Directional Derivatives and Gradients

Directional Derivatives

Given a differentiable function $f(x, y)$ and unit vector $\vec{u} = \langle a, b \rangle$, the *directional derivative of f in the direction of u* is

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

For a function of three variables $f(x, y, z)$ and unit vector $\vec{u} = \langle a, b, c \rangle$, we have

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

(Similar formulas for higher dimensions.)

Gradients

Given a function $f(x, y)$, the *gradient* of f is

$$\begin{aligned} \nabla f(x, y) &= \langle f_x(x, y),\ f_y(x, y) \rangle \\ &= \frac{\partial f}{\partial x}\hat{\mathbf{\imath}} + \frac{\partial f}{\partial y}\hat{\mathbf{\jmath}} \end{aligned}$$

For a function of three variables $f(x, y, z)$, we have

$$\begin{aligned} \nabla f(x, y, z) &= \langle f_x(x, y, z),\ f_y(x, y, z),\ f_z(x, y, z) \rangle \\ &= \frac{\partial f}{\partial x}\hat{\mathbf{\imath}} + \frac{\partial f}{\partial y}\hat{\mathbf{\jmath}} + \frac{\partial f}{\partial z}\hat{\mathbf{k}} \end{aligned}$$

(Similar formulas for higher dimensions.)

The gradient and directional derivative related to each other as follows:

$$D_{\mathbf{u}} f = \nabla f \cdot \vec{u}$$

This leads to the following theorem:

Theorem 1. *Given a function f of two or three variables and point $\mathbf{x}$ (in two or three dimensions), the maximum value of the directional derivative at that point, $D_{\mathbf{u}} f(\mathbf{x})$, is $|\nabla f(\mathbf{x})|$ and it occurs when $\vec{u}$ has the same direction as the gradient vector $\nabla f(\mathbf{x})$.*

Another property of the gradient is that: Given function f and a point (in two or three dimensions), the gradient vector at that point is perpendicular to the level curve/surface of f which passes through that point.

This leads to the following:

Tangent Planes to Level Surfaces

Given a level surface S of a function $F(x, y, z)$ of three variables (so S is of the form $F(x, y, z) = k$), we can use the property above to find the equation of the plane tangent to S at a point (x_0, y_0, z_0). We obtain the normal vector from the gradient:

$$\vec{n} = \nabla F(x_0, y_0, z_0)$$

And so the equation of the tangent plane is

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0$$

Normal Line

The normal line to the tangent plane is the line with direction vector $\nabla F(x_0, y_0, z_0)$ that passes through (x_0, y_0, z_0). Its vector equation is

$$\vec{r}(t) = \langle\, F_x(x_0, y_0, z_0) \cdot t + x_0,\ F_y(x_0, y_0, z_0) \cdot t + y_0,\ F_z(x_0, y_0, z_0) \cdot t + z_0 \,\rangle$$

and its symmetric equations are

$$\frac{x - x_0}{F_x(x_0, y_0, z_0)} = \frac{y - y_0}{F_y(x_0, y_0, z_0)} = \frac{z - z_0}{F_z(x_0, y_0, z_0)}$$

Maximum and Minimum Values for Functions of Two Variables

Finding Local Maxima and Minima of Functions of Two Variables

To find local maxima/minima of a function $f(x, y)$ of two variables, we do the following:

1. Take both partial derivatives, f_x and f_y, and set them equal to zero. This gives us a system of equations.

2. Solve the system of equations for x and y. This will give our *critical point(s)*, i.e., points (a, b) such that $f_x(a, b) = 0$ and $f_y(a, b) = 0$. (The first derivative test tells us that these are *possible* local maxima or minima of f.)

3. Use the Second Derivative Test for Functions of Two Variables to test these points:
   Let $f(x, y)$ be a function, (a, b) a point such that $f_x(a, b) = 0$ **and** $f_y(a, b) = 0$, and let

$$\begin{aligned} D(x, y) &= \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} \\ &= f_{xx} f_{yy} - f_{yx} f_{xy} \\ &= f_{xx} f_{yy} - (f_{yx})^2 \\ &= f_{xx} f_{yy} - (f_{xy})^2 \end{aligned}$$

   Then:

   i If $D(a, b) > 0$ and $f_{xx}(a, b) > 0$, then $f(x, y)$ has a local minimum at (a, b).

   ii If $D(a, b) > 0$ and $f_{xx}(a, b) < 0$, then $f(x, y)$ has a local maximum at (a, b).

   iii If $D(a, b) < 0$, then $f(x, y)$ has neither a relative minimum or minimum at (a, b). In this case, $f(x, y)$ has a *saddle point* at (a, b).

   iv If $D(a, b) = 0$, then this test is inconclusive.

Absolute Maximum and Minimum of Functions of Two Variables

To find the absolute maximum and minimum of a continuous function on a closed, bounded set B:

1. Find the value of f at any critical points of f in B.
2. Find the absolute maximum and minimum of f along the boundary of B. This typically involves finding an equation for the boundary, substituting that equation into f for one of the variables, then finding the absolute max/min using techniques from calculus I. If the boundary of B is defined by multiple functions, then find the absolute max/min for each piece of the boundary.
3. The largest value from 1. and 2. is the absolute maximum, the smallest is the absolute minimum.

Lagrange Multipliers

The method of Lagrange multipliers come from the fact that the gradient vector at a point is perpendicular to the level curve/surface of f which passes through that point. So if any two curves/surfaces pass through a point, then the gradient vectors of the associated functions will be parallel at that point.

Suppose we want to optimize the function $f(x,y,z)$ subject to a constraint $g(x,y,z)=k$. We first solve the system of equations

$$\begin{aligned} \nabla f(x,y,z) &= \lambda \nabla g(x,y,z) \\ g(x,y,z) &= k \end{aligned}$$

or, equivalently,

$$\begin{aligned} f_x &= \lambda g_x \\ f_y &= \lambda g_y \\ f_z &= \lambda g_z \\ g &= k \end{aligned}$$

for x, y, and z (we don't necessarily need to know the value of λ). Then we evaluate f at all of the solution points (x,y,z). The largest resulting value is the maximum of f, subject to the constraint, and the smallest is the minimum of f, subject to the constraint.

The method can also be used to optimize functions of two independent variables subject to a constraint:

$$\begin{aligned} \nabla f(x,y) &= \lambda \nabla g(x,y) \\ g(x,y) &= k \end{aligned}$$

and to optimize functions subject to multiple constraints:

$$\begin{aligned} \nabla f(x,y,z) &= \lambda \nabla g(x,y,z) + \mu \nabla h(x,y,z) \\ g(x,y,z) &= k \\ h(x,y,z) &= l \end{aligned}$$

or

$$\begin{aligned} \nabla f(x,y) &= \lambda \nabla g(x,y) + \mu \nabla h(x,y) \\ g(x,y) &= k \\ h(x,y) &= l \end{aligned}$$