Matrix Operations

- Linear Combination
- Vector Algebra
- Angle Between Vectors
- Projections and Reflections
- Equality of matrices, Augmented Matrix
- Matrix Addition and Matrix Scalar Multiply
- Matrix Multiply
  - Matrix multiply as a dot product extension
  - Matrix multiply as a linear combination of columns
  - How to multiply matrices on paper
- Special matrices and square matrices
- Matrix Algebra and Matrix Multiply Properties
- Transpose
- Inverse matrix, Inverse of a $\mathbf{2} \times \mathbf{2}$ Matrix

Linear Combination

A **linear combination** of vectors $\mathbf{v}_1,\ldots,\mathbf{v}_k$ is defined to be a sum

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k,$$

where $c_1,\ldots,c_k$ are constants.

Vector Algebra

The **norm** or **length** of a fixed vector $\vec{X}$ with components $x_1, \ldots, x_n$ is given by the formula

$$|\vec{X}| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

The **dot product** $\vec{X} \cdot \vec{Y}$ of two fixed vectors $\vec{X}$ and $\vec{Y}$ is defined by

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n.$$

Angle Between Vectors

If $n = 3$, then $|\vec{X}||\vec{Y}|\cos\theta = \vec{X}\cdot\vec{Y}$ where θ is the **angle between** $\vec{X}$ and $\vec{Y}$. In analogy, two n-vectors are said to be **orthogonal** provided $\vec{X}\cdot\vec{Y} = 0$. It is usual to require that $|\vec{X}| > 0$ and $|\vec{Y}| > 0$ when talking about the angle θ between vectors, in which case we *define* θ to be the acute angle $(0 \le \theta < \pi)$ satisfying

$$\cos\theta = \frac{\vec{X}\cdot\vec{Y}}{|\vec{X}||\vec{Y}|}.$$

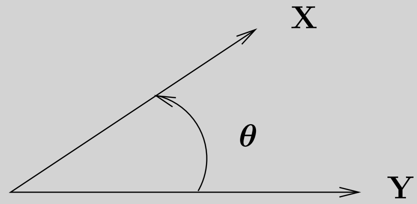

Figure 1. Angle θ between two vectors $\mathbf{X}$, $\mathbf{Y}$.

Projections

The **shadow projection** of vector $\vec{X}$ onto the direction of vector $\vec{Y}$ is the number d defined by

$$d = \frac{\vec{X} \cdot \vec{Y}}{|\vec{Y}|}.$$

The triangle determined by $\vec{X}$ and $(d/|\vec{Y}|)\vec{Y}$ is a right triangle.

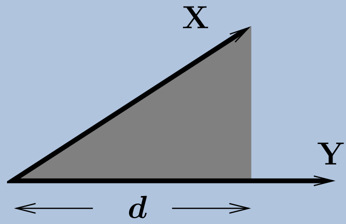

Figure 2. Shadow projection d of vector $\mathbf{X}$ onto the direction of vector $\mathbf{Y}$.

The **vector projection** of $\vec{X}$ onto the line L through the origin in the direction of $\vec{Y}$ is defined by

$$\mathbf{proj}_{\vec{Y}}(\vec{X}) = d\frac{\vec{Y}}{|\vec{Y}|} = \frac{\vec{X} \cdot \vec{Y}}{\vec{Y} \cdot \vec{Y}}\vec{Y}.$$

Reflections

The **vector reflection** of vector $\vec{X}$ in the line L through the origin having the direction of vector $\vec{Y}$ is defined to be the vector

$$\mathbf{refl}_{\vec{Y}}(\vec{X}) = 2\,\mathbf{proj}_{\vec{Y}}(\vec{X}) - \vec{X} = 2\frac{\vec{X}\cdot\vec{Y}}{\vec{Y}\cdot\vec{Y}}\vec{Y} - \vec{X}.$$

It is the formal analog of the complex conjugate map $a + ib \longrightarrow a - ib$ with the x-axis replaced by line L.

Equality of matrices

Two matrices A and B are said to be **equal** provided they have identical row and column dimensions and corresponding entries are equal. Equivalently, A and B are equal if they have identical columns, or identical rows.

