Matrices

A **matrix** is a rectangular array of values. For example, the following matrix $\mathbf{A}$ has n rows and m columns:

$$\mathbf{A} := \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}$$

.

Matrices are useful representations of data that can be interpreted and used in a variety of contexts:

1. *A matrix as a 2-dimensional array of numbers:* A matrix can be used to represent data in the form of a table.

2. *A matrix as a list of vectors:* A matrix can be used to represent a collection of vectors.

3. *A matrix characterizes a linear function on vectors:* We can define operations between matrices and vectors so that matrices can act as functions on vectors. The study of matrices as linear functions is central to the field of linear algebra.

Row and column vectors

Recall a vector, $\mathbf{x}$, can be represented as an ordered list of values $x_1, x_2, \dots, x_n$. We can treat $\mathbf{x}$ like a matrix that has either one row or one column. When treating $\mathbf{x}$ as a matrix, we can choose whether we want to orient it horizontally, thereby creating a $1 \times n$ matrix, or vertically, thereby creating a $n \times 1$ matrix. If we orient $\mathbf{x}$ horizontally, we form a matrix called a **row vector**:

$$\begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,n} \end{bmatrix}$$

.

If we orient $\mathbf{x}$ vertically, we form a matrix called a **column vector**:

$$\begin{bmatrix} x_{1,1} \\ x_{2,1} \\ \vdots \\ x_{n,1} \end{bmatrix}$$

.

Notation

We use the following notation for discussing matrices:

- Given a set C, we let $C^{m \times n}$ denote the set of all matrices of m rows and n columns consisting of items from set C
- Given a matrix $\mathbf{A} \in C^{m \times n}$, we let $a_{i,j}$ denote the item at the ith row and jth column of $\mathbf{A}$
- Given a matrix $\mathbf{A} \in C^{m \times n}$, we let $\mathbf{a}_{i,*}$ denote the ith row-vector of $\mathbf{A}$
- Given a matrix $\mathbf{A} \in C^{m \times n}$, we let $\mathbf{a}_{*,j}$ denote the jth column-vector of $\mathbf{A}$

Interpretations of matrices

A matrix as a 2-dimensional array

The most simple interpretation of matrix is as a two-dimensional array of numbers.

For example, the pixels of an image can be represented as a matrix. Let's say we have an image of $m \times n$ pixels. We can let $\mathbf{X}$ be a matrix representing this image where $x_{i,j}$ represents the pixel at row i and column j:

$$\mathbf{X} := \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ x_{2,1} & x_{2,2} & \dots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \dots & x_{m,n} \end{bmatrix}$$

.

A matrix as a list of vectors

A matrix can be thought of as a set of vectors. The ith row of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, denoted $\mathbf{a}_{i,*}$, is a row-vector in $\mathbb{R}^n$. The jth row of this matrix, denoted $\mathbf{a}_{*,j}$ is a column-vector in $\mathbb{R}^m$.

For example, say we have a data set that consists of n samples where each sample has m attributes. We can store this data in a matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$ where each column of the matrix is a data point:

$$\mathbf{X} := \begin{bmatrix} \mathbf{x}_{*,1} & \mathbf{x}_{*,2} & \dots & \mathbf{x}_{*,n} \end{bmatrix}$$

. The column-vector $\mathbf{x}_{*,j} \in \mathbb{R}^m$ is the jth sample in the dataset.

Similarly, we could orient each sample as a row-vector thereby forming a data matrix $\mathbf{X} \in \mathbb{R}^{n \times m}$:

$$\mathbf{X} := \begin{bmatrix} \mathbf{x}_{1,*} \\ \mathbf{x}_{2,*} \\ \vdots \\ \mathbf{x}_{n,*} \end{bmatrix}$$

. The row-vector $\mathbf{x}_{i,*} \in \mathbb{R}^m$ is the ith sample in the dataset.

A matrix characterizes a linear function on vectors

As will be described in another section, we can define operations between matrices and vectors so that matrices can act as functions on vectors. The study of matrices as linear functions is central to the field of linear algebra.