Linear Algebra

Vectors

Matrices

To Be Defined

Vectors

Interpretations of Vectors

- **Algebraic vectors** ($\boldsymbol{v}$, $\vec{v}$): an ordered list of numbers, e.g.,

$$\boldsymbol{v} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

  - Vectors can be written as rows (seen above) or columns (seen below), but differ only at the level of notation, some notation is more useful with certain conventions.
  - The order of elements in a vector matters:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

- **Dimensionality**: the number of elements in a vector, where each new element provides new information, or geometrically, a new direction.
- **Euclidean (geometric, spatial) vectors**: a line in geometric space that indicates the magnitude and direction from a starting point (tail) to an end point (head).
  - Geometric vectors can start at any point in space, but often represented as starting from the origin—such vectors are in standard position.
  - Coordinates are not the same as vectors, but they do indicate where the head of a vector will land if it is in standard position.

Vector Addition and Subtraction

  - Algebraically, dimensionality of vectors must be equal. When they are, then addition or subtraction vectors is done on the corresponding elements of each vector, e.g.,

$$\begin{bmatrix} 1 \\ 0 \\ 4 \\ 5 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \\ -6 \\ 11 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ -2 \\ 16 \end{bmatrix}$$

  - Geometrically, addition can be thought of translating the tail of one vector to the head of the other—resulting in a new vector.
  - Geometric interpretations of subtraction can be thought of in two ways:

  1. Multiplying one vector by -1, then applying vector addition method above.
  2. Placing both vectors in standard position, with the resulting vector between the two heads being the answer.

Vector Scalar Multiplication

- **Scalar**: typically denoted with lower case Greek letters (e.g., $\alpha, \ \lambda$) indicating an element of a field (typically real numbers) used in scalar multiplication of vectors.
- Algebraically, scalar multiplication is the multiplication of each element of a vector by a particular scalar, e.g.,

$$\lambda \boldsymbol{v} \rightarrow 7 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -7 \\ 0 \\ 7 \end{bmatrix}$$

- Geometrically, scalar-multiplication can be thought of as the extension ($\lambda > 1$) or compression ($\lambda \in (0, 1)$) of a vector.
  - When $\lambda < 0$, then it can be thought of inverting its direction with respect to the origin.

The Dot Product

- **Dot (scalar, inner) product**: an algebraic operation that takes two equal-length sequences of numbers (usually coordinate vectors), and returns a single number.
  - The result of a dot product is a scalar↑, so often it is represented as such.
    - It can also be represented as multiplication between two vectors ($\boldsymbol{a} \cdot \boldsymbol{b}$).
    - However, it is commonly represented as $\boldsymbol{a}^T \boldsymbol{b}$ – transpose↓ will be explained in more detail when dealing with matrix multiplication↓.
  - Algebraically: $\sum_{i=1}^{n} \boldsymbol{a}_i \boldsymbol{b}_i$ – where Σ denotes summation and n is the dimension of the vector space, e.g.,

$$\begin{bmatrix} 1 & 3 & 5 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix} = (1 \cdot 4) + (3 \cdot -2) + (-5 \cdot -1) = 3$$

Properties of the Dot Product

- Note: the following properties hold as long as $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are real vectors.
- ✓ Distributive: $\boldsymbol{a}^T(\boldsymbol{b} + \boldsymbol{c}) = \ \boldsymbol{a}^T \boldsymbol{b} + \boldsymbol{a}^T \boldsymbol{c}$ – vector multiplication distributes over vector addition.
- ✘ Associative: $\boldsymbol{a}^T(\boldsymbol{b}^T \boldsymbol{c}) \neq (\boldsymbol{a}^T \boldsymbol{b})\boldsymbol{c}$ – in general the associative property does not hold, as the dot product would most likely produce different scalars.
  - Additionally, $\boldsymbol{a}$ could have a different dimensionality than $\boldsymbol{b}$ and $\boldsymbol{c}$. I.e., even if $\boldsymbol{b}$ and $\boldsymbol{c}$ had the same dimensionality ($\boldsymbol{a}^T(\boldsymbol{b}^T \boldsymbol{c})$ would be valid vector scalar multiplication) then $\boldsymbol{a}^T \boldsymbol{b}$ would be invalid.
- ✓ Commutative: $\boldsymbol{a}^T \boldsymbol{b} = \boldsymbol{b}^T \boldsymbol{a}$ – the order of the vectors does not matter.

