# Math Notes

Econometrics I
NYU Stern

Professor Paul T. Scott
Email: ptscott@stern.nyu.edu

## 1 Linear Algebra

### 1.1 Inner product and outer product

There are two ways to multiply a vector by itself.

Suppose $\mathbf{x}$ is a $K \times 1$ vector.

The **inner product** of a vector is the dot product of a vector with itself, or the sum of squares of its elements:

$$\mathbf{x}'\mathbf{x} = x_1^2 + x_2^2 + \ldots x_K^2.$$

Note that the inner product involves multiplying a $1 \times K$ vector times a $K \times 1$ vector, so it is $1 \times 1$ – a scalar.

The outer product of a vector flips the order of multiplication:

$$\mathbf{x}'\mathbf{x} = \begin{pmatrix} x_1^2 & x_1 x_2 & \cdots & x_K x_1 \\ x_2 x_1 & x_2^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ x_1 x_K & \cdots & \cdots & x_K^2 \end{pmatrix}.$$

The outer product is a $K \times 1$ vector times a $1 \times K$ vector, so it results in a $K \times K$ matrix.

### 1.2 Derivatives with matrices and vectors

It is helpful to know some rules for how to do differentiation with matrices and vectors.

First, we should define what it means to take a derivative with respect to a vector.

Let $f(\mathbf{x})$ be a scalar-valued function of the $K \times 1$ vector $\mathbf{x}$.

The derivative of $f(\mathbf{x})$ with respect to $\mathbf{x}$ is defined as follows:

$$\frac{df(\mathbf{x})}{d\mathbf{x}} = \begin{pmatrix} \frac{df(\mathbf{x})}{dx_1} & \frac{df(\mathbf{x})}{dx_2} & \cdots & \frac{df(\mathbf{x})}{dx_K} \end{pmatrix},$$

where $x_k$ is the $k$th element of $\mathbf{x}$. Notice that we say that the derivative of a scalar with respect to a column vector is a row vector. We can also define derivatives with respect to vectors in the other way

– i.e., we could say that the resulting derivative is a column vector. It's not conceptually important which way we define it; it would just change the transposing in some of the following results, and we have to be careful when the function we're taking the derivative of is vector-valued because one dimension will refer to the function outputs and the other dimension will refer to the function inputs.

If $\mathbf{f}(\mathbf{x})$ is a vector-valued function with dimension $J \times 1$ , then we define its derivative with respect to the $K \times 1$ vector $\mathbf{x}$ as follows:

$$\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}} = \begin{pmatrix} \frac{d\mathbf{f}_1(\mathbf{x})}{dx_1} & \frac{d\mathbf{f}_1(\mathbf{x})}{dx_2} & \cdots & \frac{d\mathbf{f}_1(\mathbf{x})}{dx_K} \\ \frac{d\mathbf{f}_2(\mathbf{x})}{dx_1} & \frac{d\mathbf{f}_2(\mathbf{x})}{dx_2} & \cdots & \frac{d\mathbf{f}_2(\mathbf{x})}{dx_K} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d\mathbf{f}_J(\mathbf{x})}{dx_1} & \frac{d\mathbf{f}_J(\mathbf{x})}{dx_2} & \cdots & \frac{d\mathbf{f}_J(\mathbf{x})}{dx_K} \end{pmatrix},$$

noting that $\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}$ is a $J \times K$ matrix.

Now, let's turn to the following rules:

$$\frac{d\boldsymbol{a}'\boldsymbol{x}}{d\boldsymbol{x}} = \boldsymbol{a}' \qquad \frac{d\boldsymbol{A}\mathbf{x}}{d\boldsymbol{x}} = \boldsymbol{A} \qquad \frac{d\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{d\boldsymbol{x}} = \left(\boldsymbol{A} + \boldsymbol{A}'\right)\boldsymbol{x}$$

where $\mathbf{a}$ and $\mathbf{x}$ are $K \times 1$ vectors, and $\mathbf{A}$ is a $K \times K$ matrix. Note that if $\mathbf{A}'$ is a symmetric matrix, $\frac{d\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}}{d\boldsymbol{x}} = 2\boldsymbol{x}'\mathbf{A}$.

