MATH 2010E ADVANCED CALCULUS I
LECTURE 10

WING HONG TONY WONG

14.5 — Directional derivatives and gradient vectors

Definition 1. Let $D \subseteq \mathbb{R}^n$ be an open domain. Let $f : D \to \mathbb{R}$ be a real-valued function. Let $\mathbf{a} \in D$, and let $\mathbf{u}$ be a unit vector in $\mathbb{R}^n$. The **directional derivative** of f at $\mathbf{x} = \mathbf{a}$ in the direction $\mathbf{u}$ is

$$D_{\mathbf{u}}f(\mathbf{a}) = \lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a})}{t}.$$

Geometrically, when we consider the directional derivative $D_{\mathbf{u}}f(\mathbf{a})$, we first consider the curve obtained by intersecting the surface

$$x_{n+1} = f(x_1, \dots, x_n)$$

with the 2-dimensional plane

$$\{(x_1, \dots, x_n) : x_1 = a_1 + tu_1, \ \dots, \ x_n = a_n + tu_n\}.$$

The partial derivative is the slope of the curve at the point $\mathbf{x} = \mathbf{a}$ on the 2-dimensional plane.

From this, we can easily see that the directional derivative is a generalization of partial derivatives. In fact, $f_{x_i} = D_{\mathbf{e}_i}f$, where $\mathbf{e}_i$ is the vector of all zeros except a 1 in the i-th entry.

Another way to view directional derivative is

$$D_{\mathbf{u}}f(\mathbf{a}) = \lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a})}{t} = \frac{d}{dt} f(\mathbf{a} + t\mathbf{u})\bigg|_{t=0}.$$

If f is differentiable at $\mathbf{a}$, then we can view this as the case $\mathbb{R} \to \mathbb{R}^n \to \mathbb{R}$ in chain rule, so

$$\begin{aligned}
\frac{d}{dt} f(\mathbf{a} + t\mathbf{u})\bigg|_{t=0} &= \left(\frac{\partial f}{\partial x_1}\bigg|_{\mathbf{x}=\mathbf{a}} \right) \left(\frac{dx_1}{dt}\bigg|_{t=0} \right) + \cdots + \left(\frac{\partial f}{\partial x_1}\bigg|_{\mathbf{x}=\mathbf{a}} \right) \left(\frac{dx_1}{dt}\bigg|_{t=0} \right) \\
&= f_{x_1}(\mathbf{a})u_1 + \cdots + f_{x_n}(\mathbf{a})u_n \\
&= \big(f_{x_1}(\mathbf{a}), \dots, f_{x_n}(\mathbf{a})\big) \cdot (u_1, \dots, u_n) \\
&= \nabla f(\mathbf{a}) \cdot \mathbf{u}.
\end{aligned}$$

In other words, all directional derivatives are linear combinations of n partial derivatives.

Example 2. Find the directional derivative of $f(x, y) = xe^y + \cos(xy)$ at the point $(a, b) = (2, 0)$ in the direction $\mathbf{v} = (3, -4)$.

Date: Thursday, 11th June, 2015.

Solution. The unit vector in the direction of $\mathbf{v}$ is

$$\mathbf{u} = \frac{(3,-4)}{\sqrt{3^2+(-4)^2}} = \left(\frac{3}{5}, -\frac{4}{5}\right).$$

Therefore, the desired direction derivative is

$$\begin{aligned} D_{\mathbf{u}}f(2,0) &= \nabla f(2,0) \cdot \left(\frac{3}{5}, -\frac{4}{5}\right) \\ &= \left(e^y - y\sin(xy), xe^y - x\sin(xy)\right)\Big|_{(x,y)=(2,0)} \cdot \left(\frac{3}{5}, -\frac{4}{5}\right) \\ &= (1,2)\cdot\left(\frac{3}{5}, -\frac{4}{5}\right) = -1. \end{aligned}$$

