1 Introduction

The slice curves of a function graph contain information about how the function graph is changing in the direction of the slice curve. Each slice curve has an associated height function whose derivative determines the critical points for the slice curve. This derivative is called a directional derivative.

The gradient is a vector operator; applying this operator to a function generates a collection of gradient vectors called the gradient vector field. The gradient vectors of a function point in the direction of the largest directional derivative, that is the direction of greatest increase, making them useful in analyzing critical points.

2 The Directional Derivative

**Definition** The *directional derivative* for a function $f(x, y)$ in the direction $\theta$ is denoted $\nabla_\theta f(x, y)$ and is the derivative of the height function $z(t)$ in that direction.

Consider a point $P = (x_0, y_0)$ in the domain of $f(x, y)$. The line in the domain corresponding to the direction $\theta$ will have the parametric form

$$(x_0 + t \cos \theta, y_0 + t \sin \theta)$$
3.2: A geometric interpretation of the directional derivative.

and the height function \( z(t) \) associated with the slice curve along this line will have the form

\[
z(t) = f(x_0 + t \cos \theta, y_0 + t \sin \theta) .
\]

The directional derivative \( \nabla_\theta f(x_0, y_0) \) at the point \( P \) will be the derivative evaluated at \( t = 0 \) of the height function \( z(t) \)

\[
\nabla_\theta f(x_0, y_0) = z'(t)\big|_{t=0} = \frac{\partial}{\partial t} f(x_0 + t \cos \theta, y_0 + t \sin \theta)\big|_{t=0} .
\]

Note that the partial derivatives \( f_x \) and \( f_y \) are just the directional derivatives at \( \theta = 0 \) and \( \theta = \pi/2 \), respectively.

Also note that the directional derivative is a function of the direction variable \( \theta \); the variable \( t \) is only used as a parameter.

3 The Chain Rule

3.3: A geometric representation of the chain rule for two variables.
The chain rule for a function of a single variable states

If \( y(t) = f(x(t)) \), then \( y'(t) = f'(x(t))x'(t) \).

For functions of two variables it states

If \( z(t) = f(x(t), y(t)) \), then \( z'(t) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) \).

It is convenient to write the chain rule for two variables in the vector form

\[ z'(t) = (f_x(x(t), y(t)), f_y(x(t), y(t))) \cdot (x'(t), y'(t)) \cdot (x'(t), y'(t)) \cdot \]

We define \( \nabla f(x(t), y(t)) = (f_x(x(t), y(t)), f_y(x(t), y(t))) \) to be the gradient of \( f \) at \((x, y)\).

Now the chain rule for two variables reads \( z'(t) = \nabla f(x(t), y(t)) \cdot (x'(t), y'(t)) \)

Note that for a slice curve \((x(t), y_0)\) the vector form of the chain rule gives

\[ z'(t) = f_x(x(t), y_0)x'(t) + 0 \]

which simplifies to the usual chain rule for one variable.

4 The Gradient Operator

Definition The gradient operator is defined as

\[ \nabla f(x(t), y(t)) = (f_x(x(t), y(t)), f_y(x(t), y(t))) \, . \]

It is termed a "vector operator" because it is a mapping between two vectors whose components are functions.

Definition The gradient vector at a point \( P = (x_0, y_0) \) is defined as

\[ \nabla f(x_0, y_0) = (f_x(x_0, y_0), f_y(x_0, y_0)) \, . \]
At any given point $P$, the gradient vector points in the direction of the largest directional derivative and has magnitude equal to the value of the largest directional derivative.

More generally, the directional derivative in the direction $\theta$ at any given point can be expressed in terms of the gradient as

$$\nabla_{\theta} f = \nabla f \cdot (\cos \theta, \sin \theta).$$

Note that the gradient of a function of two variables is a vector in the domain of the function.

**Definition** There is a related vector called the normal vector in the graph of the function given by $(-f_x, -f_y, 1)$, perpendicular to $(1, 0, f_x)$ and $(0, 1, f_y)$.

Also note that the normal vector will be perpendicular to the tangent vector $(x'(t), y'(t), z'(t))$ since

$$
(-f_x(x(t), y(t)), -f_y(x(t), y(t)), 1) \cdot (x'(t), y'(t), z'(t))
= -f(x(t), y(t))x'(t) - f_y(x(t), y(t))y'(t) + z'(t)
= 0 \text{ by the chain rule.}
$$

5 The Gradient Vector Field

![Gradient Vector Field](image)

3.5: The gradient vector field and its relation to critical points, integral curves and level sets.

**Definition** The gradient vector field for a differentiable function $f(x, y)$ is the collection of gradient vectors for every point in the domain of $f$.

The function $f$ will have critical points whenever the gradient vector field is zero. In other words, the gradient vector at a critical point is the zero vector.

The gradient vector field is useful for determining local minima and maxima since the gradient vectors near these points will all point toward a maxima and away from a minima. So if one were to put a ball on a surface, the gradient vectors would lie in the opposite direction from the direction in which the ball would roll. A saddle point is more complicated, since some gradient vectors will point toward and others away from the saddle. If the same ball were placed exactly at a saddle point, it would not roll in any direction.
6 Introduction to LaGrange Multipliers

3.6: Finding the maximum height of a hilly highway.

**Definition** A special case of the chain rule occurs when \( z(t) = f(x(t), y(t)) = c \) is constant. In that case, \((x(t), y(t), c)\) is called a *level curve* of the function and \((x(t), y(t))\) is called a *contour* of the surface.

We then have

\[
0 = (c)' = z'(t) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) = \nabla f(x(t), y(t)) \cdot (x'(t), y'(t)).
\]

It follows that the gradient of a function of two variables at a point \((x(t), y(t))\) is perpendicular to the tangent vector to a level curve through the point. If the gradient at a point is not the zero vector, then all tangent lines to level curves through the point will be parallel. This follows since the intersection of the horizontal plane \((x, y, c)\) and the non-horizontal normal plane will be a single line and all tangent lines to level curves at the point must lie in this intersection.

The highest point on the Massachusetts Turnpike is indicated by a sign somewhere in the hilly region of Western Massachusetts. This point will occur when the gradient vector of the turnpike’s path is either the zero vector or when it is perpendicular to the turnpike’s tangent vector. In the first case, the turnpike’s maximum height occurs at a critical point of the surface of Western Massachusetts. In the second case, the turnpike’s maximum height occurs when it is tangent to a contour line.