Math 5BI: Problem Set 5
General chain rule

The chain rule for functions of one variable can be written in two ways. Suppose that $z = f(x)$ and $x = g(t)$, where f and g are differentiable functions, then $z = (f \circ g)(t)$, where $f \circ g$ is called the *composition* of f and g. For example, if

$$f(x) = \sin x \quad \text{and} \quad g(t) = e^t, \quad \text{then} \quad (f \circ g)(t) = \sin(e^t).$$

The chain rule states that

$$(f \circ g)'(t) = f'(g(t))g'(t). \tag{1}$$

But it is also customary to write this equation as

$$\frac{dz}{dt} = \frac{dz}{dx}\frac{dx}{dt}. \tag{2}$$

Here the z appearing in the numerator on the left actually stands for the function $f \circ g$ while the z in the numerator on the right stands for the function f. The x appearing in the numerator on the right stands for the function g. We use z to stand for a variable and also two different functions in order to save on usage of letters from the alphabet.

The general chain rule for functions (also called mappings) from $\mathbb{R}^n$ to $\mathbb{R}^m$ can also be written in two ways. If the dependent variables $z_1, \ldots z_m$ are functions of the intermediate variables $x_1, \ldots, x_n$,

$$\begin{aligned} z_1 &= f_1(x_1, \ldots, x_n), \\ z_2 &= f_2(x_1, \ldots, x_n), \\ &\cdots \\ z_m &= f_m(x_1, \ldots, x_n), \end{aligned}$$

and the intermediate variables $x_1, \ldots, x_n$ are in turn functions of the independent variables $t_1, \ldots, t_p$,

$$\begin{aligned} x_1 &= g_1(t_1, \ldots, t_p), \\ x_2 &= g_2(t_1, \ldots, t_p), \\ &\cdots \\ x_n &= g_n(t_1, \ldots, t_p), \end{aligned}$$

then it follows from the chain rule from the previous problem set that

$$\frac{\partial z_i}{\partial t_j} = \frac{\partial z_i}{\partial x_1}\frac{\partial x_1}{\partial t_j} + \ldots + \frac{\partial z_i}{\partial x_n}\frac{\partial x_n}{\partial t_j}. \tag{3}$$

This is the many variable version of (—refE:chainrule2). Putting these formulae for the various i and j into a matrix yields

$$\begin{pmatrix} \frac{\partial z_1}{\partial t_1} & \cdots & \frac{\partial z_1}{\partial t_p} \\ \cdots & \cdots & \cdots \\ \frac{\partial z_m}{\partial t_1} & \cdots & \frac{\partial z_m}{\partial t_p} \end{pmatrix} = \begin{pmatrix} \frac{\partial z_1}{\partial x_1} & \cdots & \frac{\partial x z_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial z_m}{\partial x_1} & \cdots & \frac{\partial z_m}{\partial x_n} \end{pmatrix} \begin{pmatrix} \frac{\partial x_1}{\partial t_1} & \cdots & \frac{\partial x_1}{\partial t_p} \\ \cdots & \cdots & \cdots \\ \frac{\partial x_n}{\partial t_1} & \cdots & \frac{\partial x_n}{\partial t_p} \end{pmatrix} \tag{4}$$

The f_i's and g_j's can be regarded as the component functions of mappings

$$\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m, \quad \text{and} \quad \mathbf{G} : \mathbb{R}^p \to \mathbb{R}^n.$$

The *composition* of these two mappings is the mapping $\mathbf{F} \circ \mathbf{G} : \mathbb{R}^p \to \mathbb{R}^m$ defined by

$$\mathbf{F} \circ \mathbf{G} \begin{pmatrix} t_1 \\ \cdot \\ t_p \end{pmatrix} = \mathbf{F}\left(\mathbf{G}\begin{pmatrix} t_1 \\ \cdot \\ t_p \end{pmatrix}\right).$$

Then (**??**) can be written as

$$D(\mathbf{F} \circ \mathbf{G})(\mathbf{c}) = D\mathbf{F}(\mathbf{G}(\mathbf{c}))D\mathbf{G}(\mathbf{c}), \quad \text{for} \quad \mathbf{c} \in \mathbb{R}^p, \tag{5}$$

where matrix multiplication occurs on the right. This is the many variable version of (**??**).

Of course, if f is a scalar-valued function, say $z = f(x_1, x_2, \ldots, x_n)$.

$$\begin{aligned} Df(x_1, \ldots, x_n) &= \nabla f(x_1, \ldots, x_n) \\ &= \left((\partial f/\partial x_1)(x_1, \ldots, x_n), \quad \cdots, \quad (\partial f/x_n)(x_1, \ldots, x_n)\right). \end{aligned}$$

Note that we can think of the gradient of f as either a column or a row vector. In the next problem it is more convenient to think of it as a row vector.

Problem 5.1. a. Suppose that $z = f(x, y)$ where $x = r\cos\theta$ and $y = r\sin\theta$. Use version (**??**) of the chain rule to find $\partial z/\partial r$ and $\partial z/\partial\theta$ in terms of $\partial z/\partial x$ and $\partial z/\partial y$.

b. Find a formula for $\partial z/\partial x$ and $\partial z/\partial y$ in terms of $\partial z/\partial r$ and $\partial z/\partial\theta$. You can do this by solving for $\partial z/\partial x$ and $\partial z/\partial y$.

c. We define a function $\mathbf{G} : \mathbb{R}^2 \to \mathbb{R}^2$ as follows:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{G}(r, \theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}.$$

Find formulae for Df, $D\mathbf{G}$ and $D(f \circ \mathbf{G})$. (Think of Df and $D(f \circ \mathbf{G})$ as row vectors.) Write out the chain rule from part a as a matrix multiplication

$$D(f \circ \mathbf{G})(r, \theta) = Df(\mathbf{G}(r, \theta)) \cdot D\mathbf{G}(r, \theta).$$

Problem 5.2. a. Suppose again that $x = r\cos\theta$ and $y = r\sin\theta$, where $r = r(t)$ and $\theta = \theta(t)$. Find a formula for dx/dt and dy/dt in terms of dr/dt and $d\theta/dt$.

b. Find a formula for dr/dt and $d\theta/dt$ in terms of dx/dt and dy/dt by solving for dr/dt and $d\theta/dt$.

c. We can define functions $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ and $\mathbf{G} : \mathbb{R} \to \mathbb{R}^2$ as follows:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{F}(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}, \mathbf{G}(t) = \begin{pmatrix} r(t) \\ \theta(t) \end{pmatrix}..$$

Find formulae for $D\mathbf{F}$, $D\mathbf{G}$ and $D(\mathbf{F}\circ\mathbf{G})$. Write out the chain rule from part a as a matrix multiplication

$$D(\mathbf{F}\circ\mathbf{G})(t) = D\mathbf{F}(\mathbf{G}(t)) \cdot D\mathbf{G}(t).$$

d. Use the formulae you have found to rewrite the system of differential equations

$$\begin{aligned} \frac{dx}{dt} &= -3x + 4y, \\ \frac{dy}{dt} &= -4x - 3y, \end{aligned}$$

in terms of r and θ.

e. Solve the resulting system of equations for $r(t)$ and $\theta(t)$.