Chapter 14.5: Gradients and Directional Derivatives

D1: Computing Derivatives. I can compute partial derivatives, total derivatives, directional derivatives, and gradients. I can use the Chain Rule for multivariable functions to compute derivatives of composite functions.

A1: Interpreting Derivatives. I can interpret the meaning of a partial derivative, a gradient, or a directional derivative of a function at a given point in a specified direction, including in the context of a graph or a contour plot.

Mechanics

1. Compute the gradients of each of the following:

   (a) $f(x, y) = x^2y$

   (b) $f(x, y, z) = x\sin(y) + z\cos(x)$

   (c) $f(x, y, z, w) = x^2 + y^2 + z^2 + w^2$

   Answer:

   (a) $\langle 2xy, x^2\rangle$

   (b) $\langle \sin(y) - z\sin(x), x\cos(y), \cos(x)\rangle$

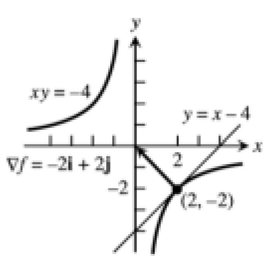
   (c) $\langle 2x, 2y, 2z, 2w\rangle$

2. Find the derivative of $g(x, y) = \dfrac{x - y}{xy + 2}$ at $(1, -1)$ in the direction of $\langle 12, 5\rangle$

   Answer: $D_{\mathbf{u}}g(1, -1) = \dfrac{21}{13}$

3. Let $f(x, y) = xy$. Sketch the curve $f(x, y) = -4$ together with $\nabla f(2, -2)$ and the tangent line at $(2, -2)$. Then, find an equation for the tangent line. What do you notice?

   Answer: Tangent line: $-2(x - 2) + 2(y + 2) = 0$



Applications

4. Gradients form the basis for "learning" in machine learning through a process called *gradient descent*. Here is a setup and overview of the thematic ideas: we are interested in minimizing a (differentiable) error function $\mathcal{L}(x, y, z)$ (*eg., $\mathcal{L}$ might represent the difference between a predicted quantity vs. the true value*). Though we are able to plug in points, the function $\mathcal{L}$ may be difficult to write down, thus we cannot do "regular calculus" (eg., first derivative test) with it.

   Gradients give us a workaround to approximate a local minimum as follows: start by randomly choosing point (x_0, y_0, z_0). Calculate the gradient at this point, and then take a small step in the direction of $-\nabla\mathcal{L}$ [why?] to arrive at a new point (x_1, y_1, z_1). Next, repeat this process with successively smaller step sizes. It turns out that if we choose our initial point and step sizes cleverly, we are (sometimes) able to get closer and closer to a local minimum, without having much knowledge of the function $\mathcal{L}$! Can you anticipate some shortcomings of this algorithm?

   Answer: Answers will vary. May include issues with stability or overshooting.

5. Suppose you are climbing a hill whose shape is given by the equation

$$z = 1000 - 0.005x^2 - 0.01y^2,$$

   where x, y, and z are measured in meters, and you are standing at a point with coordinates $(60, 40, 966)$. The positive x-axis points east and the positive y-axis points north.

   (a) If you walk due south, will you start to ascend or descend? At what rate?

   (b) If you walk northwest, will you start to ascend or descend? At what rate?

   (c) In which direction is the slope largest? What is the rate of ascent in that direction?

   Answer:

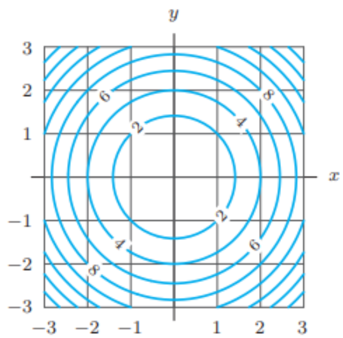
   (a) Ascend at a rate of 0.8 vertical meters per horizontal meter

   (b) Descend at a rate of $\sqrt{2}/10$ vertical meters per horizontal meter

   (c) $\langle -0.6, -0.8\rangle$ is the direction of largest slope with rate of ascent 1 vertical meter per horizontal meter.

Extensions

6. Use the contour diagram of the differentiable function f given below to decide if the specified directional derivative is positive, negative, or approximately zero.



(a) At the point $(-2, 2)$ in the direction $\mathbf{i}$

(b) At the point $(0, -2)$ in the direction $\mathbf{j}$

(c) At the point $(-1, 1)$ in the direction $\mathbf{i} + \mathbf{j}$

(d) At the point $(-1, 1)$ in the direction $-\mathbf{i} + \mathbf{j}$

(e) At the point $(0, -2)$ in the direction $\mathbf{i} - 2\mathbf{j}$

Answer:

(a) Negative

(b) Negative

(c) Approximately zero

(d) Positive

(e) Positive

7. Let $f(x, y) = -x^2y + xy^2 + xy$ and $P = (2, 1)$.

(a) Find the direction of maximal increase of f at P.

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

(c) Find the direction of maximal decrease of f at P.

(d) Find a direction $\mathbf{u}$ such that $D_{\mathbf{u}}f(P) = 0$ (note this forces $\mathbf{u}$ to be a unit vector!).

Answer:

(a) $\langle -1/\sqrt{2}, 1/\sqrt{2} \rangle$

(b) $2\sqrt{2}$

(c) $\langle 1/\sqrt{2}, -1/\sqrt{2} \rangle$

(d) $\langle 1/\sqrt{2}, 1/\sqrt{2} \rangle$