13.5 DIRECTIONAL DERIVATIVES and the GRADIENT VECTOR

Directional Derivatives

In Section 13.3 the partial derivatives f_x and f_y were defined as

$$\mathbf{f_x(x,y)} = \lim_{h\to 0} \frac{f(x+h,y)-f(x,y)}{h} \text{ (if the limit exists and is finite)} \quad \text{and}$$

$$\mathbf{f_y(x,y)} = \lim_{h\to 0} \frac{f(x,y+h)-f(x,y)}{h} \text{ (if the limit exists and is finite)}.$$

These partial derivative measured the instantaneous rate of change of $z = f(x,y)$ as we moved in the increasing x–direction (while holding y constant) and in the increasing y–direction (while holding x constant). Sometimes, however, we are interested in the rate of change of $z = f(x,y)$ as we move in some other direction, and that leads to the idea of a **directional derivative** to measure the instantaneous rate of change of $z = f(x,y)$ as we move in any direction.

Definition:
The **directional derivative** of $z = f(x,y)$ in the direction of a unit vector $\mathbf{u} = \langle a, b \rangle$ is

$$D_{\mathbf{u}}f(x,y) = \lim_{h\to 0} \frac{f(x+ah,y+bh)-f(x,y)}{h} \text{ (if the limit exists and is finite)}.$$

Figures 1a and 1b illustrate the slope of the line tangent to the curve where the plane above the vector u intersects the surface $z = f(x,y)$.

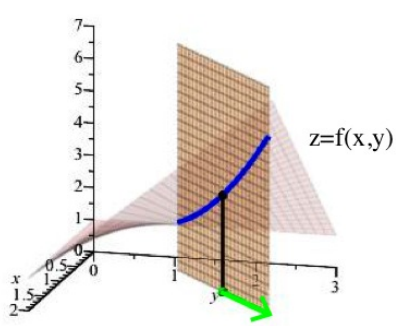

Fig. 1a

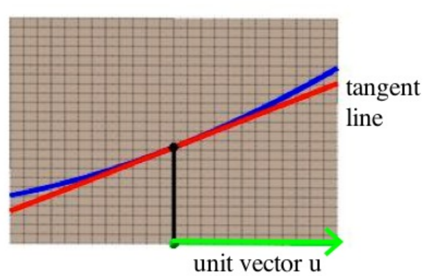

Fig. 1b

Using this definition can be tedious and algebraically messy, but it is worth doing once. Fortunately we will soon see a much more efficient method.

Example 1: Use this definition to calculate $D_{\mathbf{u}}f(1,2)$ for $f(x,y) = 1 + xy$ and $\mathbf{u} = \langle 0.6, 0.8 \rangle$

Solution: $D_{\mathbf{u}}f(1,2) = \lim_{h\to 0} \dfrac{f(1+0.6h, 2+0.8h) - f(1,2)}{h} = \lim_{h\to 0} \dfrac{[1+(1+0.6h)(2+0.8h)]-3}{h}$

$$\lim_{h\to 0} \frac{[1+2+2(0.6)h+0.8h]-3}{h} = 2(0.6)+0.8 = 2$$

Practice 1: Calculate $D_{\mathbf{u}}f(2,1)$ for $f(x,y) = x^2 + 2x + 3y + 1$ and $\mathbf{u} = \left\langle \frac{5}{13}, \frac{12}{13} \right\rangle$.

For more complicated functions it becomes extremely difficult to use the definition to calculate directional derivatives. Fortunately there is a much easier way given by the next theorem.

Theorem:
If $f(x,y)$ is differentiable,
then f has a directional derivative in the direction of every unit vector $\mathbf{u} = \langle a, b \rangle$, and

$$D_{\mathbf{u}}f(x,y) = f_x(x,y)\, a + f_y(x,y)\, b .$$

A proof of this result uses a multivariable "Chain Rule" that we will discuss in Section 13.8.

Example 2: Verify that this theorem gives the same answer for $D_{\mathbf{u}}f(x,y)$ as our use of the directional derivative definition in Example 1.

Solution: $f_x(x,y)=y$ and $f_y(x,y)=x$ so $f_x(1,2)= 2$ and $f_y(1,2)= 1$. $\mathbf{u} = \langle 0.6, 0.8 \rangle$ so

$D_{\mathbf{u}}f(1,2) = (2)(0.6) + (1)(0.8) = 2$, the same result as in Example 1.

