Math 18.02 – Notes on the Chain Rule and Implicit Differentiation – Fall 2011

The Chain Rule

The chain rule is an algebraic rule that describes how to calculate rates of change of functions built from other functions through composition. For example, in a first semester calculus course we learn that if $y = y(u)$ and $u = u(x)$, then we can calculate $\frac{dy}{dx}$ by the chain rule: $\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$. In a multivariable setting, we might have $z = z(x, y)$ and $x = x(t)$, $y = y(t)$. We then have $\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}$, the basic chain rule.

The chain rule gets more interesting when you apply it to situations where there are more input variables and output variables. For example, let us suppose we have a situation where there are two parameters, ϕ and θ, and that for any ϕ and θ we have equations giving $\left\{\begin{matrix} x = x(\phi,\theta) \\ y = y(\phi,\theta) \\ z = z(\phi,\theta) \end{matrix}\right\}$. Let us further suppose that for any choices of the variables x, y, and z we have two other variables, u and v, defined by equations $\left\{\begin{matrix} u = u(x,y,z) \\ v = v(x,y,z) \end{matrix}\right\}$.

In this case we can think of this functionally as

$$(\phi,\theta) \xrightarrow{G} (x,y,z) \xrightarrow{F} (u,v).$$

In this context, the general chain rule gives that

$$\frac{\partial u}{\partial \phi} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \phi} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial \phi} \qquad \frac{\partial u}{\partial \theta} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \theta} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial \theta}$$
$$\frac{\partial v}{\partial \phi} = \frac{\partial v}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial \phi} + \frac{\partial v}{\partial z}\frac{\partial z}{\partial \phi} \qquad \frac{\partial v}{\partial \theta} = \frac{\partial v}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial \theta} + \frac{\partial v}{\partial z}\frac{\partial z}{\partial \theta}$$

These can be organized into a statement about the Jacobian matrices of the two functions and of their composition. A Jacobian matrix may be thought of simply as an array of (partial) derivatives of the various output variables with respect to the various input variables, where the outputs are listed from top to bottom and the inputs are listed from left to right. If you know about matrix multiplication, we have

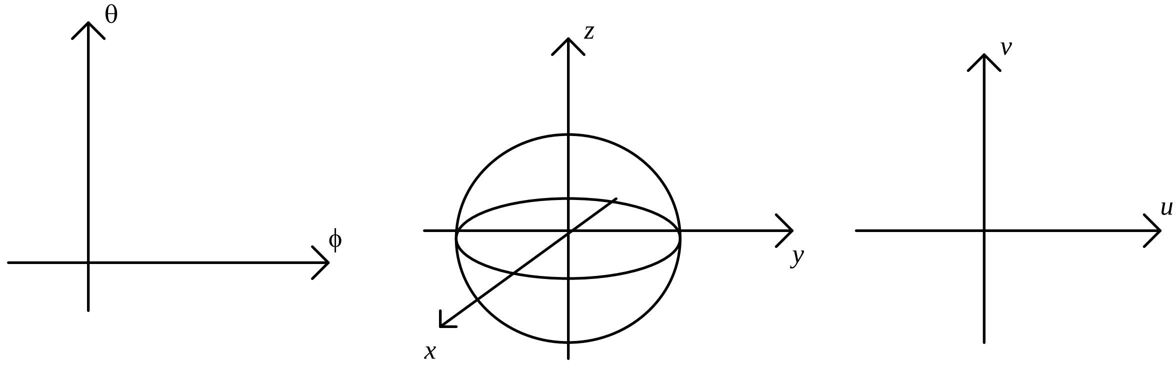
$$\begin{bmatrix} \frac{\partial u}{\partial \phi} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial \phi} & \frac{\partial v}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial \theta} \end{bmatrix}$$ or, more succinctly, $\mathbf{J}_{F\circ G} = \mathbf{J}_F\mathbf{J}_G$.

To picture what this is telling us, let's specifically look at the situation where ϕ and θ represent latitude and longitude with the minor change that latitude will be measured from the north pole as 0°, the equator as 90°, and the south pole as 180°. We can then describe a sphere of radius R by the parametric equations

$$\left\{\begin{matrix} x = R\cos\theta\sin\phi \\ y = R\sin\theta\sin\phi \\ z = R\cos\phi \end{matrix}\right\}.$$

[We'll derive these later when we look at spherical coordinates in detail.]