□

In the definition of differentiability of f at $\mathbf{x} = \mathbf{a}$, we have

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - \nabla f(\mathbf{a})\cdot\mathbf{h}}{\|\mathbf{h}\|} = 0$$

$$\lim_{\mathbf{h}\to\mathbf{0}} \left(\frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})}{\|\mathbf{h}\|} - \nabla f(\mathbf{a})\cdot\frac{\mathbf{h}}{\|\mathbf{h}\|}\right) = 0$$

$$\lim_{\substack{\|\mathbf{u}\|=1 \\ t\to 0^+}} \left(\frac{f(\mathbf{a}+t\mathbf{u}) - f(\mathbf{a})}{t} - \nabla f(\mathbf{a})\cdot\mathbf{u}\right) = 0 \quad \text{by letting } t = \|\mathbf{h}\|.$$

$$\Leftrightarrow \quad \text{For all unit vectors } \mathbf{u},\ \lim_{t\to 0^+} \frac{f(\mathbf{a}+t\mathbf{u}) - f(\mathbf{a})}{t} = \nabla f(\mathbf{a})\cdot\mathbf{u}.$$

By using $-\mathbf{u}$ in the last line, we have

$$\begin{aligned} \text{For all unit vectors } \mathbf{u},\ \lim_{t\to 0^+} \frac{f(\mathbf{a}+t(-\mathbf{u})) - f(\mathbf{a})}{t} &= \nabla f(\mathbf{a})\cdot(-\mathbf{u}) \\ \lim_{t\to 0^+} \frac{f(\mathbf{a}+(-t)\mathbf{u}) - f(\mathbf{a})}{-t} &= \nabla f(\mathbf{a})\cdot\mathbf{u} \\ \lim_{t\to 0^-} \frac{f(\mathbf{a}+t\mathbf{u}) - f(\mathbf{a})}{t} &= \nabla f(\mathbf{a})\cdot\mathbf{u} \\ \therefore \text{Combining with above,} \quad \lim_{t\to 0} \frac{f(\mathbf{a}+t\mathbf{u}) - f(\mathbf{a})}{t} &= \nabla f(\mathbf{a})\cdot\mathbf{u}. \end{aligned}$$

In other words, f is differentiable at $\mathbf{x} = \mathbf{a}$ if and only if all directional derivatives $D_{\mathbf{u}}f(\mathbf{a})$ exist at $\mathbf{a}$ and can be calculated as $\nabla f(\mathbf{a})\cdot\mathbf{u}$.

Warning: This proof looks pretty, but it is WRONG!! This is because once I let $\mathbf{h} = t\mathbf{u}$, and shrink $t \to 0$ for each $\mathbf{u}$, I have already restricted myself to use straight paths for $\mathbf{h}$ to get to $\mathbf{0}$, but "$\lim\limits_{\mathbf{h}\to\mathbf{0}}$" needs to check every possible path. Therefore, this proof is totally worthless. The conclusion is WRONG as well. Please refer to Example 4. I am very sorry for the confusion.

Example 3. Let $f(x,y) = \begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}} & \text{if } (x,y) \neq 0 \\ 0 & \text{otherwise} \end{cases}$. We have shown in Lecture notes 9 that this function is not differentiable at $(0,0)$. Find $D_{\mathbf{u}}f(0,0)$ for all unit vectors $\mathbf{u}$, and compare with $\nabla f(0,0)\cdot\mathbf{u}$.

