Columns of I

A vector is simply an element of a vector space. In the vector space, we can add vectors to each other and multiply vectors by scalars, and always get another vector in that space (closure properties). If our vector space has finite dimension n, then we can associate a vector with a sequence of n numbers called the *coordinate* representation of the vector (you may have used the term n-tuple).

Enabling this is the idea of a basis: If $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ is a basis for our space, then any v in the space can be written (uniquely) as the *linear combination* $v = x_1v_1 + x_2v_2 + \cdots + x_nv_n$, and the coordinates of v (in the basis \mathcal{B}) is the n-tuple (x_1, x_2, \ldots, x_n) . The coordinate vector of v (wrt \mathcal{B}) is this sequence of n numbers (scalars); I've written them in a row here, but could have written them in a column. It isn't a row or a column, it is just this sequence of scalars.

This was review (I hope), so now to the point: Any finite dimensional vector space over the real numbers is essentially \mathbb{R}^n (where *n* is its dimension). In numerical linear algebra, this is where we live. For convenience, we simply decide to think of our n-tuples as either rows or columns. We will go with columns. So in these pages, we will often confuse a vector in $x \in \mathbb{R}^n$ with a matrix $x \in \mathbb{R}^{n \times 1}$. We simply *identify* x with its coordinate vector in $\mathbb{R}^{n \times 1}$. If we want a row version of x, we would write x^T (the matrix transpose), and it lives in $\mathbb{R}^{1 \times n}$.

The standard ordered basis for \mathbb{R}^n are the columns of the $n \times n$ identity matrix: $I = [e_1, e_2, \ldots, e_n]$, where e_j has a 1 in row j, and 0 elsewhere (if you prefer: $e_j = [\delta_{i,j}] \in \mathbb{R}^{n \times 1}$ (the Kronecker delta)).

Let $A \in \mathbb{R}^{m \times n}$. We might want to refer to the columns of A individually, as in $A = [a_1, a_2, \ldots, a_n]$. But $A = AI = A[e_1, e_2, \ldots, e_n] = [Ae_1, Ae_2, \ldots, Ae_n]$, so the j^{th} column of A is simply Ae_j . Likewise, the j^{th} row of A is $e_j^T A$, and the element $a_{i,j}$ is just $e_i^T Ae_j$. Notice the double-duty of the e_j notation: in $a_{ij} = e_i^T Ae_j$, e_i is the i^{th} column of $I_{m \times m}$, while e_j is the j^{th} column of $I_{n \times n}$.

So you can write $A = [a_1, a_2, \ldots, a_n]$, and a_j is the j^{th} column of A, or simply denote the j^{th} column of A as Ae_j .

Now convince yourself that we can also write $I_{n \times n} = e_1 e_1^T + e_2 e_2^T + \dots + e_n e_n^T$. Actually do it: let n = 3, write $e_1 e_1^T + e_2 e_2^T + e_3 e_3^T$ as vectors, multiply the vectors, and see the sum. Seriously, if you haven't already, please do this.

Then for any A and B than can be multiplied, define C = AB and $c_{ij} = e_i^T (AB)e_j = (e_i^T A)(Be_j)$ gives the dot-product definition $c_{ij} = \sum a_{ik}b_{kj}$, $Ce_j = (AB)e_j = A(Be_j)$ gives j^{th} col of AB as A times j^{th} col of B, $e_i^T C = e_i^T (AB) = (e_i^T A)B$ gives i^{th} row of AB as i^{th} row of A times B, and $C = AIB = A(\sum e_i e_i^T)B = \sum (Ae_i)(e_i^T B)$ is the 'outer-product' product.

We will use all 4 perspectives above this semester. In fact, this is so basic and important that you will soon feel lost if you don't make it part of you.