Contents

0.1 Matrix Definitions and Operations

Definition 1.1 (Matrix). A real-valued (m, n) matrix A with $m, n \in \mathbb{N}$ is an $m \cdot n$-tuple of elements a_{ij}, where $i = 1, \ldots, m$ and $j = 1, \ldots, n$, arranged in a rectangular pattern with m rows and n columns:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Matrix addition is an operation performed on two matrices of the same size. Given matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$, adding the matrices together involves adding their corresponding elements (Element-Wise):

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

Matrix multiplication is performed between two matrices, $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times k}$, and results in $C = AB \in \mathbb{R}^{m\times k}$. The product C is calculated by taking the dot product of each row of matrix A with each column of matrix B. We sum the products of the multiplication to obtain an entry in the corresponding row i and column j of C.

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$$

- a_{ik} is the element at row i and column k of matrix A,
- b_{kj} is the element at row k and column j of matrix B,
- n is the number of columns in matrix A (which is also the number of rows in matrix B).

such that

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \in \mathbb{R}^{3\times 2}, \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \in \mathbb{R}^{2\times 3}$$

$$C = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} & a_{31}b_{13} + a_{32}b_{23} \end{bmatrix}$$

Definition 1.2 (Identity Matrix). A matrix in $\mathbb{R}^{n\times n}$ that contains only 1's on its diagonal entries.

$$I_n := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Definition 1.3 (Linear Independence). In a linearly independent matrix, each column holds the property that it cannot be expressed by performing linear combinations of the other columns. i.e. if a matrix has 3 columns, you cannot multiply columns 2 or 3 by any value $c \in \mathbb{R}$ such that the columns would be the same as column 1 and vice versa. It also holds that adding to columns together with the ability to scale them also cannot equal another column in the matrix.
The formal definition states: A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V over a field F is said to be linearly independent if the only scalars $c_1, c_2, \ldots, c_n$ satisfying the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$$

the only solution to this should be $c_1 = c_2 = \ldots = c_n = 0$. In other words, the only way to obtain the zero vector $\mathbf{0}$ as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is by setting all the scalar coefficients to zero.

0.2 Tensors