Solution. Every unit vector in $\mathbb{R}^2$ can be written as $\mathbf{u} = (\cos\theta, \sin\theta)$ for some θ. Hence,

$$\begin{aligned} D_{\mathbf{u}}f(0,0) &= \lim_{t\to 0} \frac{f\big((0,0) + t(\cos\theta, \sin\theta)\big) - f(0,0)}{t} \\ &= \lim_{t\to 0} \frac{\dfrac{t^2\cos\theta\sin\theta}{\sqrt{t^2\cos^2\theta + t^2\sin^2\theta}} - 0}{t} = \lim_{t\to 0} \cos\theta\sin\theta\frac{t}{|t|} \end{aligned}$$

which exists if and only if $\cos\theta\sin\theta = 0$, and at that moment, the directional derivatives are precisely the partial derivatives.

$$f_x(0,0) = D_{(1,0)}f(0,0) = \lim_{t\to 0} \cos 0 \sin 0 \frac{t}{|t|} = 0,$$

and

$$f_y(0,0) = D_{(0,1)}f(0,0) = \lim_{t\to 0} \cos\frac{\pi}{2}\sin\frac{\pi}{2}\frac{t}{|t|} = 0.$$

In other words, $\nabla f(0,0) = (0,0)$. Therefore, $\nabla f(0,0)\cdot\mathbf{u} = 0$ for all unit vectors $\mathbf{u}$, which do not agree with the directional derivatives $D_{\mathbf{u}}f(0,0)$.

□

Example 4. Let $f(x,y) = \begin{cases} 1 & \text{if } 0 < y < x^2 \\ 0 & \text{otherwise} \end{cases}$.

(a) Determine whether f is differentiable at $(0,0)$.
(b) Find $D_{\mathbf{u}}f(0,0)$ for all unit vectors $\mathbf{u}$, and compare with $\nabla f(0,0)\cdot\mathbf{u}$.

Solution. (a) $\displaystyle\lim_{\substack{(h,k)=(2t,2t^2)\\ t\to 0}} f(x,y) = 1$ since $2t^2 < (2t)^2$, while $\displaystyle\lim_{\substack{(h,k)=(t,0)\\ t\to 0}} f(x,y) = 0$. Hence, f is not continuous at $(0,0)$, implying that f is not differentiable at $(0,0)$.
(b)

$$D_{\mathbf{u}}f(0,0) = \lim_{t\to 0} \frac{f\big((0,0) + t(\cos\theta,\sin\theta)\big) - f(0,0)}{t} = \lim_{t\to 0}\frac{f\big(t(\cos\theta,\sin\theta)\big)}{t}.$$

Case 1. $\sin\theta = 0$.
$f\big(t(\cos\theta,\sin\theta)\big) = f(t,0) = 0$. Hence, the directional derivative is 0.
Case 2. $\sin\theta > 0$.
For all $t \in [0, \sin\theta]$, $(t\cos\theta)^2 = t^2\cos^2\theta \le t^2 \le t\sin\theta$, so $f\big(t(\cos\theta,\sin\theta)\big) = 0$; for all $t \in [-\sin\theta, 0)$, $t\sin\theta < 0$, so $f\big(t(\cos\theta,\sin\theta)\big) = 0$. Hence, the directional derivative is 0.
Case 3. $\sin\theta < 0$.
For all $t \in [0, -\sin\theta]$, $t\sin\theta < 0$, so $f\big(t(\cos\theta,\sin\theta)\big) = 0$; for all $t \in [\sin\theta, 0)$, $(t\cos\theta)^2 = t^2\cos^2\theta \le t^2 \le t\sin\theta$, so $f\big(t(\cos\theta,\sin\theta)\big) = 0$. Hence, the directional derivative is 0.

Therefore, $D_{\mathbf{u}}f(0,0) = 0$ for all unit vectors $\mathbf{u}$. In particular, $f_x(0,0) = f_y(0,0) = 0$, so $\nabla f(0,0)\cdot\mathbf{u} = 0 = D_{\mathbf{u}}f(0,0)$.

