Intro to Econometrics: Recitation 5

A quick introduction to vector calculus

Gustavo Pereira

October 22, 2019

Contents

1 Intro

In these notes I try to introduce some notation regarding calculus with functions that map vectors into vectors. One reason why things get a bit messy is that when we write $x \in \mathbf{R}^N$ we don't distinguish between

$$\begin{bmatrix} x_1 & \cdots & x_N \end{bmatrix}$$

and

$$\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

but crucially, the operation $L(\mathbf{x})$ where L is a linear map is represented differently by means of matrix multiplication notation; in the top case, $L(x)$ corresponds to a matrix acting on x "on the right", whereas in the bottom case, the matrix acts "on the left". Of course, there is an operation that takes us from the "row world" to the "column world": transposition.

Since derivatives are in fact linear maps, losing track of which side the derivative matrix operates on can lead to dimension inferno. So here I provide a few examples that might shed light on how to deal with this.

2 The meaning of a derivative

It will be useful to recall how derivatives and linear maps are connected. Because these aren't notes in analysis, I won't be as general as I could, neither will I provide any proofs. For proofs and generalizations, check any undergraduate real analysis textbook. I also assume that you are familiar with linear maps and their connection with matrix operations.

Now let's recall the definition of a derivative.

Definition 1. Let $F : \mathbf{R}^n \to \mathbf{R}^p$. The function f is called differentiable at $x_0 \in \mathbf{R}^n$ if there exists a linear map $L : \mathbf{R}^n \to \mathbf{R}^p$ such that

$$\lim_{h \to 0} \frac{\|F(x_0 + h) - F(x_0) - L(h)\|}{\|h\|} = 0$$

we will denote $L = DF(x_0)$. The linear map $DF(x_0)$ is called the *derivative* of F. ◁

An important thing to note is that $DF(x_0)$ is a linear map, so it applies to vectors in $\mathbf{R}^n$. This leaves us with the awkward notation

$$L(h) = DF(x_0)(h)$$

which becomes (maybe?) a bit less ambiguous by adding even more parentheses:

$$L(h) = (DF(x_0))(h)$$

It is sometimes useful to divide $\mathbf{R}^n$ into two sets of coordinates, say $\mathbf{R}^n = \mathbf{R}^{n_1} \times \mathbf{R}^{n_2}$, so we study functions like $F(x, y)$. The partial derivative with respect to the first set of n_1 coordinates, evaluated at (x_0, y_0), is denoted $D_1 F(x_0, y_0)$. It just means the derivative of the map

$$x \mapsto F(x, y_0)$$

evaluated at x_0 (whenever it exists). Whenever we consistently refer to the first set of coordinates as x, we can also write

$$D_x F(x_0, y_0)$$

to denote the same partial derivative.

One important caveat is that both $D_x F(x_0, y_0)$ and $D_y F(x_0, y_0)$ might be defined at a point where F is not differentiable.

2.1 Facts about derivatives to have in mind

I state a proposition that summarizes all that I will use about derivatives. As mentioned earlier, I don't give any proofs but they should be contained in any basic real analysis textbook.

Theorem 1. *Let $F_1 : \mathbf{R}^n \to \mathbf{R}^p$ and $F_2 : \mathbf{R}^n \to \mathbf{R}^P$ be differentiable at $x_0 \in \mathbf{R}^n$, and let $G : \mathbf{R}^k \to \mathbf{R}^n$ be differentiable z_0, where $x_0 = G(z_0)$. Then:*

1. *$F(x) = F_1(x) + F_2(x)$ is differentiable at x_0 and $DF(x_0) = DF_1(x_0) + DF_2(x_0)$*

2. *$H(z) = F_1(G(z))$ is differentiable at z_0 and $DH(z_0) = DF_1(x_0) \circ DG(z_0)$*

[TODO: More in-depth about meaning of the two items. Especially the composition above, and the fact that $DH(z_0)$ is a linear map.]

Theorem 2. *Let $g_1 : \mathbf{R}^m \to \mathbf{R}^{n_1}$ and $g_2 : \mathbf{R}^m \to \mathbf{R}^{n_2}$ both be differentiable at $t_0 \in \mathbf{R}^m$. Let $F : \mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \to \mathbf{R}^p$ be differentiable at $(x_0, y_0) = (g_1(t_0), g_2(t_0))$.*

Then

$$\phi(t) = F(g_1(t), g_2(t))$$

is differentiable at t_0, and

$$D\phi(t_0) = D_x F(x_0, y_0) \circ Dg_1(t_0) + D_y F(x_0, y_0) \circ Dg_2(t_0)$$

3 Seven examples

TODO: finish

1. Take $f_1(x) = Ax$. What is the derivative of f_1? Take $L(h) = Ah$.

   $$f_1(x + h) - f_1(x) - L(h) \equiv 0$$

   hence $Df_1(x) = A$ for all x.

   Importantly, we specified the action above but it's good to repeat it: $(Df_1(x))(h) = Ah$.

   That is, the derivative is A and the action is on the left.

2. Let $f_2(x) = x'B$. Take $T(h) = h'B$.