Notes on Differential Calculus of Surfaces (Chapter 14)

by Alec Johnson

fall, 2009

This document discusses smooth real-valued functions of multiple variables. We use $f(x, y)$ to denote a generic function from $\mathbb{R}^2$ to $\mathbb{R}$, and we use $f(x, y, z)$ to denote a generic function from $\mathbb{R}^3$ to $\mathbb{R}$.

1 Anonymous functions.

We use *anonymous function* notation to specify a function without having to give it a name. For example, $x \mapsto x^3$ is the cube function, and $(x, y) \mapsto x^2y + b$ is the anonymous way to refer to the function $f(x, y) = x^2y + b$. To evaluate an anonymous function, we use a vertical bar: e.g., $(x \mapsto x^3)\big|_2 = 8$.

2 Slices

A *slice* of a multivariable function $f(x, y, z)$ is a function obtained by freezing some of the variables. If we freeze all the variables except for one, we get a function of one variable, called a *one-dimensional slice*. For example, the function $x \mapsto f(x, y_0)$ is the slice of f along the line $y = y_0$, and the function $x \mapsto f(x, y_0, z_0)$ is the slice of f along the line $y = y_0, z = z_0$.

3 Partial derivatives

By considering one-dimensional slices we can use the tools of one-variable calculus to do multivariable calculus. We refer to old-fashioned one-variable derivatives as *ordinary derivatives*. The *partial derivative of f with respect to x*, written $\frac{\partial f}{\partial x}$, is defined to be the ordinary derivative of f with respect to x holding the other variables constant. It is the rate of change of f as you move parallel to the x axis. It is denoted $\frac{\partial f}{\partial x}$, $D_x f$, or f_x. So

$$\frac{\partial f}{\partial x}(x_0, y_0, z_0) := \frac{d\big(x \mapsto f(x, y_0, z_0)\big)}{dx}\bigg|_{x=x_0}.$$

(Recall that we write $A := B$ (or sometimes $B =: A$) to mean that A is defined by B.)

4 Quantity notation

Physicists often use "quantity notation" to specify multivariable functions. A common example is the ideal gas equation $PV = NkT$. We write $P(T, V, N)$ to refer to the quantity named P as a function of the quantities named T and V, and N.

When we are using quantity notation, we write something like $\left(\frac{\partial P}{\partial T}\right)_{V,N}$ to mean "the partial derivative of the quantity P as T changes and V and N are held constant", where it is implied that $P(T, V, N)$, i.e., that the quantity P is a function of the independent quantities T, V, and N.

Remark 4.1 (Notation for partial derivatives)**.** *The notation for partial derivatives can be problematic. When you take a partial derivative, you must be very clear what function and what argument you are talking about. $\frac{\partial f}{\partial x}$ means "the partial derivative of f with respect to the formal argument named x holding the other arguments constant". The problem can arise when you take the derivative of an* expression *for a quantity, where the function is not explicitly defined. Then you need to be clear "which arguments you are holding constant", i.e., which function you are talking about. Here is an example of ambiguity that arises when you are doing implicit partial differentiation: Does $\frac{\partial f(x,y,z(x,y))}{\partial x}$ mean*

$$\frac{\partial}{\partial x}\Big[(x, y) \mapsto f(x, y, z(x, y))\Big]\bigg|_{(x,y)} =: \frac{\partial}{\partial x}\Big[f(x, y, z(x, y))\Big]$$

or

$$\frac{\partial}{\partial x}\Big[(x, y, z) \mapsto f(x, y, z)\Big]\bigg|_{(x,y,z(x,y))} =: (f_x)(x, y, z(x, y))?$$

The problem is that our notation is ambiguous and uses the letter x to refer to two different things:

(1) the first argument of the function f and (2) the first coordinate of the point *where we are evaluating our partial derivative. One way to deal with this would be to rename the formal arguments of f as (u, v, w) and write something like $\frac{\partial f}{\partial u}\big|_{(x,y,z(x,y))}$. Another way would be be to use the notation D_n, the partial derivative with respect to the nth argument and write something like $D_1 f\big|_{(x,y,z(x,y))}$.*

Exercise. Let $f(x, y, z) = x^2y^3 + z$. Find $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}(1, 2, 3)$, and $\frac{\partial f}{\partial z}(1, 2, 3)$. Answers: $2xy^3$, 12, and 1.