□

If f is differentiable at $\mathbf{x} = \mathbf{a}$, then the equation of the tangent plane of $x_{n+1} = f(\mathbf{x})$ at $\mathbf{x} = \mathbf{a}$ is

$$\begin{aligned} x_{n+1} &= f(\mathbf{a}) + \nabla f(\mathbf{a})\cdot(\mathbf{x} - \mathbf{a}) \\ &= f(\mathbf{a}) + \nabla f(\mathbf{a})\cdot\frac{\mathbf{x}-\mathbf{a}}{\|\mathbf{x}-\mathbf{a}\|}\cdot\|\mathbf{x}-\mathbf{a}\| \\ &= f(\mathbf{a}) + D_{\frac{\mathbf{x}-\mathbf{a}}{\|\mathbf{x}-\mathbf{a}\|}}f(\mathbf{a})\cdot(\|\mathbf{x}-\mathbf{a}\|). \end{aligned}$$

Theorem 5. *Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a differentiable real-valued function. At $\mathbf{a} \in D$,*
(a) the value of f increases most rapidly along the direction ∇f.
(b) the value of f decreases most rapidly along the direction $-\nabla f$.
(c) the value of f does not change along any direction $\mathbf{u}$ that is orthogonal to ∇f.

Proof. Let $\mathbf{u}$ be a unit vector. First, we note that the rate of change of f along the direction $\mathbf{u}$ at $\mathbf{x} = \mathbf{a}$ is given by $D_{\mathbf{u}}f(\mathbf{a})$. Since

$$D_{\mathbf{u}}f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{u} = \|\nabla f(\mathbf{a})\| \|\mathbf{u}\| \cos\theta = \|\nabla f(\mathbf{a})\| \cos\theta,$$

(a) the maximum increasing rate of f occurs when $\theta = 0$, i.e. $\mathbf{u}$ is in the same direction as $\nabla f(\mathbf{a})$;
(b) the maximum decreasing rate of f occurs when $\theta = \pi$, i.e. $\mathbf{u}$ is in the same direction as $-\nabla f(\mathbf{a})$;
(c) the rate of change of f is 0 when $\theta = \dfrac{\pi}{2}$, i.e. $\mathbf{u}$ is orthogonal to $\nabla f(\mathbf{a})$.

□

For $\mathbf{a} \in D$, let $f(\mathbf{a}) = k$ for some constant $k \in \mathbb{R}$. In view of Theorem 5(c), the equation of the tangent plane of the level surface $f(\mathbf{x}) = k$ at the point $\mathbf{x} = \mathbf{a}$ (in case of $n = 2$, it is the equation of the tangent line of the level curve) is

$$\nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) = 0,$$

or

$$f_{x_1}(\mathbf{a})(x_1 - a_1) + \cdots + f_{x_n}(\mathbf{a})(x_n - a_n) = 0.$$

Remark: Compare this with the equation of the tangent plane of the surface $x_{n+1} = f(\mathbf{x})$ at $\mathbf{x} = \mathbf{a}$

$$x_{n+1} = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}).$$

We can consider this surface as the level surface

$$g(\mathbf{x}, x_{n+1}) = x_{n+1} - f(\mathbf{x}) = 0,$$

then the equation of the tangent plane of this level surface at $(\mathbf{x}, x_{n+1}) = \big(\mathbf{a}, f(\mathbf{a})\big)$ is

$$\begin{aligned}
\nabla g\big(\mathbf{a}, f(\mathbf{a})\big) \cdot \big(\mathbf{x} - \mathbf{a},\ x_{n+1} - f(\mathbf{a})\big) &= 0 \\
\big(-\nabla f(\mathbf{a}), 1\big) \cdot \big(\mathbf{x} - \mathbf{a},\ x_{n+1} - f(\mathbf{a})\big) &= 0 \\
-\nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) + x_{n+1} - f(\mathbf{a}) &= 0 \\
x_{n+1} &= f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}).
\end{aligned}$$

Example 6. Let $D \subseteq \mathbb{R}^n$ be an open domain, and let $\mathbf{a} \in D$. Let $f : D \to \mathbb{R}$ be a real-valued function differentiable at $\mathbf{a}$. Let $k \in \mathbb{R}$ be a constant such that $f(\mathbf{a}) = k$. Let Γ be the level surface $f(\mathbf{x}) = k$. Let $\mathbf{x} : [a, b] \to \Gamma$ be a curve such that $\mathbf{x}(t_0) = \mathbf{a}$ for some $t_0 \in (a, b)$. Also, assume that $\mathbf{x}$ is differentiable at $t = t_0$. Show that $\nabla f(\mathbf{a})$ is orthogonal to $\mathbf{x}'(t_0)$.