Augmented Matrix

If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are fixed vectors, then the augmented matrix A of these vectors is the matrix package whose columns are $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and we write

$$A = \mathbf{aug}(\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

Similarly, when two matrices A and B can be appended to make a new matrix C, we write

$$C = \mathbf{aug}(A, B).$$

Matrix Addition __ Addition of two matrices is defined by applying fixed vector addition on corresponding columns. Similarly, an organization by rows leads to a second definition of matrix addition, which is exactly the same:

$$\begin{pmatrix} \boldsymbol{a}_{11} & \cdots & \boldsymbol{a}_{1n} \\ \boldsymbol{a}_{21} & \cdots & \boldsymbol{a}_{2n} \\ & \vdots & \\ \boldsymbol{a}_{m1} & \cdots & \boldsymbol{a}_{mn} \end{pmatrix} + \begin{pmatrix} \boldsymbol{b}_{11} & \cdots & \boldsymbol{b}_{1n} \\ \boldsymbol{b}_{21} & \cdots & \boldsymbol{b}_{2n} \\ & \vdots & \\ \boldsymbol{b}_{m1} & \cdots & \boldsymbol{b}_{mn} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_{11} + \boldsymbol{b}_{11} & \cdots & \boldsymbol{a}_{1n} + \boldsymbol{b}_{1n} \\ \boldsymbol{a}_{21} + \boldsymbol{b}_{21} & \cdots & \boldsymbol{a}_{2n} + \boldsymbol{b}_{2n} \\ & \vdots & \\ \boldsymbol{a}_{m1} + \boldsymbol{b}_{m1} & \cdots & \boldsymbol{a}_{mn} + \boldsymbol{b}_{mn} \end{pmatrix}.$$

Matrix Scalar Multiply

Scalar multiplication of matrices is defined by applying scalar multiplication to the columns or rows:

$$\boldsymbol{k}\begin{pmatrix} \boldsymbol{a}_{11} & \cdots & \boldsymbol{a}_{1n} \\ \boldsymbol{a}_{21} & \cdots & \boldsymbol{a}_{2n} \\ & \vdots & \\ \boldsymbol{a}_{m1} & \cdots & \boldsymbol{a}_{mn} \end{pmatrix} = \begin{pmatrix} \boldsymbol{k}\boldsymbol{a}_{11} & \cdots & \boldsymbol{k}\boldsymbol{a}_{1n} \\ \boldsymbol{k}\boldsymbol{a}_{21} & \cdots & \boldsymbol{k}\boldsymbol{a}_{2n} \\ & \vdots & \\ \boldsymbol{k}\boldsymbol{a}_{m1} & \cdots & \boldsymbol{k}\boldsymbol{a}_{mn} \end{pmatrix}.$$

Both operations on matrices are motivated by considering a matrix to be a long single array or *vector*, to which the standard fixed vector definitions are applied. The operation of addition is properly defined exactly when the two matrices have the same row and column dimensions.

Matrix Multiply

College algebra texts cite the definition of matrix multiplication as *the product $\boldsymbol{AB}$ equals a matrix $\boldsymbol{C}$ given by the relations*

$$\boldsymbol{c_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}, \quad 1 \leq i \leq m,\ 1 \leq j \leq k.}$$

Matrix multiply as a dot product extension

The college algebra definition of $\boldsymbol{C = AB}$ can be written in terms of dot products as follows:

$$\boldsymbol{c_{ij} = \mathbf{row}(A, i) \cdot \mathbf{col}(B, j).}$$

The general scheme for computing a matrix product $\boldsymbol{AB}$ can be written as

$$\boldsymbol{AB = \mathbf{aug}(A\,\mathbf{col}(B, 1), \ldots, A\,\mathbf{col}(B, n)).}$$

Each product $\boldsymbol{A\,\mathbf{col}(B, j)}$ is computed by taking dot products. Therefore, matrix multiply can be viewed as a dot product extension which applies to packages of fixed vectors. A matrix product $\boldsymbol{AB}$ is properly defined only in case the number of matrix rows of $\boldsymbol{B}$ equals the number of matrix columns of $\boldsymbol{A}$, so that the dot products on the right are defined.

Matrix multiply as a linear combination of columns

The identity

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1\begin{pmatrix} a \\ c \end{pmatrix} + x_2\begin{pmatrix} b \\ d \end{pmatrix}$$

implies that $A\mathrm{x}$ is a linear combination of the columns of A, where A is the 2×2 matrix on the left.