5 Linear functions

The essential idea of calculus is that the graph of a smooth function locally looks flat. Another word for flat is linear. So the starting point for understanding the calculus of multivariable functions is to understand linear functions of multiple variables.

In this section we let $f(x, y)$ represent a linear function from $\mathbb{R}^2$ to $\mathbb{R}$. Its graph is a *plane*. Its partial derivatives are constants. What does the formula for such a linear function look like? I claim that it is given by the "point-slopes" formula

$$\boxed{f(x, y) = f(x_0, y_0) + A(x - x_0) + B(y - y_0)},$$

where the "slopes" are the partial derivatives:

$$A = \frac{\partial f}{\partial x}, \quad B = \frac{\partial f}{\partial y}.$$

To check this claim, simply observe that the right hand side agrees with f at the base point (x_0, y_0) and has the same partial derivatives. A little thought should convince you that if two functions f and g agree at a base point and have the same partial derivatives, then they must be identical. (Hint: what can you say about the function $f - g$?) If you want a derivation, here are two to pick from:

1. **Integrating along slices.** To obtain an algebraic expression for f, observe that that since the graph is a plane, a slice of f parallel to the x axis is a straight line, and *the slope of this line is independent of the slice.* Let A be the slope of slices along the x axis, and let B be the slope of slices along the y axis. Pick a base point (x_0, y_0). We want to find $f(x, y)$ in terms of A, B, and f_0. We do so by moving from (x_0, y_0) to (x, y) along slices. Moving parallel to the x-axis first and then parallel to the y-axis, we have: $f(x, y_0) = f(x_0, y_0) + A(x - x_0)$ and $f(x, y) = f(x, y_0) + B(y - x_0)$. Putting these together, we have:

$$f(x, y) = f_0 + A(x - x_0) + B(y - y_0)$$

2. **Starting with the equation of a plane.** Suppose that the graph of $z = f(x, y)$ is a plane. Let $\mathbf{r}_0 = (x_0, y_0, z_0)$ be a point on this plane, and let $\widetilde{\mathbf{n}} = (\widetilde{A}, \widetilde{B}, \widetilde{C})$ be a normal vector for this plane. Since the plane is the graph of a function, the normal vector must have a component in the direction of the z-axis, i.e., $\widetilde{\mathbf{n}} \cdot \mathbf{k} = \widetilde{C} \neq 0$. So we can rescale $\widetilde{\mathbf{n}}$ by $-1/\widetilde{C}$, giving a normal vector of the form $\mathbf{n} = (A, B, -1)$. Then the equation of the plane is:

$$\begin{aligned}
&\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0, \text{ i.e.,} \\
&z - z_0 = A(x - x_0) + B(y - y_0), \text{ i.e.,} \\
&f(x, y) = f_0 + A(x - x_0) + B(y - y_0).
\end{aligned}$$

If we denote the change in f by $\Delta f := f(x, y) - f(x_0, y_0)$, the change in x by $\Delta x := x - x_0$, and the change in y by $\Delta y := y - y_0$, then the formula for a linear function says:

$$\boxed{\Delta f = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y};$$

i.e., the change in f is the rate of change of f as you move in the x direction times the displacement in the x direction plus the rate of change of f as you move in the y direction times the displacement in the y direction.

If we choose the base point to be the origin $(0, 0)$, then our formula for a linear function says that we can write f as

$$\boxed{f(x, y) = f(0, 0) + Ax + By}.$$

Exercise. Find the linear function f that satisfies $f(0, 0) = 3$, $f(1, 0) = 8$, and $f(0, 2) = -1$. Answer: $f(x, y) = 3 + 5x - 2y$.

6 Linear approximation

In this section we let $f(x, y)$ be a smooth function near the point (x_0, y_0). This means that you can approximate f by a linear function. More formally:

Proposition 6.1 (Linear approximation). *Let $f(x, y)$ be differentiable at (x_0, y_0). Then $f(x, y) \approxeq L(x, y)$, where the linear approximation $L(x, y)$ is given by*

$$\begin{aligned} L(x, y) := f(x_0, y_0) + (x - x_0)\frac{\partial f}{\partial x}\Big|_{(x_0,y_0)} \\ + (y - y_0)\frac{\partial f}{\partial y}\Big|_{(x_0,y_0)}. \end{aligned}$$

This should seem intuitively clear to you. It says that when you move from (x_0, y_0) to (x, y), the change in f is approximately the rate at which f changes as you move along the x axis times the change in x plus the rate at which f changes as you move along the y axis times the change in y.

