

Topic 6: Matrices

02-680: Essentials of Mathematics and Statistics

September 19, 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 $a \in \mathbb{R}$:

$$O = M + N \rightarrow O_{i,j} = M_{i,j} + N_{i,j} \quad \forall 1 \leq i \leq n, 1 \leq j \leq m$$

$$O = aM \rightarrow O_{i,j} = aM_{i,j} \quad \forall 1 \leq i \leq n, 1 \leq j \leq m$$

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:

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

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

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

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:

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

We can actually rewrite this using dot product, but lets quickly introduce some notation. Lets first say for a matrix $A \in \mathbb{R}^{n \times m}$ we could say that

$$A = A_{[n],[m]}$$

remember here that $[n] \iff [1, n] \iff \{1, 2, \dots, n\}$. So really the equation above redefines A using a list of columns and a list of rows. That means we can let $C_{i,[p]}$ is the i -th column of C , and $D_{[p],j}$ is the j -th column of D . In that case

$$E_{i,j} = C_{i,[p]} \cdot D_{[p],j}^T.$$

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 $tr(I_n) = n$.

Note also that for any matrix $A \in \mathbb{R}^{n \times m}$

$$AI_n = A \quad \text{and} \quad I_n A = A.$$

3.2 Determinants

We define the *determinant* of a square matrix $det : \mathbb{R}^{n \times n} \mapsto \mathbb{R}$ as a function with the range of all square matrices and a codomain of real numbers. We often write this as $|A|$ for $A \in \mathbb{R}^{n \times n}$.

We define determinant *recursively* (meaning it is a function makes a reference to itself), but we first need to define a method for constructing sub-matrices.

Using the notation of sets of column/row indexes ($A = A_{[n],[n]}$) can then use set math to manipulate those rows/columns (mainly using \setminus):

$$A_{[n] \setminus i, [n] \setminus j}.$$

Which is A with all but row i and all but column j .

For instance:

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \quad A_{[3] \setminus 2, [3]} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

To make this easier we will actually shorten this to:

$$A_{[n] \setminus i, [n] \setminus j} \iff A_{\setminus i, \setminus j}.$$

We need that notation to more easily define the determinate for any chosen j :

$$|A| := (-1)^{(i+j)} \sum_{i=1}^n A_{ij} |A_{\setminus i, \setminus j}|.$$

(It can also be defined for a fixed i and sum over i .)

Some explicit examples:

$$|[A_{11}]| = A_{11}$$

$$\left| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \right| = A_{11}A_{22} - A_{21}A_{12}$$

$$\left| \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \right| = A_{11} \left| \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix} \right| - A_{12} \left| \begin{bmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{bmatrix} \right| + A_{13} \left| \begin{bmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} \right|$$

Useful References

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

Wilder, “10-606-f23:Lecture 3” GitHub repository, https://github.com/bwilder0/10606-f23/blob/main/files/notes_linalg.pdf

Kolter, “Linear Algebra Review and Reference”, <https://www.cs.cmu.edu/~zkolter/course/15-884/linalg-review.pdf> §1.1,2.3,3.1-3.5