This result holds in general. Assume $A = \mathbf{aug}(\mathrm{v}_1, \ldots, \mathrm{v}_n)$ and $\vec{X}$ has components $x_1, \ldots, x_n$. Then the definition of matrix multiply implies

$$A\vec{X} = x_1\mathrm{v}_1 + x_2\mathrm{v}_2 + \cdots + x_n\mathrm{v}_n.$$

This relation is used so often, that we record it as a formal result.

Theorem 1 (Linear Combination of Columns)

The product of a matrix A and a vector x satisfies

$$A\mathrm{x} = x_1\,\mathrm{col}(A, 1) + \cdots + x_n\,\mathrm{col}(A, n)$$

where $\mathrm{col}(A, i)$ denotes column i of matrix A.

How to multiply matrices on paper

Most persons make arithmetic errors when computing dot products

$$\begin{pmatrix} -7 & 3 & 5 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 3 \\ -5 \end{pmatrix} = -9,$$

because alignment of corresponding entries must be done mentally. It is visually easier when the entries are aligned.

On paper, the work for a matrix times a vector can be arranged so that the entries align. The transcription above the matrix columns is temporary, erased after the dot product step.

$$\begin{array}{c} \begin{array}{ccc} -1 & \;\;3 & -5 \end{array} \\ \begin{pmatrix} -7 & 3 & 5 \\ -5 & -2 & 3 \\ 1 & -3 & -7 \end{pmatrix} \end{array} \cdot \begin{pmatrix} -1 \\ 3 \\ -5 \end{pmatrix} = \begin{pmatrix} -9 \\ -16 \\ 25 \end{pmatrix}$$

Special matrices

The **zero matrix**, denoted $\mathbf{0}$, is the $m \times n$ matrix all of whose entries are zero. The **identity matrix**, denoted I, is the $n \times n$ matrix with ones on the diagonal and zeros elsewhere: $a_{ij} = 1$ for $i = j$ and $a_{ij} = 0$ for $i \neq j$.

$$\mathbf{0} = \begin{pmatrix} 0\,0 \cdots 0 \\ 0\,0 \cdots 0 \\ \vdots \\ 0\,0 \cdots 0 \end{pmatrix}, \quad I = \begin{pmatrix} 1\,0 \cdots 0 \\ 0\,1 \cdots 0 \\ \vdots \\ 0\,0 \cdots 1 \end{pmatrix}.$$

The **negative** of a matrix A is $(-1)A$, which multiplies each entry of A by the factor (-1):

$$-A = \begin{pmatrix} -a_{11} \cdots -a_{1n} \\ -a_{21} \cdots -a_{2n} \\ \vdots \\ -a_{m1} \cdots -a_{mn} \end{pmatrix}.$$

Square matrices

An $n \times n$ matrix A is said to be **square**. The entries a_{kk}, $k = 1, \ldots, n$ of a square matrix make up its **diagonal**. A square matrix A is **lower triangular** if $a_{ij} = 0$ for $i > j$, and **upper triangular** if $a_{ij} = 0$ for $i < j$; it is **triangular** if it is either upper or lower triangular. Therefore, an upper triangular matrix has all zeros below the diagonal and a lower triangular matrix has all zeros above the diagonal. A square matrix A is a **diagonal matrix** if $a_{ij} = 0$ for $i \neq j$, that is, the off-diagonal elements are zero. A square matrix A is a **scalar matrix** if $A = cI$ for some constant c.

$$\text{upper triangular} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \quad \text{lower triangular} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ & & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

$$\text{diagonal} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \quad \text{scalar} = \begin{pmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & c \end{pmatrix}.$$

Matrix algebra

A matrix can be viewed as a single long array, or fixed vector, therefore the toolkit for fixed vectors applies to matrices.

Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ be matrices of the same row and column dimensions and let $\boldsymbol{k_1}$, $\boldsymbol{k_2}$, $\boldsymbol{k}$ be constants. Then

Closure	The operations $\boldsymbol{A}+\boldsymbol{B}$ and $\boldsymbol{kA}$ are defined and result in a new matrix of the same dimensions.	
Addition rules	$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$	commutative
	$\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})=(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}$	associative
	Matrix $\boldsymbol{0}$ is defined and $\boldsymbol{0}+\boldsymbol{A}=\boldsymbol{A}$	zero
	Matrix $-\boldsymbol{A}$ is defined and $\boldsymbol{A}+(-\boldsymbol{A})=\boldsymbol{0}$	negative
Scalar multiply rules	$\boldsymbol{k}(\boldsymbol{A}+\boldsymbol{B})=\boldsymbol{kA}+\boldsymbol{kB}$	distributive I
	$(\boldsymbol{k_1}+\boldsymbol{k_2})\boldsymbol{A}=\boldsymbol{k_1A}+\boldsymbol{k_2B}$	distributive II
	$\boldsymbol{k_1}(\boldsymbol{k_2A})=(\boldsymbol{k_1k_2})\boldsymbol{A}$	distributive III
	$\boldsymbol{1}\,\boldsymbol{A}=\boldsymbol{A}$	identity

These rules collectively establish that the set of all $\boldsymbol{m}\times\boldsymbol{n}$ matrices is an abstract vector space.