The language of **differentials** is designed to make this clear. A small change in x and y is represented by the "differentials" $dx := x - x_0$ and $dy := y - y_0$ and the resulting small change in f is approximated by the "differential" df:

Definition 6.2 (Differential).

$$f(x, y) - f(x_0, y_0) \approx df, \quad \text{where}$$
$$\boxed{df := dx\frac{\partial f}{\partial x}\Big|_{(x_0,y_0)} + dy\frac{\partial f}{\partial y}\Big|_{(x_0,y_0)}}$$

Note that the differential df is *not* the change Δf. The differential df is technically defined to be the change in the *linear approximation* of f. That is, $L(x, y) = f(x_0, y_0) + df$.

In my opinion, the best way to make sure that you understand a calculus relationship is to see what it tells you in the linear case. That is, if you want to understand why a rule is true, try it out on a linear function. So take a linear function $f(x, y) = Ax + By + C$ and plug it into the equations above. These approximate equalities should become exact equalities.

The formal definition of differentiability basically says that the error of the linear approximation is small:

Definition 6.3 (Derivative). $f(\mathbf{r})$ is **differentiable** at $(\mathbf{r}_0)$ if there is a linear approximation $L(\mathbf{r}) = C + \mathbf{n} \cdot \mathbf{r} = C + Ax + By$. ($L$ is a linear approximation at $\mathbf{r}_0$ if as $\mathbf{r}$ goes to $\mathbf{r}_0$ the error $f(\mathbf{r}) - L(\mathbf{r})$ goes to zero even faster (i.e., $\lim_{\|\mathbf{r}-\mathbf{r}_0\|\to 0} \frac{f(\mathbf{r})-L(\mathbf{r})}{\|\mathbf{r}-\mathbf{r}_0\|} = 0$)).

Exercise. Let $f(x, y, z) = x^2y^3 + z$.

1. What is the linear approximation of the function f near the point $(1, 2, 3)$?
   Answer: $11 + 16(x - 1) + 12(y - 2) + (z - 3)$.
2. What is the *change* in f as $\mathbf{r}$ goes from $(1, 2, 3)$ to $(1.1, 2, 3.1)$? Answer: 1.78
3. What is the *differential* of the linear approximation of f near $(1, 2, 3)$ as $\mathbf{r}$ goes from $(1, 2, 3)$ to $(1.1, 2, 3.1)$? Answer: 1.7.

7 Chain Rule

To derive a chain rule for differentiating functions of multiple variables, we work with differentials.

Proposition 7.1 (One-variable differentials). *Let $u(t)$ be differentiable. Then*

$$du = \frac{du}{dt}dt.$$

For a small differential du we can substitute this in to the multivariable formula for differentials to get a differential form of the chain rule:

Proposition 7.2 (Chain rule for differentials). *Given the functions $f(u, v)$, $u(t)$, and $v(t)$,*

$$\begin{aligned} df &= \frac{\partial f}{\partial u}du \quad + \frac{\partial f}{\partial v}dv \\ &= \frac{\partial f}{\partial u}\frac{du}{dt}dt + \frac{\partial f}{\partial v}\frac{dv}{dt}dt. \end{aligned}$$

Dividing the chain rule for differentials by dt gives:

Proposition 7.3 (Chain rule (I)).

$$\boxed{\frac{df}{dt} = \frac{\partial f}{\partial u}\frac{du}{dt} + \frac{\partial f}{\partial v}\frac{dv}{dt}}.$$

Proposition 7.4 (Chain rule (II))**.** *Let $u(t,s)$ and $v(t,s)$. Holding s constant and applying chain rule (I) gives*

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial t}.$$

Again, to see why these propositions are true, check that they hold for linear functions.

Exercises.

1. Suppose that $u(1) = 3$, $v(1) = 2$, $\frac{\partial f}{\partial u}(3,2) = 7$, $\frac{\partial f}{\partial v}(3,2) = 2$, $\frac{du}{dt}(1) = 2$, and $\frac{dv}{dt}(1) = -1$, and Find $\frac{df}{dt}$. Answer: 12.