Vector Length

- **Vector norm (magnitude, length)**: denoted with double vertical bars $\|\boldsymbol{v}\|$, indicating length of a vector in euclidean space. Not to be confused with absolute value $|x|$ of a scalar's "norm." However, sometimes the notation $|\boldsymbol{v}|$ is used.
- Calculating $\|\boldsymbol{v}\|$ is done using the Euclidean norm:

$$\|\boldsymbol{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}$$

  - This is a consequence of the Pythagorean theorem, since the basis vectors↓ $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3$ are orthogonal↓ unit vectors↓.
- Thus, the norm can easily be found by taking the square root of the dot product of the vector with itself:

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v}^T v}$$

Geometric Interpretation of the Dot Product

- The dot product of two Euclidean vectors↑ $\boldsymbol{a}$ and $\boldsymbol{b}$ is defined by:

$$\lambda = \boldsymbol{a}^T \boldsymbol{b} = \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| \cos(\theta)$$

  where θ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.
- **Features based on θ**:
  - When $\cos\theta > 0$ ($\theta < 90° -$ acute) then $\lambda > 0\ (+)$
  - When $\cos\theta < 0$ ($\theta > 90° -$ obtuse) then $\lambda < 0\ (-)$
  - When $\cos\theta = 0$ ($\theta = 90° -$ perpendicular) then $\lambda = 0$
    - This represents a special case where the vectors are said to be orthogonal.
    - **Orthogonality**: the generalization of the notion of perpendicularity to the linear algebra of bilinear forms.
  - When $\cos\theta = 1$ then the vectors are codirectional:

$$\boldsymbol{a}^T \boldsymbol{b} = \|\boldsymbol{a}\| \, \|\boldsymbol{b}\|$$

    - Thus, the dot product with a vector $\boldsymbol{v}$ with itself is

$$\boldsymbol{v}^T \boldsymbol{v} = \|\boldsymbol{v}\|^2$$

    - Which gives us the norm↑ as defined above, i.e., $\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v}^T v}$
    - If $\cos\theta = -1$, then really vectors are still codirectional, but point in opposite directions with respect to the origin.

Other Properties of Vectors

- Matrices↓ are properly defined later. However, vectors are technically single row or column matrices, so many of the following operations also work on vectors, thus use of matrices often appears in this section.

Vector Hadamard Multiplication

- **Hadamard (element-wise) product**: a binary operation (only takes two operands) that matrices of the same dimensions and produces another matrix of the same dimension as the operands, e.g., vector Hadamard multiplication:

$$\begin{bmatrix}1\\0\\4\\5\end{bmatrix} + \begin{bmatrix}2\\3\\-6\\11\end{bmatrix} = \begin{bmatrix}2\\0\\-24\\55\end{bmatrix}$$

Outer Product

- Recall that the dot (scalar, inner) product↑ produces a 1×1 matrix, or rather a scalar, hence "scalar" product.
  - Note: the typical notation used for dot products is $\boldsymbol{v}^T \boldsymbol{w}$ — part of reasoning for the "inner" product.
- **Outer product**: an $n \times m$ matrix that results from the product of two vectors with dimensions n and m.

$$\boldsymbol{v}\boldsymbol{w}^T = n \times m$$

- The subtle change in notation matters in contrast to the dot product, as both represent distinct operations (assuming they are column vectors).
- The outer product allows for the multiplication of vectors with different dimensionality
- Can be thought of in two different ways:
  - The row perspective:

$$\begin{bmatrix}1\\0\\4\\2\end{bmatrix} \begin{bmatrix}a & b & c\end{bmatrix} = \begin{bmatrix}1a & 1b & 1c\\0a & 0b & 0c\\4a & 4b & 4c\\2a & 2b & 2c\end{bmatrix}$$

  - The column perspective:

$$\begin{bmatrix}1\\0\\4\\2\end{bmatrix} \begin{bmatrix}a & b & c\end{bmatrix} = \begin{bmatrix}1a & 1b & 1c\\0a & 0b & 0c\\4a & 4b & 4c\\2a & 2b & 2c\end{bmatrix}$$