Practice 2: Verify that this theorem gives the same answer for $D_{\mathbf{u}}f(x,y)$ as your use of the directional derivative definition in Practice 1.

Example 3: Calculate the directional derivatives of $z = f(x,y) = x + 5x^2y^3$ at the point $(2,1)$ in the directions of the unit vectors (a) $\mathbf{u} = \langle 0.6, 0.8 \rangle$, (b) $\mathbf{u} = \langle -0.6, -0.8 \rangle$, (c) $\mathbf{u} = \langle 0.8, 0.6 \rangle$, (d) $\mathbf{u} = \mathbf{i} = \langle 1, 0 \rangle$, and (e) $\mathbf{u} = \mathbf{j} = \langle 0, 1 \rangle$.

Solution: $f_x(x,y) = 1 + 10xy^3$ and $f_y(x,y) = 15x^2y^2$ so $f_x(2,1) = 21$ and $f_y(2,1) = 60$. Then, by the previous Theorem,

(a) for $\mathbf{u} = \langle 0.6, 0.8 \rangle$, $D_{\mathbf{u}}f(2,1) = f_x(2,1)\, a + f_y(2,1)\, b = (21)(0.6) + (60)(0.8) = 60.6$.

(b) for $\mathbf{u} = \langle -0.6, -0.8 \rangle$, $D_{\mathbf{u}}f(2,1) = (21)(-0.6) + (60)(-0.8) = -60.6$.

(c) for $\mathbf{u} = \langle 0.8, 0.6 \rangle$, $D_{\mathbf{u}}f(2,1) = (21)(0.8) + (60)(0.6) = 52.8$.

(d) for $\mathbf{u} = \mathbf{i} = \langle 1, 0 \rangle$, $D_{\mathbf{u}}f(2,1) = (21)(1) + (60)(0) = 21$ ($= f_x(2,1)$) .

(e) for $\mathbf{u} = \mathbf{j} = \langle 0, 1 \rangle$, $D_{\mathbf{u}}f(2,1) = (21)(0) + (60)(1) = 60$ ($= f_y(2,1)$) .

The Gradient Vector

You might have noticed that the pattern $D_{\mathbf{u}}f(x,y) = f_x(x,y)\, a + f_y(x,y)\, b$ for calculating the directional derivative can be viewed as the dot product of the vector $\langle f_x(x,y), f_y(x,y) \rangle$ with the unit direction vector $\mathbf{u} = \langle a, b \rangle$. This vector $\langle f_x(x,y), f_y(x,y) \rangle$ shows up in a variety of contexts and is called the **gradient** of f.

Definition of the **Gradient Vector:**

The **gradient vector** of $f(x,y)$ is $\nabla f(x,y) = \langle f_x(x,y), f_y(x,y) \rangle = \frac{\partial f}{\partial x}\,\mathbf{i} + \frac{\partial f}{\partial y}\,\mathbf{j}$.

The symbol " ∇f " is read as "grad f" or "del f."

Example 4: Calculate $\nabla f(x,y)$ and $\nabla f(0,1)$ for (a) $f(x,y) = x + 5x^2y^3$ and (b) $f(x,y) = y^2 + \sin(x)$.

Solution: (a) For $f(x,y) = x + 5x^2y^3$, $f_x(x,y) = 1 + 10xy^3$ and $f_y(x,y) = 15x^2y^2$ so

$\nabla f(x,y) = \langle 1 + 10xy^3, 15x^2y^2 \rangle$ and $\nabla f(0,1) = \langle 1, 0 \rangle$.

(b) For $f(x,y) = y^2 + \sin(x)$, $f_x(x,y) = \cos(x)$ and $f_y(x,y) = 2y$ so

$\nabla f(x,y) = \langle \cos(x), 2y \rangle$ and $\nabla f(0,1) = \langle 1, 2 \rangle$.

Practice 3: Calculate $\nabla f(x,y)$ and $\nabla f(1,2)$ for (a) $f(x,y) = e^{xy} + 2x^3y + y^2$ and (b) $f(x,y) = \cos(2x+3y)$.

