Matrix Arithmetic

If you don't remember how to add matrices, you should look it up now. Here we are going to talk about matrix products.

Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$. Let's also say that the matrix A is the coordinate representation of a linear transformation $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$, and likewise for B and $\mathcal{B} : \mathbb{R}^p \to \mathbb{R}^n$. Then linear transformation

$$\mathcal{C} = \mathcal{AB} : \mathbb{R}^p \to \mathbb{R}^n \to \mathbb{R}^m$$

has as its coordinate representation the matrix $C = AB$. While it is true that *a matrix is a rectangular array of numbers*, it will be useful for us to remember that a matrix represents a linear function from one vector space to another: *a matrix is a linear transformation.* This is precisely why the natural product of two matrices isn't entrywise, like addition, but instead has the (maybe not as intuitive) form

$$C = [c_{ij}]\,, \text{ where } c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Let's let a_i^T be the i^{th} row of A, and b_j be the j^{th} column of B. Then $c_{ij} = a_i^T b_j$. If we now let c_j be the j^{th} column of C, we can write $c_j = Ab_j$. Running this backward, you see that AB is a collection of matrix-vector products, each of which is a collection of vector-vector products.

The row-oriented version of this is simply $\chi_i^T = a_i^T B$, where χ_i is the i^{th} row of C.

Now let α_j be the j^{th} column of A and β_i^T the i^{th} row of B. Then $C = \sum_{k=1}^{n} \alpha_k \beta_k^T$. Here is C as a sum of rank 1 matrices (outer products: $(m \times 1) * (1 \times p) = (m \times p)$).

So we can think of a matrix product as a collection of inner products, a collection of matrix-vector products, a collection of vector-matrix products, or a sum of vector-vector products. And these perspectives all come from only partioning A and/or B into columns or rows.

We may find it useful to partition A and B as $A = [A_{ij}]_{i=1:d}^{j=1:e}$, where $A_{ij} \in \mathbb{R}^{m_i \times m_j}$ and $B = [B_{ij}]_{i=1:f}^{j=1:g}$, where $B_{jk} \in \mathbb{R}^{n_j \times n_k}$. This partioning is *conformal* to the product if $e = f$ and $m_j = n_j, j = 1, 2, \ldots, e$. In this setting all of the perspectives above for the matrix product are valid with the elements a_{ij} and b_{ij} replaced by the submatrices A_{ij} and B_{ij}, and this includes replacing columns by block-columns and rows by block-rows.

If you haven't already, now is the time to write down a 2×3 matrix and a 3×4 matrix and try all of the ways above of finding the product. Then partition each and repeat with the partitioned versions. Then partition them differently and do again.