**Rule 1:** $\frac{d\boldsymbol{a}'\boldsymbol{x}}{d\boldsymbol{x}} = \boldsymbol{a}'$

Notice that

$$\mathbf{a}'\mathbf{x} = a_1x_1 + a_2x_2 + \ldots + a_Kx_K = \sum_{k=1}^{K} a_kx_k.$$

From this formula, it is clear that $\frac{d\mathbf{a}'\mathbf{x}}{dx_k} = a_k$. Therefore, from the definition of the derivative above (with $f(\mathbf{x}) = \mathbf{a}'\mathbf{x}$) we have

$$\frac{d\boldsymbol{a}'\boldsymbol{x}}{d\boldsymbol{x}} = \begin{pmatrix} \frac{d\mathbf{a}'\mathbf{x}}{dx_1} & \frac{d\mathbf{a}'\mathbf{x}}{dx_2} & \ldots & \frac{d\mathbf{a}'\mathbf{x}}{dx_K} \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & \ldots & a_K \end{pmatrix} = \mathbf{a}'.$$

**Rule 2:** $\frac{d\boldsymbol{A}\mathbf{x}}{d\boldsymbol{x}} = \boldsymbol{A}$

Notice that $\boldsymbol{A}\mathbf{x}$ is a $K \times 1$ vector. Also notice that we can write

$$\boldsymbol{A}\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_K \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}_1\mathbf{x} \\ \mathbf{a}_2\mathbf{x} \\ \vdots \\ \mathbf{a}_K\mathbf{x} \end{bmatrix}$$

where $\mathbf{a}_k$ refers to the $k$th row of $\mathbf{A}$. In this case, we can apply the definition of $\frac{d\mathbf{f}(\mathbf{x})}{d\mathbf{x}}$ and Rule 1 conclude that

$$\frac{d\boldsymbol{Ax}}{d\boldsymbol{x}} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_K \end{bmatrix} = \mathbf{A}$$

**Rule 3:** $\frac{d\boldsymbol{x}'\boldsymbol{Ax}}{d\boldsymbol{x}} = \left(\boldsymbol{A} + \boldsymbol{A}'\right)\boldsymbol{x}$

Notice that $\boldsymbol{x}'\boldsymbol{Ax}$ is a $1 \times K$ vector times a $K \times K$ matrix times a $K \times 1$ vector, so it is a scalar. Let's use a change in variables. Define $\mathbf{u}(\mathbf{x}) = \mathbf{A}'\mathbf{x}$, and consider

$$\frac{d\mathbf{u}(\mathbf{x})'\,\boldsymbol{x}}{d\boldsymbol{x}} = \frac{d\boldsymbol{x}'\boldsymbol{Ax}}{d\boldsymbol{x}}.$$

Taking the total derivative, we have

$$\frac{d\mathbf{u}'\boldsymbol{x}}{d\boldsymbol{x}} = \frac{\partial\left(\mathbf{u}'\boldsymbol{x}\right)}{\partial\mathbf{x}} + \left(\frac{d\mathbf{u}(\mathbf{x})}{d\mathbf{x}}\right)'\boldsymbol{x}.$$

Using Rule 1, $\frac{\partial\left(\mathbf{u}'\boldsymbol{x}\right)}{\partial\mathbf{x}} = \mathbf{u}'$. Using Rule 2, $\frac{d\mathbf{u}(\mathbf{x})}{d\mathbf{x}} = \frac{d\mathbf{A}'\mathbf{x}}{d\mathbf{x}} = \mathbf{A}'$. Thus,

$$\frac{d\mathbf{u}'\boldsymbol{x}}{d\boldsymbol{x}} = \mathbf{u}' + \left(\mathbf{A}'\right)'\boldsymbol{x} = \mathbf{A}'\boldsymbol{x} + \mathbf{A}\boldsymbol{x} = \left(\mathbf{A}' + \mathbf{A}\right)\boldsymbol{x},$$

which is equal to $2\mathbf{Ax}$ if $\mathbf{A}$ is a symmetric matrix.