1 Partial derivatives and Differentiability

Definition 1.1 *Given a function f of two variables given by $z = f(x, y)$,* **the first partial derivative of f with respect to** *x at the point (a, b), denoted by either $f_x(a, b)$ or $\frac{\partial f}{\partial x}(a, b)$, is defined as*

$$f_x(a,b) = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h} = \lim_{x\to a} \frac{f(x,b) - f(a,b)}{x-a},$$

and **the first partial derivative of f with respect to** *y at the point (a, b), denoted $f_y(a, b)$ or $\frac{\partial f}{\partial y}(a, b)$, is defined as*

$$f_x(a,b) = \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h} = \lim_{y\to b} \frac{f(a,y) - f(a,b)}{y-b}.$$

Note 1.2 *In practice, these partial derivatives (for arbitrary points (x, y)) are usually computed with the same "short-cuts" as in first semester calculus by treating the variable you're not differentiating with respect to as a constant. HOWEVER, just because you can't compute a partial derivative with respect to these shortcuts doesn't mean the partial derivative doesn't exist (you can see this, for example, in the solution to Exercise 58 in section 13.4); in this case, we appeal to the limit definition given above.*

Notation 1.3 *When considering higher order partial derivatives, we abide by the following notational conventions: if we differentiate with respect to x first and then differentiate with respect to y, we denote this by $\frac{\partial^2 f}{\partial y \partial x} := \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$ or $f_{xy} := (f_x)_y$; notation is similar for the other partial derivatives. Notice that in these two notations for the same thing the order of the variables are in opposite orders - take note of this and BE CAREFUL.*

Theorem 1.4 (Clairut) *Assume that f is defined on an open set D of $\mathbb{R}^2$ and that f_{xy} and f_{yx} are continuous throughout D. Then, $f_{yx} = f_{xy}$ on D.*

The book's definition of differentiability is not typically very useful or computationally practical in practice, so we're usually going to use the theorem below to determine differentiability (you may use it as a definition).

Theorem 1.5 *Suppose that a function f has partial derivatives f_x and f_y defined on an open set containing (a, b) and that f_x and f_y are both continuous at the point (a, b). Then, f is* **differentiable** *at the point (a, b).*

WARNING 1.6 *Note that this theorem says the existence of partial derivatives at (a, b) is NOT enough to conclude that f is differentiable at the point (a, b). These partial derivatives MUST be continuous as well at (a, b). This is illustrated in 13.4 Exercise 58, for instance. Moreover, partial derivatives at (a, b) may exist despite the fact f isn't continuous at (a, b).*

Theorem 1.7 *If a function f is differentiable at the point (a, b), then f is continuous at (a, b). In particular, if f is not continuous at (a, b), it cannot be differentiable at (a, b).*

2 The Chain Rule

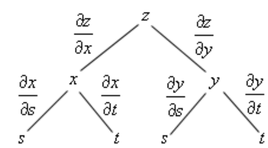
First Scenario: We consider the case where z is a function of two variables x and y, which are both functions of one variable t. Then,

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

Second Scenario: We consider the case where f is a function of a pair of variables x and y, where both x and y are both functions of another pair of s and t. Then,

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \quad \text{and} \quad \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}.$$

In general, variable dependence could look different from this (it could be more complicated, in fact!). The best way to determine how the Chain Rule works in computing a given partial derivative is by setting up a **dependence tree** and looking at all possible paths from the top to your desired variable. For example, the following is the dependence tree for $z = f(x, y)$ where x and y are functions of s and t (as in the second scenario):



Notice that from this tree the formulas on the previous page for $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$ follow quite naturally when considering all possible paths from z to s and z to t, respectively, on the tree above, and multiplying the legs of the path.

Procedure 2.1 *Consider a function $F : \mathbb{R}^2 \to \mathbb{R}$ that is differentiable on its domain, and suppose that the equation $F(x, y) = 0$ defines y as a differentiable function of x (in general, we'll assume this; a result from advanced calculus called the Implicit Function Theorem determines when this is true). Then, $\frac{dy}{dx} = -\frac{F_x}{F_y}$.*

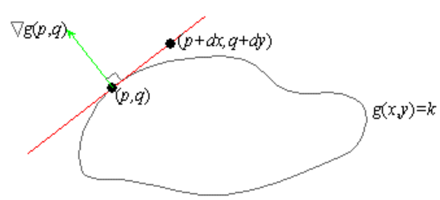
3 Directional Derivatives and the Gradient

Definition 3.1 *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function at the point (a, b) and let $\mathbf{u} = \langle u_1, u_2\rangle$ be a unit vector. Then, the* **directional derivative** *of f at the point (a, b) in the direction of $\mathbf{u}$ is $D_{\mathbf{u}}f(a, b) = \langle f_x(a, b), f_y(a, b)\rangle \cdot \mathbf{u}$, and the vector $\nabla f(a, b) := \langle f_x(a, b), f_y(a, b)\rangle$ is called the* **gradient** *of f at (a, b). In particular, if f is differentiable on its domain, ∇f is a function defining each point (a, b) to a vector $\nabla f(a, b)$; we call such functions* **vector fields**.

The gradient is actually has enormous significance besides its use in computing directional derivatives. We list some results below.

Theorem 3.2 *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function at the point (a, b).*
1. *f has its maximum rate of increase from the point (a, b) in the direction of its gradient $\nabla f(a, b)$. The rate of that increase in this direction is $|\nabla f(a, b)|$.*
2. *f has its maximum rate of decrease from the point of (a, b) in the direction opposite from its gradient, i.e. in the direction of $-\nabla f(a, b)$. The rate of that decrease in this direction is $-|\nabla f(a, b)|$.*
3. *The directional derivative is 0 in any direction orthogonal to $\nabla f(a, b)$.*
4. *Provided $\nabla f(a, b) \neq \mathbf{0}$, if $f(a, b) = c$, then $\nabla f(a, b)$ is perpendicular to the tangent vector at the point (a, b) on the level curve $f(x, y) = c$ in the xy-plane.*

Here's a picture illustrating what 4. means visually:



As we'll learn in 13.9, the result 4. stated in this theorem is extremely important: it's the reason why Lagrange multipliers work!

Definitions and results for functions of 3 variables are more or less the same, so we won't state them here.