The directional derivative can now be written simply as the dot product of the gradient and the unit direction vector .

$$D_{\mathbf{u}}f(x,y) = \nabla f(x,y) \cdot \mathbf{u}$$

The gradient vector $\nabla f(x,y)$ is useful for much more than a compact notation for directional derivatives. $\nabla f(x,y)$ has a number of special features that make it useful for investigating the behavior of surfaces.

Three very important properties of the gradient vector ∇f(x,y) :

(1) At a point (x,y), the maximum value of the directional derivative $D_{\mathbf{u}}f(x,y)$ is $|\nabla f(x,y)|$.

(2) At a point (x,y), the maximum value of $D_{\mathbf{u}}f(x,y)$ occurs when **u** has the same direction as the gradient vector $\nabla f(x,y)$. (At each point (x,y), the gradient vector $\nabla f(x,y)$ "points" in the direction of maximum increase for f(x,y).)

(3) At a point (x,y), the gradient vector $\nabla f(x,y)$ is normal (perpendicular) to the level curve that goes through the point (x,y).

One of the beauties of mathematics is that sometimes a result like the powerful and non-obvious properties of the gradient can be proven in rather simple ways.

Proof of (1) and (2):
$$D_{\mathbf{u}}f(x,y) = \nabla f(x,y) \cdot \mathbf{u}$$
$$= |\nabla f(x,y)|\,|\mathbf{u}|\cos(\theta) \qquad \theta \text{ is the angle between the vectors } \nabla f(x,y) \text{ and } \mathbf{u}$$
$$= |\nabla f(x,y)|\cos(\theta)$$

The maximum value of $\cos(\theta)$ is 1 (when $\theta = 0$), so the maximum value of $D_{\mathbf{u}}f(x,y)$ is $|\nabla f(x,y)|$ and this maximum occurs when $\theta = 0$, when $\nabla f(x,y)$ and **u** have the same direction.

Proof of (3): f is constant along the level curve at the point (x,y) so $D_{\mathbf{u}}f(x,y) = 0$ when **u** is the direction of the level curve. But $D_{\mathbf{u}}f(x,y) = \nabla f(x,y) \cdot \mathbf{u}$ so $\nabla f(x,y)$ is perpendicular to **u**, the direction of the level curve.

Example 5: Find (a) the maximum rate of change of $f(x,y) = xe^y$ at the point (2,0,) and (b) the direction in which this maximum rate of change occurs.

Solution: $f_x(x,y) = e^y$ and $f_y(x,y) = xe^y$ so $f_x(2,0) = e^0 = 1$ and $f_y(2,0) = 2e^0 = 2$.

(a) The maximum value of the rate of change of f is $D_{\mathbf{u}}f(x,y) = |\nabla f(1,2)| = |\langle 1, 2 \rangle| = \sqrt{5}$.

(b) This maximum value occurs when **u** is in the direction of $\nabla f(1,2)$: $\mathbf{u} = \langle 1/\sqrt{5}, 2/\sqrt{5} \rangle$.

Practice 4: Find (a) the maximum rate of change of $f(x,y) = \sqrt{2x+3y}$ at the point (5, 2) and (b) the direction in which this maximum rate of change occurs.

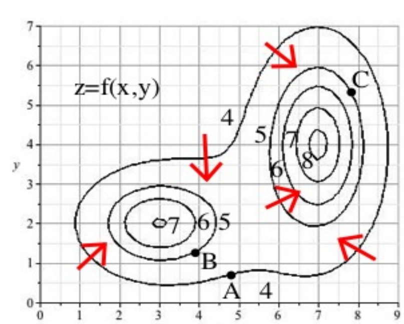

Fig. 2

Fig. 2 shows several level curves for a function z = f(x,y) and the gradient vector at several locations. (Note: the lengths of these gradient vectors are exaggerated.)

Practice 5: Sketch the gradient vector $\nabla f(x,y)$ for the function f in Fig. 2 at A, B and C.

A ball placed at (x,y) will begin to roll in the direction $u = -\nabla f(x,y)$.

Climbing to a (local) maximum

Property (2) is the foundation for using the gradient vector $\nabla f(x,y)$ in iterative methods for finding local maximums of functions of several variables:

(i) At any point (x,y) we take a short "step" in the direction of $\nabla f(x,y)$ — this takes us "uphill" along the steepest route at that point.

