14.5 and 14.6: Chain Rule, Directional Derivatives and the Gradient

In these sections we pick up some tools that are generally important in the understanding of multivariable calculus and specifically will be important in chapter 16.

14.5: The Chain Rule

1. Recall: If $y = f(x)$, then $\frac{dy}{dx} = f'(x)$ is the slope of the tangent line to $f(x)$ which is the same as instantaneous rate of change of y with respect to x.
   If $z = f(x, y)$, then $\frac{\partial z}{\partial x} = f_x(x, y)$ is the 'slope' in the direction parallel to the x-axis which is the same as the instantaneous rate of change of z with respect to x.
   In this section we consider these rates in situations when input variables for a multi-variable function are also variables of other parameters. These considerations will be important as we consider parametric curves and parametric surfaces along 2 and 3 variable functions in chapter 16.

2. CASE 1: Assume $z = f(x, y)$ is a surface and $x = g(t)$ and $y = h(t)$.

   We can think of $x = g(t)$, $y = h(t)$ as motion along some parameterized curve, C. If we want to know the change in z with respect to t as we move along this curve, we can use the generalized chain rule:
   $$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

3. CASE 2: Assume $z = f(x, y)$ is a surface, $x = g(s, t)$ and $y = h(s, t)$.

   Here x and y are defined in terms of two parameters. If we wish to measure the rate of change of z with respect to one of the parameters s or t, we get the very similar rules:
   $$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \quad \text{and} \quad \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}.$$

4. These results generalize to more variables. To help explain the generalization, a quick discussion of variable names. In CASE 2, we had

   - z is the **dependent variable**
   - x and y are **intermediate variables**
   - s and t are **independent variables**

   We can summarize these relationships in a tree diagram.

   When we want to find $\dfrac{\partial z}{\partial t}$ for example, we go through the tree to get to every possible t (multiplying the derivatives along a branch and adding separate branches together). This is an easy way to explain the more general form of the chain rule which follows the same pattern.

14.6: Directional Derivatives and Gradients

1. If this section we are given a multivariable function, such as the surface $z = f(x, y)$ and we wish to measure the rate of change in various directions (not just parallel to the x and y axis). If we are given a direction by a unit vector $\mathbf{u} = \langle a, b\rangle$ (meaning $a^2+b^2 = 1$), then we define the **directional derivative of $f(x, y)$ in the direction of u** by

$$D_{\mathbf{u}}f(x, y) = \lim_{h\to 0} \frac{f(x + ha, y + hb) - f(x, y)}{h} \quad \text{or} \quad D_{\mathbf{u}}f(\mathbf{x}) = \lim_{h\to 0} \frac{f(\mathbf{x} + h\mathbf{u}) - f(\mathbf{x})}{h}$$

where $\mathbf{x} = \langle x, y\rangle$.

This definition is the same for 3, 4, 5, ... variable functions.

Unit Direction Vector Note: Often a unit direction vector $\mathbf{u} = \langle a, b\rangle$ is described in terms of an angle, θ. If θ is measured counterclockwise on the xy-plane from the positive x-axis (in the same way as we do in polar, cylindrical and spherical coordinates), then we know from some basic trig that the unit direction vector is given by $\mathbf{u} = \langle\cos(\theta), \sin(\theta)\rangle$. In any case, the direction vector will satisfy $\mathbf{u} = \langle a, b\rangle = \langle\cos(\theta), \sin(\theta)\rangle$ with $a^2 + b^2 = 1$.

Partial Derivatives Note: If $\mathbf{u} = \langle 1, 0\rangle$, then the direction vector is parallel to the x-axis an we get $D_{\langle 1,0\rangle}f(x, y) = f_x(x, y)$. Similarly $D_{\langle 0,1\rangle}f(x, y) = f_y(x, y)$.

2. The **gradient vector** is the vector containing each partial derivative as components. That is,

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle = \left\langle \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y} \right\rangle \text{ and } \nabla f(x, y, z) = \langle f_x(x, y), f_y(x, y), f_z(x, y)\rangle$$

The symbol ∇ is often called the **gradient operator**. For a multivariable function, it means take all partial derivatives and express them in a vector as above.

3. In the book, and in class, we quickly derived the connection between the directional derivatives and gradient. We found:

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

This results extend naturally to three dimensions.

4. **Significance of Gradient**: If f is a function of two or three variables, then
   (a) $|\nabla f(\mathbf{x})|$ = the maximum values of $D_{\mathbf{u}}f(\mathbf{x})$ over all possible directions $\mathbf{u}$.
   (b) And this maximum occurs when $\mathbf{u}$ is in the same direction as $\nabla f(\mathbf{x})$.

5. **Notes on the Gradient for a function of two variables**: Let $z = f(x, y)$ be a surface. Consider the contour map we get by plotting the level curves $f(x, y) = k$ in the xy-plane for various values of k. At any point, (x_0, y_0), the gradient vector $\nabla f(x_0, y_0)$ gives the direction of steepest ascent along the surface. So the gradient vector is always perpendicular to any level curve $f(x, y) = k$.

6. **Notes on the Gradient for a function of three variables**: Let $w = f(x, y, z)$ be a function. Consider the level *surfaces* you get by plotting $f(x, y, z) = k$ for various values of k. At any point (x_0, y_0, z_0), the gradient vector $\nabla f(x_0, y_0, z_0)$ gives the direction of greatest increase. So the gradient is always orthogonal to the level surface $f(x, y, z) = k$. Another way to say this is that if $\mathbf{r}(t)$ is any curve along the level surface $f(x, y, z) = k$ that goes through (x_0, y_0, z_0) at the parameter value t_0, then by differentiating $f(x(t), y(t), z(t)) = k$ we get $\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} = 0$, so $\nabla f(x_0, y_0, z_0)\cdot \mathbf{r}'(t_0) = 0$. The gradient vector is orthogonal to any curve traveling along a level surface.