1, 0)$ has the value 1.

Solution: We compute ∇f first:

$$\nabla f(x, y) = (2x + y \cos xy, x \cos xy),$$

so at $(1, 0)$ we get, $\nabla f(1, 0) = (2, 1)$.

To find the directional derivative in the direction $v = (v_1, v_2)$ (where v is a unit vector), we simply take the dot product with the gradient:

$$\nabla_v f(1, 0) = 2v_1 + v_2.$$

This will have value "1" when $2v_1 + v_2 = 1$, subject to $v_1^2 + v_2^2 = 1$, which yields $v_1 = 0, v_2 = 1$ or $v_1 = 4/5, v_2 = -3/5$.

Problems involving the gradient, continued

Exercise 3: Find $\nabla_u F(2,2,1)$ where $\nabla_u F$ denotes the directional derivative of the function $F(x,y,z) = 3x - 5y + 2z$ and u is the unit vector in the outward normal to the sphere $x^2 + y^2 + z^2 = 9$ at the point $(2,2,1)$.

Solution: The unit outward normal to the sphere $g(x,y,z) = 9$ at $(2,2,1)$ is given by

$$\frac{\nabla g(2,2,1)}{\|\nabla g(2,2,1)\|}.$$

We see that $\nabla g(2,1,1) = (4,4,2)$ so the corresponding unit vector is $(2,2,1)/3$.

To get the directional derivative we simply take the dot product of ∇F with u:

$$(3,-5,2)\cdot(2,2,1)/3 = -2/3$$

Comments: Also, there is no need to compute the gradient to find the normal vector to the sphere - it is obviously the radial vector at the point $(2,2,1)$!.

Problems involving the gradient, continued

Exercise 4: Find the equations of the tangent plane and the normal line to the surface

$$F(x, y, z) := x^2 + 2xy - y^2 + z^2 = 7$$

at $(1, -1, 3)$.

Solution: We first compute the gradient of F to get $\nabla F(x, y, z) = (2x + 2y, 2x - 2y, 2z)$. At $(1, -1, 3)$, this yields the vector $\lambda(0, 4, 6)$ which is normal to the given surface at $(1, -1, 3)$. The point $(1, 3, 9)$ also lies on the normal line so its equation are

$$x = 1, \frac{y + 1}{4} = \frac{z - 3}{6}.$$

The equation of the tangent plane is given by

$$4(y + 1) + 6(z - 3) = 0,$$

since it consists of all lines orthogonal to the normal and passing through the point $(1, -1, 3)$.