(ii) Repeat step (i) until a (local) maximum is reached.

Property (3) provides an easy way to geometrically determine the direction of the gradient from the level curves of a surface.

Fig. 3 shows level curves for a function z = f(x,y) and the "uphill gradient" paths for several starting points.

Practice 6: Sketch the "uphill gradient" path for the function f in Fig. 3 at starting points A, B and C.

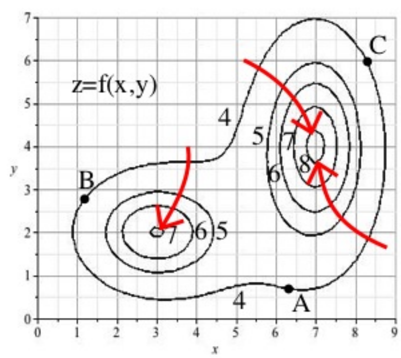

Fig. 3

Beyond z = f(x,y)

So far all of the examples have dealt with functions of two variables, z = f(x,y), but that was just for convenience. The definitions and ideas of gradient vectors and directional derivatives and their properties extend in a natural way to functions of three (or more) variables.

Extensions to w = f(x,y,z)

Definition: $\nabla f(x,y,z) = \langle f_x(x,y,z), f_y(x,y,z), f_z(x,y,z) \rangle = \frac{\partial f}{\partial x}\,\mathbf{i} + \frac{\partial f}{\partial y}\,\mathbf{j} + \frac{\partial f}{\partial z}\,\mathbf{k}$.

Theorem: For a differentiable function f(x,y,z) and a unit direction vector $\mathbf{u} = \langle a, b, c \rangle$,

$$D_{\mathbf{u}} f(x,y,z) = \nabla f(x,y,z) \cdot \mathbf{u}$$

Features: (1) The maximum value of the directional derivative $D_{\mathbf{u}} f(x,y,z)$ is $|\nabla f(x,y,z)|$.

(2) The maximum value of $D_{\mathbf{u}} f(x,y,z)$ occurs when $\mathbf{u}$ has the same direction as $\nabla f(x,y,z)$.

(2) The gradient vector $\nabla f(x,y,z)$ is normal (perpendicular) to the level surface through the point (x,y,z).

These same ideas also extend very naturally to functions of more than three variables.
For example, if x, y and z (all in meters) give the location in a room then $w=f(x,y,z)$ could be the temperature (ºC) at that location. Then instead of a level curve (in 2D), the points (x,y,z) where $w=70$ would be a level **surface** in 3D. $D_{\mathbf{u}}f(x,y,z)$ would be the instantaneous rate of change of temperature at location (x,y,z) in the direction $\mathbf{u}$, and the units of $D_{\mathbf{u}}f(x,y,z)$ would be ºC/m . The gradient vector $\nabla f(x,y,z)$ would still give the maximum value of the directional derivative and would point in the direction of maximum rate of temperature increase. A heat-seeking flying bug in the room would follow a path in the direction of the gradient at each point.

PROBLEMS

In problems 1 – 4, find the directional derivative of f at the given point in the direction indicated by the given angle θ. (Note: θ is the angle the direction vector $\mathbf{u}$ makes with the positive x–axis, so the components of $\mathbf{u}$ are $\langle \cos(\theta), \sin(\theta) \rangle$.)

1. $f(x, y) = x^2y^2 + 2x^4y$ at $(1, -2)$ with $\theta = \pi/3$
2. $f(x, y) = (x^2 - y)^3$ at $(3, 1)$ with $\theta = 3\pi/4$
3. $f(x, y) = y^x$ at $(1, 2)$ with $\theta = \pi/2$
4. $f(x, y) = \sin(x + 2y)$ at $(4, -2)$ with $\theta = -2\pi/3$

In problems 5 – 8, (a) find the gradient of f, (b) evaluate the gradient at the given point P, and (c) find the rate of change of f at P in the direction of the given vector $\mathbf{u}$.

