Topic 6: Matrices

02-680: Essentials of Mathematics and Statistics

September 5, 2024

You can almost think of a ***matrix*** as a 2-dimension vector. We say that an "n-by-m" matrix $M \in \mathbb{R}^{n\times m}$ has n rows and m columns and we usually write it as:

$$M = \begin{bmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,m} \\ M_{2,1} & M_{2,2} & \dots & M_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,m} \end{bmatrix}$$

1 Simple Matrix Operations

1.1 Addition and Scalar Multiplication.

Like with vectors, addition of two matrices as well as scalar multiplication are element-wise operations, so for matrices $M, N \in \mathbb{R}^{n\times m}$ and scalar $a \in \mathbb{R}$:

$$O = M + N \to O_{i,j} = M_{i,j} + N_{i,j} \;\; \forall 1 \le i \le n, 1 \le j \le m$$

$$O = aM \to O_{i,j} = aM_{i,j} \;\; \forall 1 \le i \le n, 1 \le j \le m$$

1.2 Transpose

For a given matrix $M \in \mathbb{R}^{n\times m}$, the transpose $M^T \in \mathbb{R}^{m\times n}$ is defined such that:

$$\forall I \in [0, n-1], j \in [0, m-1] : M^T_{j,i} = M_{i,j}$$

This operation works for both matrixes and vectors (which are really $n \times 1$ matrices). Some examples:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \qquad\qquad \begin{bmatrix} 7 \\ 8 \\ 9 \\ 10 \end{bmatrix}^T = \begin{bmatrix} 7 & 8 & 9 & 10 \end{bmatrix}$$

2 Matrix Multiplication

Just like with vectors, multiplying two matrices is more complicated than scalars. The first question is the size of the result, if we multiply $C \in \mathbb{R}^{n \times p}$ with $D \in \mathbb{R}^{p \times m}$ we get a matrix $E \in \mathbb{R}^{n \times m}$; notice that the *inner* dimensions are the same. And the values in E are defined as follows:

$$E_{i,j} = \sum_{k=1}^{m} C_{i,k} Dk, j$$

We can actually rewrite this using dot product, lets say that $C_{i,*}$ is the i-th column of C, and $D_{*,j}$ is the j-th column of D. In that case

$$E_{i,j} = C_{i,*} \cdot D^T_{*,j}.$$

What can we do with it? Lets define the following:

- G is an n-by-m matrix where $G_{i,j} = 1$ if actor i was in an episode of the show j (and 0 otherwise)
- H be an m-by-p matrix where $H_{j,k} = 1$ if the show j is available to stream on service k (and 0 otherwise)

3 Square Matrices

Square matrices (that is, matrices where $m = n$) come up a lot, possibly because of this or vice versa there are several properties and operations that exist only on these.

In a square matrix $N \in \mathbb{R}^{n \times n}$, we define the ***main diagonal*** as the entries where the horizontal and vertical component are equal; i.e. $\{N_{i,i} \mid 1 \leq i \leq n\}$.

Symmetry. We say a square matrix is ***symmetric*** if $A = A^T$ (and *anti-symmetric* is $A = -A^T$). That is, A is symmetric if it is mirrored across the main diagonal which often happens for things like distance matrices (though not always as we'll see). Similarly, it is anti-symmetric if it's mirrored across the *anti-diagonal.*

Trace. The ***trace*** of a matrix $tr(A)$ is the sum of the diagonal elements:

$$tr(A) := \sum_{I=1}^{n} A_{i,i}.$$

The trace does not change under transpose, and is distributive across sum and scalar product.

3.1 Identity Matrix

The ***identity*** matrix $I_n \in \mathbb{R}^{n\times n}$ (sometimes simplified to just I when the size is implied from context) is a special symmetric matrix where the main diagonal values are 1 and all other values are 0.

$$\forall I, j \in [1, n] : I_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Note, I_n is symmetric and $tr(I_n) = n$.

Useful References

Liben-Nowell, "Connecting Discrete Mathematics and Computer Science, 2e". §2.4