The Gradient of the Quadratic Form of a Symmetric Matrix

450A Political Methodology I, handout

Based on a note by Will Marble

Let A be an $n \times n$ symmetric matrix and let x be a column vector of length n. Consider the quadratic form $x'Ax$. We want to show that its gradient is $\frac{\partial x'Ax}{\partial x} = 2x'A$. We do this by writing out the elements of $x'Ax$ one by one, computing the partial derivatives, putting them into a vector, and showing that this vector equals the matrix product $2x'A$.

Start by writing out x and the columns of A:

$$x'Ax = [x_1 \; x_2 \; \cdots \; x_n]_{1\times n}[A_1 \; A_2 \; \cdots \; A_n]_{n\times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1} \tag{1}$$

where A_i is the i-th column of A, a vector of length n, and A_{ji} is its j-th element (row j, column i of A). The subscripts show right away that $x'Ax$ is 1×1, a scalar. So its gradient will be $1 \times n$, one partial derivative for each of the n elements of x.

Next, multiplying the first two terms, we have

$$x'Ax = [x \cdot A_1 \quad x \cdot A_2 \quad \cdots \quad x \cdot A_n]_{1\times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1}, \tag{2}$$

where $x \cdot A_i = x'A_i$ is the dot product of x and the i-th column of A. Finally, doing the last multiplication, we get

$$\begin{aligned} x'Ax &= (x \cdot A_1)x_1 + (x \cdot A_2)x_2 + \cdots + (x \cdot A_n)x_n \qquad (3) \\ &= (x_1A_{11} + x_2A_{21} + \cdots + x_nA_{n1})x_1 + (x_1A_{12} + x_2A_{22} + \cdots + x_nA_{n2})x_2 \\ &\quad + \cdots + (x_1A_{1n} + x_2A_{2n} + \cdots + x_nA_{nn})x_n. \qquad (4) \end{aligned}$$

In the first line, note the difference between the dot product $x \cdot A_i$ and the product of that scalar with x_i.

To get the gradient with respect to x, we take the vector of partial derivatives. Here is the partial derivative with respect to x_1. The first group of terms contains x_1 twice (once inside the parentheses, once outside); every other group contains it once, inside the parentheses:

$$\begin{aligned} \frac{\partial x'Ax}{\partial x_1} &= (2x_1A_{11} + x_2A_{21} + x_3A_{31} + \cdots + x_nA_{n1}) \\ &\quad + (x_2A_{12} + x_3A_{13} + \cdots + x_nA_{1n}). \qquad (5) \end{aligned}$$

(Write this step out yourself if you are not convinced.)

Because A is symmetric, $A_{1i} = A_{i1}$, so we can rewrite (5) as

$$\frac{\partial x'Ax}{\partial x_1} = 2x_1A_{11} + 2x_2A_{21} + 2x_3A_{31} + \cdots + 2x_nA_{n1} \tag{6}$$

$$= 2\sum_{i=1}^{n} x_iA_{i1}. \tag{7}$$

How do you write that sum in matrix notation? It is $x'A_1$, the dot product of x and the first column of A. So the partial derivative with respect to x_1 is

$$\frac{\partial x'Ax}{\partial x_1} = 2x'A_1.$$

The other partial derivatives work the same way, so in general

$$\frac{\partial x'Ax}{\partial x_i} = 2x'A_i, \tag{8}$$

where again A_i is the i-th column of A. The gradient puts these partial derivatives into a vector:

$$\nabla x'Ax = [2x'A_1 \quad 2x'A_2 \quad \cdots \quad 2x'A_n].$$

This is exactly $2x'$ multiplied by A, so $\nabla x'Ax = 2x'A$, and we are done.

Two remarks. First, some texts write the gradient as a column vector; it is then the transpose, $2Ax$ (using $A' = A$). Second, if A is not symmetric, the same steps give $\nabla x'Ax = x'(A + A')$, which reduces to $2x'A$ when $A' = A$.