5. $f(x, y) = x^3 - 4x^2y + y^2$ at $P = (0, -1)$ with $\mathbf{u} = \langle 3/5, 4/5 \rangle$.
6. $f(x, y) = e^x \cdot \sin(y)$ at $P = (1, \pi/4)$ with $\mathbf{u} = \langle -1/\sqrt{5}, 2/\sqrt{5} \rangle$.
7. $f(x, y, z) = xy^2z^3$ at $P = (1, -2, 1)$ with $\mathbf{u} = \langle 1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3} \rangle$.
8. $f(x, y, z) = xy + yz^2 + xz^3$ at $P = (2, 0, 3)$ with $\mathbf{u} = \langle -2/3, -1/3, 2/3 \rangle$.

In problems 9 – 14, find the directional derivative of the given function at the given point in the direction of the given vector $\mathbf{v}$.

9. $f(x, y) = \sqrt{x - y}$ at $(5, 1)$ with $\mathbf{v} = \langle 12, 5 \rangle$.
10. $f(x, y) = x/y$ at $(6, -2)$ with $\mathbf{v} = \langle -1, 3 \rangle$.
11. $g(x, y) = x \cdot e^{xy}$ at $(-3, 0)$ with $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$.
12. $g(x, y) = e^x \cos(y)$ at $(1, \pi/6)$ with $\mathbf{v} = \mathbf{i} - \mathbf{j}$.
13. $f(x, y, z) = \sqrt{xyz}$ at $(2, 4, 2)$ with $\mathbf{v} = \langle 4, 2, -4 \rangle$.
14. $f(x, y, z) = z^3 - x^2y$ at $(1, 6, 2)$ with $\mathbf{v} = \langle 3, 4, 12 \rangle$.

In problems 15 – 20, find the maximum rate of change of f at the given point and the direction in which it occurs.

15. $f(x, y) = x \cdot e^{-y} + 3y$ at the point $(1, 0)$

16. $f(x, y) = \ln(x^2 + y^2)$ at the point $(1, 2)$

17. $f(x, y) = \sqrt{x^2 + 2y}$ at the point $(4, 10)$

18. $f(x, y, z) = x + y/z$ at the point $(4, 3, -1)$

19. $f(x, y) = \cos(3x + 2y)$ at the point $(\pi/6, -\pi/8)$

20. $f(x, y, z) = \frac{x}{y} + \frac{y}{z}$ at the point $(4, 2, 1)$

21. At each dot in Fig. 4 sketch the gradient vector.

22. At each dot in Fig. 5 sketch the gradient vector.

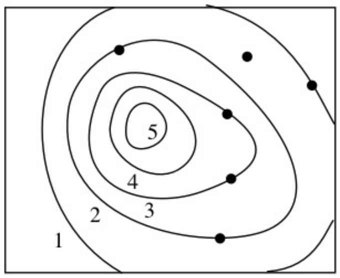

Fig. 4

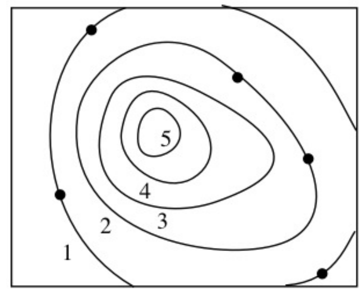

Fig. 5

23. At each dot in Fig. 6 sketch the "uphill gradient" path.

24. At each dot in Fig. 7 sketch the "uphill gradient" path.

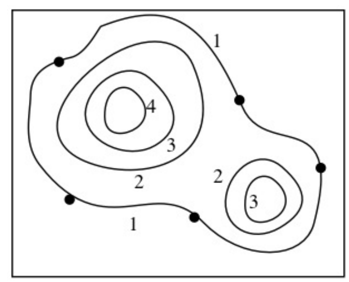

Fig. 6

25. Show that a differentiable function f decreases most rapidly at (x,y) in the direction opposite to the gradient vector, that is, in the direction $-\nabla f(x,y)$.

26. Use the result of Problem 23 to find the direction in which the function $f(x,y) = x^4y - x^2y^3$ decreases fastest at the point $(2, -3)$.

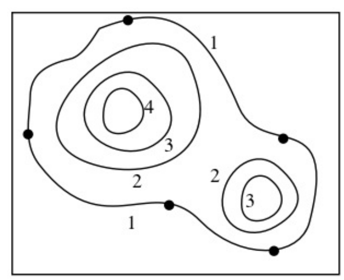

Fig. 7

27. The temperature T in a metal ball is inversely proportional to the distance from the center of the ball, which we take to be the origin. The temperature at the point $(1, 2, 2)$ is 120^o.