2. Let $f(u,v) = u^2v^3$ Let $u(t) = \cos(t)$ and let $v(t) = \sin(t)$. Find $\frac{d}{dt}f(u(t),v(t))$, (a) using the multivariable chain rule and (b) by substituting and differentiating. Answer: $-2\cos(t)\sin^4(t) + 3\cos^3(t)\sin^2(t)$.

8 Directional derivatives and the gradient

Definition 8.1 (Directional derivative)**.** The derivative of $f(x,y)$ in the direction $\hat{\mathbf{u}}$ at the point $\mathbf{r}_0 = (x_0, y_0)$ is just the rate of change of f as you travel through $\mathbf{r}_0$ along a line in the direction $\hat{\mathbf{u}}$, i.e., $\frac{df(\mathbf{r}(t))}{dt}\Big|_{t=0}$, where $\mathbf{r}(t) = \mathbf{r}_0 + t\hat{\mathbf{u}}$:

$$\begin{aligned} D_{\hat{\mathbf{u}}}f\Big|_{\mathbf{r}_0} &= \frac{df(\mathbf{r}_0 + t\hat{\mathbf{u}})}{dt} \\ &= \frac{dx}{dt}\frac{\partial f}{\partial x} + \frac{dy}{dt}\frac{\partial f}{\partial y} \\ &= u_1\frac{\partial f}{\partial x} + u_2\frac{\partial f}{\partial y} \\ &= \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \end{aligned}$$

(where derivatives of f with respect to position coordinates are evaluated at $\mathbf{r}_0$). So we can write:

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

where the vector $\nabla f := (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$ is called the **gradient** of f and θ is the angle between $\hat{\mathbf{u}}$ and ∇f. Thus *the directional derivative of f varies between $-|\nabla f|$ and $|\nabla f|$ and is maximized in the direction of ∇f.*

Exercises. Let $f(x,y) = x^2y^3$.

1. Find ∇f. Answer: $\langle 2xy^3, 3x^2y^2\rangle$.

2. Find $\nabla f\big|_{(1,2)}$. Answer: $\langle 16, 12\rangle$.

3. Find the derivative of the function $f(x,y) = x^2y^3$ near $(1,2)$ in the direction of the vector $(3,4)$. Answer: $96/5$. (Hint: what is the direction vector of $(3,4)$?)

4. Find the directions for which $D_{\hat{\mathbf{u}}}f$ is maximized and minimized. Answer: $\langle 4/5, 3/5\rangle$, $\langle -4/5, -3/5\rangle$.

5. Find the directions for which $D_{\hat{\mathbf{u}}}f$ is zero. Answer: $\pm\langle 3/5, -4/5\rangle$.

9 Implicit differentiation

Problem 9.1 (Implicit partial derivative)**.** *Suppose that near $\mathbf{r}_0 = (x_0, y_0, z_0)$ the function $f(x,y,z)$ is smooth and $\frac{\partial f}{\partial z} \neq 0$. Then the equation*

$$f(x,y,z) = 0 \tag{1}$$

implicitly defines a function $z(x,y)$ near $\mathbf{r}_0$. To find the partial derivatives of z, we can use the chain rule to differentiate $f(x,y,z(x,y))$, being very clear with notation. As a shortcut, take the differential of (1) *and get*

$$\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = 0.$$

If we hold y constant, then $dy = 0$. Then $\frac{dz}{dx} = \frac{\partial z}{\partial x}$, and so

$$\boxed{\frac{\partial z}{\partial x} = \frac{-\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}}}. \tag{2}$$

Exercise. Let $f(x,y,z) := xz^2 + yz^3$. The equation $f = 0$ implicitly defines z as a function of x and y near points on the solution set where $\frac{\partial f}{\partial z} \neq 0$. Verify that $\mathbf{r}_0 := (2,-1,2)$ is on the solution set. Find $\frac{\partial z}{\partial x}\Big|_{\mathbf{r}_0}$ (a) using the implicit differentiation formula (2), and (b) by assuming that z is a function of x and y, partially differentiating both sizes with respect to x (which will require you to use the chain rule!), solving for $\frac{\partial z}{\partial x}$, and evaluating at $\mathbf{r}_0$. Answer: $\frac{\partial z}{\partial x}\Big|_{\mathbf{r}_0} = -1$.