10.6 Directional Derivatives and Gradients

In this section, we will introduce the concepts of the directional derivative and gradient.

The Directional Derivative

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. The **directional derivative of f at the point $\mathbf{x}$ in the direction of the unit vector $\hat{\mathbf{u}}$** denoted by $D_{\hat{\mathbf{u}}}f(\mathbf{x})$ and is defined as

$$D_{\hat{\mathbf{u}}}f(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x} + h\hat{\mathbf{u}}) - f(\mathbf{x})}{h}.$$

In particular, if $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at the point $\mathbf{x} = (x_1, \ldots, x_n)$ we have that

$$D_{\hat{\mathbf{u}}}f(\mathbf{x}) = \frac{\partial f}{\partial x_1}u_1 + \frac{\partial f}{\partial x_2}u_2 + \cdots + \frac{\partial f}{\partial x_n}u_n.$$

The Gradient

Suppose $f(\mathbf{x})$ is differentiable at the point $\mathbf{p} = (p_1, \ldots, p_n)$. The **gradient of $f(\mathbf{x})$ at the point $\mathbf{p}$** is the vector $\boldsymbol{\nabla} f(\mathbf{p})$ defined as

$$\boldsymbol{\nabla} f(\mathbf{p}) = \frac{\partial f}{\partial x_1}\hat{\mathbf{x_1}} + \cdots + \frac{\partial f}{\partial x_n}\hat{\mathbf{x_n}}.$$

In particular, the gradient vector points in the direction of greatest positive change in the function value.

Theorem: Gradient Adjunction Formula

If $f(\mathbf{x})$ is differentiable at a point $\mathbf{p} = (p_1, \ldots, p_n)$ and $\hat{\mathbf{u}} = (u_1, \ldots, u_n)$ is any unit vector, then we have that

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

where θ is the angle between the vectors $\boldsymbol{\nabla} f(\mathbf{p})$ and $\hat{\mathbf{u}}$.

Question 1. For each of the following, find the directional derivative **using only the limit definition.**

(a) $f(x, y) = 5 - 2x^2 - \frac{1}{2}y^2$ at the point $(3, 4)$ in the direction of the unit vector $(\sqrt{2}/2, \sqrt{2}/2)$.

(b) $h(x, y) = e^x \sin(y)$ at the point $(1, \pi/2)$ in the direction of the unit vector $\hat{\mathbf{u}} = -\hat{\mathbf{x}}$.

(c) $f(x, y) = y^2 - x^3 - x^2$ at the point $(0, 0)$ in the direction of the unit vector $(\sqrt{2}/2, -\sqrt{2}/2)$.

(d) $f(x, y) = y^2 - x^3 - x^2$ at the point $(-1, 0)$ in the direction of the unit vector $(0, 1)$.

Question 2. Let $h(x, y) = x^2 + y^2 - 1$.

(a) Sketch a contour plot of $h(x, y)$. Make sure to include contours corresponding to at least $h = -1$, $h = 0$, $h = 1$, $h = 2$, and $h = 3$.

(b) Find the graident of $h(x, y)$.

(c) Find the gradient of $h(x, y)$ at the each points $(\pm 1, 0)$, $(0, \pm 1)$. Plot these vectors on your contour plot.

(d) What do you notice about the magnitude of $\nabla h(x, y)$ at these points? The direction in which the gradient vectors point?

Question 3. For each of the following, compute the gradient.

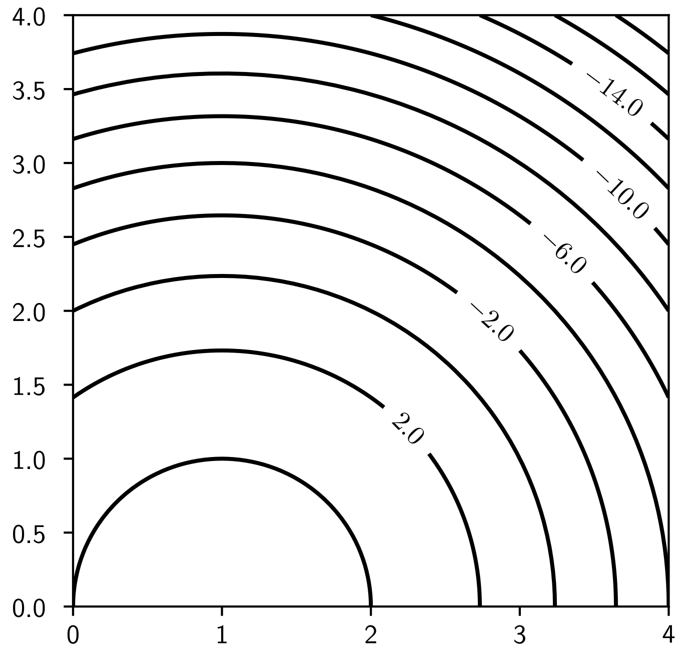
(a) $f(x, y) = x^2 + y^2$

(b) $g(x, y, z) = \sqrt{x^2 + y^2 + z^2}$

(c) $h(x, y) = \sin^2(\pi x) + \sin^2(\pi y)$.

(d) $u(x, y) = x^2 + 6xy - y^2$

Question 4. **(Application:)** In this problem, we will explore how we can use the gradient to maximize a function. For this example, let $f(x, y) = 4 - x^2 - y^2 + 2x$.



A contour plot of $f(x, y) = 4 - x^2 - y^2 + 2x$

(a) Find the gradient of $f(x, y)$.