Matrix Multiply Properties ___ The operation of matrix multiplication gives rise to some new matrix rules, which are in common use, but do not qualify as vector space rules.

Associative $\boldsymbol{A}(\boldsymbol{B}\boldsymbol{C}) = (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C}$, provided products $\boldsymbol{B}\boldsymbol{C}$ and $\boldsymbol{A}\boldsymbol{B}$ are defined.

Distributive $\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A}\boldsymbol{B} + \boldsymbol{A}\boldsymbol{C}$, provided products $\boldsymbol{A}\boldsymbol{B}$ and $\boldsymbol{A}\boldsymbol{C}$ are defined.

Right Identity $\boldsymbol{A}\boldsymbol{I} = \boldsymbol{A}$, provided $\boldsymbol{A}\boldsymbol{I}$ is defined.

Left Identity $\boldsymbol{I}\boldsymbol{A} = \boldsymbol{A}$, provided $\boldsymbol{I}\boldsymbol{A}$ is defined.

Transpose

Swapping rows and columns of a matrix $\boldsymbol{A}$ results in a new matrix $\boldsymbol{B}$ whose entries are given by $\boldsymbol{b}_{ij} = \boldsymbol{a}_{ji}$. The matrix $\boldsymbol{B}$ is denoted $\boldsymbol{A}^T$ (pronounced "$\boldsymbol{A}$-*transpose*"). The transpose has these properties:

$$(\boldsymbol{A}^T)^T = \boldsymbol{A} \qquad \text{Identity}$$
$$(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T \qquad \text{Sum}$$
$$(\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T \qquad \text{Product}$$
$$(\boldsymbol{k}\boldsymbol{A})^T = \boldsymbol{k}\boldsymbol{A}^T \qquad \text{Scalar}$$

A matrix $\boldsymbol{A}$ is said to be **symmetric** if $\boldsymbol{A}^T = \boldsymbol{A}$, which implies that the row and column dimensions of $\boldsymbol{A}$ are the same and $\boldsymbol{a}_{ij} = \boldsymbol{a}_{ji}$.

Inverse matrix

A square matrix $\boldsymbol{B}$ is said to be an **inverse** of a square matrix $\boldsymbol{A}$ provided $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$. The symbol $\boldsymbol{I}$ is the identity matrix of matching dimension. A given matrix $\boldsymbol{A}$ may not have an inverse, for example, $\mathbf{0}$ times any square matrix $\boldsymbol{B}$ is $\mathbf{0}$, which prohibits a relation $\mathbf{0}\boldsymbol{B} = \boldsymbol{B}\mathbf{0} = \boldsymbol{I}$. When $\boldsymbol{A}$ does have an inverse $\boldsymbol{B}$, then the notation $\boldsymbol{A}^{-1}$ is used for $\boldsymbol{B}$, hence $\boldsymbol{AA}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}$.

Theorem 2 (Inverses)

Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ denote square matrices. Then

(a) A matrix has at most one inverse, that is, if $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$ and $\boldsymbol{AC} = \boldsymbol{CA} = \boldsymbol{I}$, then $\boldsymbol{B} = \boldsymbol{C}$.

(b) If $\boldsymbol{A}$ has an inverse, then so does $\boldsymbol{A}^{-1}$ and $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$.

(c) If $\boldsymbol{A}$ has an inverse, then $(\boldsymbol{A}^{-1})^{T} = (\boldsymbol{A}^{T})^{-1}$.

(d) If $\boldsymbol{A}$ and $\boldsymbol{B}$ have inverses , then $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

Inverse of a 2×2 Matrix

Theorem 3 (Inverse of a 2×2)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

In words, the theorem says:

> Swap the diagonal entries, change signs on the off-diagonal entries, then divide by the determinant $ad - bc$.