Let us further suppose that the variables u and v measure, for example, temperature and barometric pressure at any point (x, y, z) in $\mathbf{R}^3$ and, in particular, at points on this parametrized sphere in $\mathbf{R}^3$.



We might ask questions about how temperature would vary as we changed latitude or longitude, or how barometric pressure would vary as we changed latitude or longitude. These are the quantities in the Jacobian matrix $\mathbf{J}_{F\circ G} = \begin{bmatrix} \frac{\partial u}{\partial \phi} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial \phi} & \frac{\partial v}{\partial \theta} \end{bmatrix}$. The <u>rows</u> of the Jacobian matrix $\mathbf{J}_F = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{bmatrix}$ are just the gradient vectors (in $\mathbf{R}^3$) of the temperature and barometric pressure functions. (Note that these are functions defined on $\mathbf{R}^3$ and not just on the spherical surface.)

The two <u>columns</u> of the Jacobian matrix $\mathbf{J}_G = \begin{bmatrix} \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial \theta} \end{bmatrix}$ represent "velocity" vectors tangent to the longitudes (ϕ varying) and latitudes (θ varying). These two column vectors are tangent to curves lying in the sphere and are therefore tangent to the sphere. They are, essentially, the "south vector" and the "east vector" at any point of the sphere (except at the poles). You might further observe that their cross product will be normal to this spherical surface at any given point, a fact which will be useful later in this course when we look at surface integrals.

The two <u>columns</u> of the Jacobian matrix $\mathbf{J}_{F\circ G} = \begin{bmatrix} \frac{\partial u}{\partial \phi} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial \phi} & \frac{\partial v}{\partial \theta} \end{bmatrix}$ represent vectors in the (u, v) plane and indicate the directions of change if we slightly vary the latitude or the longitude.

Implicitly Defined Functions and Implicit Differentiation

Often it is the case that an equation (or several equations) relate some variables and we wish to consider one variable (or several) as depending on the rest. For example, given the equation of a circle

$$x^2 + y^2 = 16$$

we may wish to consider $y = y(x)$. If we solve <u>explicitly</u>, we get either $y = \sqrt{16 - x^2}$ or $y = -\sqrt{16 - x^2}$, whose graphs are the upper and lower semicircles. Though we could calculate the derivatives directly, there is an alternate approach. Think of x as a <u>parameter</u> and use it to parametrize either one of the semicircles as

$x \to (x, y(x))$, where the dependence of y on x is defined <u>implicitly</u> by the given curve (semi-circle). If we let $F(x, y) = x^2 + y^2$, then we can view the circle as just the $F = 16$ contour, or level set, of the function F. Composing these functions, we have

$$x \to (x, y(x)) \to F(x, y(x)) = \text{constant}$$

Applying the chain rule (and using F_x and F_y to denote the partial derivatives of F), we have

$$\frac{d}{dx} F(x, y(x)) = F_x 1 + F_y \frac{dy}{dx} = 0 .$$

Here we used the fact that $\frac{dx}{dx} = 1$ and that the composite function was constant. Solving for $\frac{dy}{dx}$, we get that $\frac{dy}{dx} = -\frac{F_x}{F_y}$, so as long as we avoid those places where $F_y = 0$ (where the two semicircles meet), we have a valid formula for calculating $\frac{dy}{dx}$. In the above example, this gives $\frac{dy}{dx} = -\frac{2x}{2y} = -\frac{x}{y}$. This may be used for either the upper or the lower semicircle.

This formulation, and the formula $\frac{dy}{dx} = -\frac{F_x}{F_y}$, are valid whenever we have a relation of the form $F(x, y) = \text{constant}$, where F is a differentiable function and where we can consider $y = y(x)$ as being implicitly defined by the equation. The only exception is at those points where $F_y = 0$, i.e. at points where the tangent line to the relation is vertical.