    (a) Find the rate of change of T at $(1, 2, 2)$ in the direction toward the point $(2, 1, 3)$.

    (b) Show that at any given point in the ball the direction of greatest increase in temperature is given by a vector that points toward the origin.

28. The temperature at a point (x, y, z) is given by $T(x,y,z) = 200 \cdot e^{(-x^2 - 3y^2 - 9z^2)}$ where T is measured in ^oC and x, y, and z in meters.

    (a) Find the rate of change of temperature at the point $P(2, -1, 2)$ in the direction toward the point $(3, -3, 3)$.

    (b) In which direction does the temperature increase fastest at P?

(c) Find the maximum rate of increase at P.

29. Suppose that over a certain region of space the electrical potential V is given by $V(x,y,z) = 5x^2 - 3xy + xyz$.

(a) Find the rate of change of the potential at P(3, 4, 5) in the direction of the vector $\mathbf{v} = \mathbf{i} + \mathbf{j} - \mathbf{k}$.

(b) In which direction does V change most rapidly at P?

(c) What is the maximum rate of change at P?

30. Suppose that you are climbing a hill whose shape is given by the equation $z = 1000 - 0.01x^2 - 0.02y^2$ and you are standing at the point with coordinates (60, 100, 764).

(a) In which direction should you proceed initially in order to reach the top of the hill fastest?

(b) If you climb in that direction, at what angle above the horizontal will you be climbing initially?

31. Let F be a function of two variables that has continuous partial derivatives and consider the points A(1,3), B(3,3), C(1,7), and D(6,15). The directional derivative of A in the direction of the vector AB is 3, and the direction derivative at A in the direction of AC is 26. Find the direction derivative of f at A in the direction of the vector AD.

Practice Answers

Practice 1: $f(x,y) = x^2 + 2x + 3y + 1$ and $\mathbf{u} = \left\langle \frac{5}{13}, \frac{12}{13} \right\rangle$

$$D_{\mathbf{u}} f(2,1) = \lim_{h \to 0} \frac{f\left(2 + \frac{5}{13}h, 2 + \frac{12}{13}h\right) - f(2,1)}{h} = \lim_{h \to 0} \frac{\left[\left(2 + \frac{5}{13}h\right)^2 + 2\left(2 + \frac{5}{13}h\right) + 3\left(1 + \frac{12}{13}h\right) + 1\right] - 12}{h}$$

$$= \lim_{h \to 0} \frac{\left[\left(4 + \frac{20}{13}h + \frac{25}{13}h^2\right) + \left(4 + \frac{10}{13}h\right) + \left(3 + \frac{36}{13}h\right) + 1\right] - 12}{h} = \lim_{h \to 0} \frac{\frac{66}{13}h + \frac{25}{13}h^2}{h} = \frac{66}{13}$$

Practice 2: $f_x(x,y) = 2x+2$ and $f_y(x,y) = 3$ so $f_x(2, 1) = 6$ and $f_y(2, 1) = 3$. $\mathbf{u} = \left\langle \frac{5}{13}, \frac{12}{13} \right\rangle$ so

$D_{\mathbf{u}} f(2, 1) = (6)\left(\frac{5}{13}\right) + (3)\left(\frac{12}{13}\right) = \frac{66}{13}$, the same result as in Example 2 but much easier.

Practice 3: (a) For $f(x,y) = e^{xy} + 2x^3y + y^2$, $f_x(x,y) = y \cdot e^{xy} + 6x^2y$ and $f_y(x,y) = x \cdot e^{xy} + 2x^3 + 2y$ so

$\nabla f(x,y) = \left\langle y \cdot e^{xy} + 6x^2y,\ x \cdot e^{xy} + 2x^3 + 2y \right\rangle$ and $\nabla f(1,2) = \left\langle 2e^2 + 12,\ e^2 + 6 \right\rangle$.

(b) For $f(x,y) = \cos(2x+3y)$, $f_x(x,y) = -2\sin(2x+3y)$ and $f_y(x,y) = -3\sin(2x+3y)$ so

$\nabla f(1, 2) = \left\langle -2\sin(2x+3y), -3\sin(2x+3y) \right\rangle$ and $\nabla f(1, 2) = \left\langle -2\sin(8), -3\sin(8) \right\rangle$.