Solution. Method 1 (*geometric interpretation*). $\nabla f(\mathbf{a})$ is normal to the tangent plane of the level surface $f(\mathbf{x}) = k$ at $\mathbf{x} = \mathbf{a}$. Since $\mathbf{x}(t)$ is a curve on the level surface $f(\mathbf{x}) = k$, and $\mathbf{x}'(t_0)$ is a tangent vector of the level surface at $\mathbf{x} = \mathbf{a}$, i.e. $\mathbf{x}'(t_0)$ lies on the tangent plane. Therefore, $\nabla f(\mathbf{a})$ is orthogonal to $\mathbf{x}'(t_0)$.
Method 2 (*chain rule*). First, note that $f(\mathbf{x}(t)) = k$, which is the $\mathbb{R} \to \mathbb{R}^n \to \mathbb{R}$ case. By chain rule at $t = t_0$, we have

$$\nabla f(\mathbf{a}) \cdot \mathbf{x}'(t_0) = 0.$$

□

Example 7. Let $f(x,y,z) = x^2 - xy + \frac{1}{2}y^2 + 3$.
(a) Find the equation of the tangent plane to the surface $z = f(x,y)$ at $(3,2)$.
(b) Find the equation of the tangent line to the level curve $f(x,y) = 8$ at $(3,2)$.

Solution. (a) The equation is

$$\begin{aligned}
z &= f(3,2) + \nabla f(3,2) \cdot (x-3, y-2) \\
z &= 8 + (2x - y, -x + y)\big|_{(x,y)=(3,2)} \cdot (x-3, y-2) \\
z &= 8 + (4,-1) \cdot (x-3, y-2) \\
z &= 8 + 4x - 12 - y + 2 \\
z &= 4x - y - 2.
\end{aligned}$$

(b) The equation is

$$\begin{aligned}
\nabla f(3,2) \cdot (x-3, y-2) &= 0 \\
4x - 12 - y + 2 &= 0 \\
4x - y - 10 &= 0.
\end{aligned}$$

□

As the gradient operator ∇ is so useful, it will be helpful to know some of its properties. Let $f, g : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be two real-valued functions such that $\nabla f(\mathbf{a})$ and $\nabla g(\mathbf{a})$ exist for some $\mathbf{a} \in D$, and let $k \in \mathbb{R}$, then

- (sum and difference rule) $\nabla(f \pm g)(\mathbf{a}) = \nabla f(\mathbf{a}) \pm \nabla g(\mathbf{a})$.
- (product rules)
  - $\nabla(kf)(\mathbf{a}) = k\nabla f(\mathbf{a})$.
  - $\nabla(fg)(\mathbf{a}) = f(\mathbf{a})\nabla g(\mathbf{a}) + g(\mathbf{a})\nabla f(\mathbf{a})$.
- (quotient rule) $\nabla\left(\frac{f}{g}\right)(\mathbf{a}) = \dfrac{g(\mathbf{a})\nabla f(\mathbf{a}) - f(\mathbf{a})\nabla g(\mathbf{a})}{g^2(\mathbf{a})}$ if $g(\mathbf{a}) \neq 0$.