This same approach can be used for relations of the form $F(x, y, z) = \text{constant}$, where we may wish to consider one of the variables as being dependent on the others. For example, if we choose to think of $z = z(x, y)$, then it is useful to consider x and y as <u>parameters</u> and to formulate the situation as

$$(x, y) \to (x, y, z(x, y)) \to F(x, y, z(x, y)) = \text{constant}$$

Here we can think of the relation as a surface in $\mathbf{R}^3$, and what this is saying is that by choosing (x, y) we may find one point (or several points) on the graph. We can apply the chain rule to calculate the partial derivatives of the composition with respect to the parameters x and y. What makes this a bit tricky is the fact that x and y are playing dual roles as parameters and as coordinates in $\mathbf{R}^3$. Nonetheless, we have

$$\frac{\partial}{\partial x} F(x, y, z(x, y)) = F_x 1 + F_y 0 + F_z \frac{\partial z}{\partial x} = 0$$
$$\frac{\partial}{\partial y} F(x, y, z(x, y)) = F_x 0 + F_y 1 + F_z \frac{\partial z}{\partial y} = 0$$

which enable us to solve for $\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$ and $\frac{\partial z}{\partial y} = -\frac{F_y}{F_z}$. These expressions will be valid wherever F is differentiable and where $F_z \neq 0$. It should be relatively clear that this same formulation could be done for relations with any number of variables and would give analogous expressions for the partial derivatives of the implicitly defined functions.

Exercises:

(1) Given the relation $7x^2y^6 - 5xy^5 + 3y^3 = 5$, find a formula for $\frac{dy}{dx}$ that is valid for points on the curve in the vicinity of the point (1, 1). Then use the tangent line to approximate the y-coordinate of a point on the graph corresponding to $x = 1.05$.

(2) Find expressions for $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ and state where they are valid if $z(x, y)$ is defined implicitly by the relation $ye^z + xz - x^2 - y^2 = 0$.

(3) Suppose that quantities *u* and *v* are given in terms of *x*, *y*, and *z* by the equations $\left\{\begin{matrix} u = x^2 + y^2 - z^2 \\ v = xy + xz + yz \end{matrix}\right\}$. Further suppose that *x*, *y*, and *z* are points on the sphere of radius *R* parametrized by ϕ and θ as previously described. Calculate each of the Jacobian matrices $\mathbf{J}_F$, $\mathbf{J}_G$, and $\mathbf{J}_{F\circ G}$ in the case where $(\phi, \theta) = (30°, 60°) = (\frac{\pi}{6}, \frac{\pi}{3})$. [Here the functions *F* and *G* are also as previously described.]

(4) Let $z = f(x, y)$ where the Cartesian coordinates (x, y) are related to the polar coordinates (r, θ) by the equations $x = r\cos\theta$ and $y = r\sin\theta$. Do the following:

(a) Find expressions for $\frac{\partial z}{\partial r}$ and $\frac{\partial z}{\partial \theta}$ involving *r*, θ, and the partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

(b) Show that $\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 = \left(\frac{\partial z}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2$.

(5) Two surfaces, S_1 and S_2, are described by the equations:

$$S_1: \quad xy - x^2 + z^2 = 1$$

$$S_2: \quad 3y = 2xz + y^3$$

These surfaces intersect in a curve C that contains the point (1,1,1).

(a) Find equations of the tangent planes to S_1 and S_2 respectively at the point (1,1,1).

(b) Find parametric equations for the line tangent to C at (1,1,1).

(c) On surface S_1, its equation implicitly defines the variable *x* as a function of the other two variables. Give expressions for the two partial derivatives of this function and evaluate these expressions at the point (1,1,1).

(6) Suppose $p = f(x, y, z) = 3xy + yz$ and that *x*, *y*, *z* are functions of *u* and *v*:

$$x = \ln u + \cos v\,,\ y = 1 + u\sin v\,,\ z = uv$$

(a) Use the Chain Rule to find $\frac{\partial p}{\partial u}$ and $\frac{\partial p}{\partial v}$ at $(u, v) = (1, \pi)$.

(b) Suppose now that *u* and *v* are also functions of *t*:

$$u = 1 + \sin(\pi t)\,,\ v = \pi t^2.$$

Use your answer to part (a) and the chain rule to find $\frac{dp}{dt}$ at $t = 1$.

(7) The relations $\left\{\begin{matrix} F(x,y,u,v) = x^2 - y^2 - u^3 + v^2 + 4 = 0 \\ G(x,y,u,v) = 2xy + y^2 - 2u^2 + 3v^4 + 8 = 0 \end{matrix}\right\}$ each define 3-dimensional "hypersurfaces" in $\mathbf{R}^4$ with coordinates (x, y, u, v). The intersection of the two hypersurfaces is 2-dimensional and can be parametrized by (x, y). That is, we can, in theory, express $u = u(x,y)$ and $v = v(x,y)$. Using the chain rule, give expressions for $\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial y}$. (You may leave your answers in terms of *x*, *y*, *u*, and *v*.)