Practice 4: $f_x(x,y) = \frac{1}{\sqrt{2x+3y}}$ and $f_y(x,y) = \frac{3}{2\sqrt{2x+3y}}$ so $f_x(5, 2) = \frac{1}{4}$ and $f_y(5, 2) = \frac{3}{8}$.

(a) The maximum value of the rate of change of f is
$D_{\mathbf{u}}f(5,2) = |\nabla f(5,2)| = \left\langle \frac{1}{4}, \frac{3}{8} \right\rangle = \frac{\sqrt{13}}{8} \approx 0.45$.

(b) This maximum value occurs when **u** is in the direction of $\nabla f(5,2)$: $\mathbf{u} = \frac{\nabla f(5,2)}{|\nabla f(5,2)|} = \left\langle \frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}} \right\rangle$.

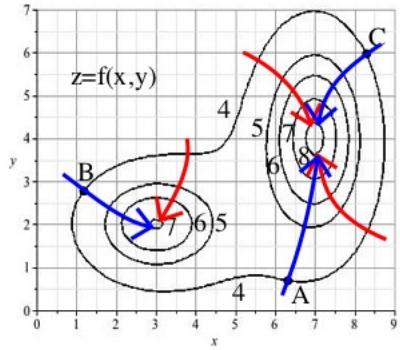

Fig. 8

Practice 5: See Fig. 9. Note that each gradient vector is perpendicular to the level curve and points uphill.

Practice 6: See Fig. 8. Note that "uphill gradient" path is always perpendicular to the level curves.

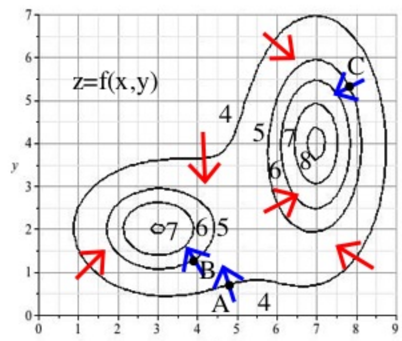

Fig. 9

Selected Answers

1. $-4 + 14\sqrt{3}$ 3. 1

5. (a) $\nabla f(x,y) = \langle 3x^2 - 8xy, -4x^2 + 2y \rangle$ (b) $\langle 0, -2 \rangle$
(c) $-8/5$

7. (a) $\nabla f(x,y,z) = \langle y^2z^3, 2xyz^3, 3xy^2z^2 \rangle$ (b) $\langle 4, -4, 12 \rangle$
(c) $20/\sqrt{3}$

9. 7/52 11. $29/\sqrt{13}$

13. 1/6 15. $\sqrt{5}$, $\langle 1/\sqrt{5}, 2/\sqrt{5} \rangle$

17. $\sqrt{17}\ /6$, $\langle 4/\sqrt{17}, 1/\sqrt{17} \rangle$

19. $\sqrt{(13/2)}$, $\langle -3/\sqrt{13}, -2/\sqrt{13} \rangle$

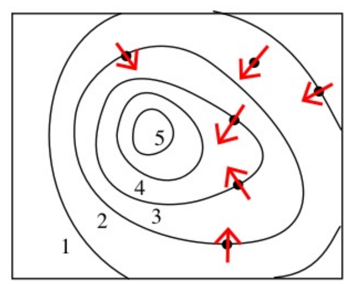
Fig. 10

21. See Fig. 10. Note that each gradient vector is perpendicular to the level curve and points uphill.

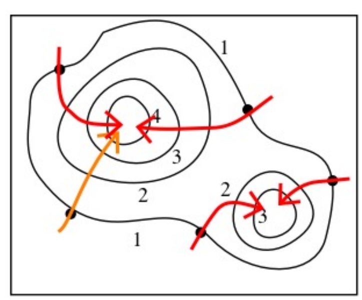
Fig. 11

23. See Fig. 11. Note that "uphill gradient" path is always perpendicular to the level curves.

27. (a) $-40/(3\sqrt{3})$

29. (a) $32/\sqrt{3}$ (b) $u = \langle 38, 6, 12 \rangle/(2\sqrt{406})$ (c) $2\sqrt{406}$

31. 327/13