(b) The first step in the algorithm is to pick an arbitrary point $\mathbf{p}_0 = (x_0, y_0)$ in the domain of our function. Let's pick the point $\mathbf{p}_0 = (2, 1)$ as our initial guess. Compute $\nabla f(\mathbf{p}_0)$. Plot the gradient vector at $\mathbf{p}_0$ on the contour plot.

(c) We will now "update" our guess by moving a small distance h along the gradient vector and then see where we land. For this example, we'll work with a fairly large step size, say $h = 0.25$. Let $\mathbf{p}_1 = \mathbf{p}_0 + h\nabla f(\mathbf{p}_0)$ be the next point. Compute $\mathbf{p}_1$ and $\nabla f(\mathbf{p}_1)$. Then plot $\nabla f(\mathbf{p}_1)$ on the contour plot.

(d) We can continue this process. The general process we follow is that $\mathbf{p}_{n+1} = \mathbf{p} + h\nabla f(\mathbf{p}_n)$. Using this method, compute $\mathbf{p}_n$ and $\nabla f(\mathbf{p}_n)$ for $n = 3, 4, 5$, and plot each of the gradient vectors on the contour plot.

(e) This process will either continue until we reach a point where the gradient vector is zero, or will run off to infinity. Why would the process stop if we had $\nabla f(\mathbf{p}) = \mathbf{0}$ for some $\mathbf{p}$?

10.7 Optimization

In this section, we will examine critical points of multivariable functions and how to find them.

Definition: Critical Points of Multivariable Functions

Let $f(x, y)$ be a differentiable function. A point $\mathbf{p} = (x_0, y_0)$ is a **critical point** of $f(\mathbf{x})$ if $\nabla f(\mathbf{p}) = \mathbf{0}$. Let $\mathbf{p} = (x_0, y_0)$ be a critical point of $f(x, y)$.

- If $f(\mathbf{p}) \geq f(\mathbf{x})$ for all points $\mathbf{x}$ *near* $\mathbf{p}$, then $\mathbf{p}$ is a **local maximum** of f.
- If $f(\mathbf{p}) \leq f(\mathbf{x})$ for all points $\mathbf{x}$ *near* $\mathbf{p}$, then $\mathbf{p}$ is a **local minumum** of f.
- The critical point $\mathbf{p}$ is a **saddle point** of $f(x, y)$ if it is neither a local maximum nor a local minimum.

Theorem: The Multivariable Second Derivative Test

Suppose we have a function $f(x, y)$ that is differentiable everywhere. Let $\mathbf{p} = (x_0, y_0)$ be a critical point of $f(x, y)$. Then, let $D = f_{xx}(\mathbf{p})f_{yy}(\mathbf{p}) - f_{xy}(\mathbf{p})^2$. (This D is called the *discriminant* or *Hessian* of $f(x, y)$ at $\mathbf{p}$.)

Then:

- If $D > 0$ and $f_{xx}(\mathbf{p}) < 0$ or $f_{yy}(\mathbf{p}) < 0$, then f has a local maximum at $\mathbf{p}$
- If $D > 0$ and $f_{xx}(\mathbf{p}) > 0$ or $f_{yy}(\mathbf{p}) > 0$, then f has a local minumum at $\mathbf{p}$
- If $D < 0$ then f has a saddle point at $\mathbf{p}$
- If $D = 0$ then this test yields no information about what happens at (x_0, y_0).

Theorem: The Extreme Value Theorem

If $f(x, y)$ is continuous on a closed, bounded region $\Omega \subset \mathbb{R}^2$, then $f(x, y)$ attains its global maximum and minimum inside Ω.

Question 1. Find and classify all critical points of $f(x, y) = x^2y^2 - 6x^2y - 4xy^2 + 24xy$.

Question 2. Let a and b be any two real numbers, and let $f(x, y) = xy - ax - by$.

(a) How many local minumum points must $f(x, y)$ have?

(b) How many local maximum points must $f(x, y)$ have?

(c) How many saddle points must $f(x, y)$ have?

Question 3. Find and classify all critical points of the funciton $f(x, y) = 22xye^{-x^2-y^2}$ using the second derivative test.

Question 4. A twice-differentiable function $h(x, y)$ is called **harmonic** on a closed, bounded region Ω if $f_{xx}(x, y) + f_{yy}(x, y) = 0$ for each point (x, y) in Ω. Harmonic functions satisfy the extrema principle— all of their minima and maxima lie on the boundary of Ω and the only critical points h can have on the interior of Ω are the saddle points.

(a) Use the Extreme Value Theorem to show that $h(x, y)$ must have maximum **and** minumum values on the boundary of Ω.

(b) Use the extrema principle to argue that, if $h(x, y)$ is constant on the boundary of Ω, then $h(x, y)$ must be constant on all of Ω.

Question 5. Airlines place stringent restrictions on what is allowed carry-on luggage. In addition to requirements on weight, airlines require that length of the bag plus its height cannot exceed 45in, so tha the bag is able to fit in an overhead bin.

Let $\ell, w,$ and h be the length, width, and height (in inches) of the bag. In this problem, we will find the dimensions of the bag of largest volume $V = f(\ell, w, h) = \ell wh$ that meets these requirements.

(a) Use the constraint $\ell + h = 45$ to rewrite the volume as a function $V = V(\ell, w)$.

(b) Explain why the domain over which V is defined is the triangular region R with vertices $(0, 0)$, $(45, 0)$ and $(0, 45)$ (in (ℓ, w)-coordinates).

(c) Find the critical points, if any, of V on the *interior* of R.

(d) Find the maximum value of V on the boundary of the region R, and then determine the dimensions of a bag with maximum value on the entire region R.