14.6 — Tangent planes and differentials

Let $D \subseteq \mathbb{R}^n$ be an open domain. Let $f : D \to \mathbb{R}$ be a differentiable function at $\mathbf{a} \in D$. We have already derived the formula for tangent plane of the surface $x_{n+1} = f(\mathbf{x})$ at $\mathbf{x} = \mathbf{a}$ in Section 14.3 in Lecture notes 8 as

$$x_{n+1} = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a})$$

or

$$x_{n+1} = f(\mathbf{a}) + f_{x_1}(\mathbf{a})(x_1 - a_1) + \cdots + f_{x_n}(\mathbf{a})(x_n - a_n).$$

Also, from our discussions in Section 14.3, the tangent plane is a **linearization** of the function $x_{n+1} = f(\mathbf{x})$ near $\mathbf{x} = \mathbf{a}$, i.e.

$$L(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}),$$

and $f(\mathbf{x}) \approx L(\mathbf{x})$ is the **standard linear approximation** of f at $\mathbf{x} = \mathbf{a}$. The upper bound of the error of this approximation is given by

$$E(\mathbf{x}) = |f(\mathbf{x}) - L(\mathbf{x})| \leq \frac{1}{2}M\big(|x_1 - a_1| + \cdots + |x_n - a_n|\big)^2,$$

where M is an upper bound of all $|f_{x_i x_j}|$ over D, $1 \leq i, j \leq n$, if it exists.

Example 8. Let $f(x,y,z) = x^2 - xy + 3\sin z$.
(a) Find the linearization $L(x,y,z)$ at the point $(2,1,0)$.
(b) If we are to use $L(x,y,z)$ to approximate $f(x,y,z)$ in

$$R = \{(x,y,z) : |x-2| \le 0,01, |y-1| \le 0,02, |z| \le 0.01\},$$

find an upper bound of the error.

Solution. (a)

$$\begin{aligned} L(2,1,0) &= f(2,1,0) + \nabla f(2,1,0) \cdot (x-2, y-1, z-0) \\ &= 2 + (2x-y, -x, 3\cos z)\big|_{(x,y,z)=(2,1,0)} \cdot (x-2, y-1, z) \\ &= 2 + (3,-2,3) \cdot (x-2, y-1, z) \\ &= 2 + 3x - 6 - 2y + 2 + 3z \\ &= 3x - 2y + 3z - 2. \end{aligned}$$

(b) $f_{xx} = 2$, $f_{yy} = 0$, $f_{zz} = -3\sin z$, $f_{xy} = f_{yx} = -1$, $f_{yz} = f_{zy} = 0$, $f_{xz} = f_{zx} = 0$. The maximum of all these functions in R is 2, so we can take $M = 2$.

$$E(x,y,z) \le \frac{1}{2}M(|x-2| + |y-1| + |z-0|)^2 \le \frac{1}{2}(2)(0.01 + 0.02 + 0.01)^2 = 0.0016.$$

□

If we rearrange the terms in

$$x_{n+1} = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a})$$

into

$$x_{n+1} - f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}),$$

we obtain the following two results in differentials.

(1) $df = \nabla f(\mathbf{a}) \cdot \mathbf{u}\, ds$, where ds denotes the small distance from $\mathbf{a}$ in the direction of the unit vector $\mathbf{u}$.
(2) $df = f_{x_1}(\mathbf{a})\, dx_1 + \cdots + f_{x_n}(\mathbf{a})\, dx_n$, which is called the **total differential** of f at $\mathbf{x} = \mathbf{a}$.

Example 9. A cylinder has radius 1 and height 5. If the radius increases by 0.03, and the height decreases by 0.1, estimate the absolute change in the volume of the can.

Solution. Let r denote the radius, h denote the height, and V denote the volume of the cylinder. Then $V = \pi r^2 h$. The total differential of f at $(r,h) = (1,5)$ is

$$\begin{aligned} dV &= (2\pi r h)\big|_{(r,h)=(1,5)}\, dr + (\pi r^2)\big|_{(r,h)=(1,5)}\, dh \\ &= 10\pi dr + \pi dh. \end{aligned}$$

Now, $dr \approx 0.03$ and $dh \approx -0.1$, so

$$dV \approx 10\pi(0.03) + \pi(-0.1) = 0.2\pi.$$

□