Cross Product

- **Cross (vector, directed area) product**: denoted by the symbol $\times$, indicating a binary operation on two vectors in three-dimensional space $\mathbb{R}^3$.
  - Given two linearly independent↓ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, then the cross product $\boldsymbol{a} \times \boldsymbol{b}$ produces a new vector that is orthogonal to both $\boldsymbol{a}$ and $\boldsymbol{b}$, or normal to the plane containing them.
  - The direction of the vector is given by the right-hand rule ($\boldsymbol{a}$ = pointer, $\boldsymbol{b}$ = index, thumb = direction).
  - The magnitude of the vector represents the area of the parallelogram that the vectors span.
- The cross product can be defined by the formula:

$$\boldsymbol{a} \times \boldsymbol{b} = \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| \sin(\theta) \, \boldsymbol{v}$$

  - Notice this is similar to the geometric interpretations of the dot product↑ — the contrast between the two leads to an intuitive interpretation:
    - $\cos(\theta)$ in the dot product is used to measure how "parallel" the two vectors are, i.e., they are codirectional when $\theta = 1$, allowing for calculation of the norm↑.
    - $\sin(\theta)$ in the cross product is used to measure how "perpendicular" two vectors are, i.e., they are orthogonal when $\theta = 1$. There are multiple directions of the orthogonal vector, so calculation of the signed area returns vector $\boldsymbol{v}$ that describes both magnitude and direction as described above.
    - The intuition described here will be more clear when the determinant↓ is discussed in more detail.
- An algebraic example of vector e.g.,

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \times \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2c & - & 3b \\ 3a & - & 1c \\ 1b & - & 2a \end{bmatrix}$$

Primer on Complex Numbers

- **Complex number**: a number that can be expressed in the form $a + bi$, where a and b are real numbers, and i (j) is a symbol called the imaginary unit that satisfies the equation $i = \sqrt{-1}$.
- This imaginary unit allows for the imaginary set $(a + bi \in \mathbb{C})$ to be described.
- The combination of a real part and an imaginary part ($a + bi$) gives imaginary numbers both a direction and magnitude on the 2-D plane created between the two parts—thus they can be described as 2-D vectors in this space.

- Multiplication of complex numbers can be done by factoring the imaginary unit, e.g.,

$$\begin{aligned} z &= a + bj \\ w &= c + dj \end{aligned}$$

$$\begin{aligned} zw &= (a + bj)(c + dj) \\ &= ac + adj + cbj + bdj^2 \\ &= ac + adj + cbj - bd && (j^2 = -1) \end{aligned}$$

- Computing the dot product with complex vectors is the same, just including the factoring mentioned above when necessary, e.g.,

$$\begin{aligned} &\begin{bmatrix} 1 + 3j \\ -2j \\ 4 \\ 5 \end{bmatrix}^T \begin{bmatrix} 6 + 2j \\ 8 \\ 3j \\ -5 \end{bmatrix} \\ &= (1 + 3j)(6 + 2j) + -16j + 12j + 25 \\ &= 6 + 2j + 18j - 6 - 16j + 12j + 25 \\ &= 25 + 16j \end{aligned}$$

Conjugate Transpose

- **Conjugate (Hermitian) transpose** M^H, M^*: the n-by-m matrix obtained by taking the transpose↓ and then taking the complex conjugate of each entry.
- **Complex conjugate**: the number with both equal real and imaginary parts equal in magnitude but opposite in sign.

$$a + bi = a - bi \quad \longleftrightarrow \quad a - bi = a + bi$$

- Multiplying a vector by the conjugate transpose allows us to calculate the magnitude of the vector containing imaginary numbers by using the complex conjugate to "canceling out" all the imaginary units and give a real number answer (or rather, $a + 0i$).
- Matrices will return a matrix of real numbers—this is part of what makes using imaginary numbers useful for future computations.

Unit Vectors

- **Unit vectors** $\hat{\imath}, \hat{\jmath}, \hat{k}$: a vector scaled such that $(:)$ the length of the vector equals 1 in a normed vector space.

$$\hat{\imath} = \lambda \boldsymbol{v} : \|\lambda \boldsymbol{v}\| = \frac{1}{\|\boldsymbol{v}\|}\|\boldsymbol{v}\| = 1$$

- **Normed vector space**: a vector space, over the real or complex numbers, on which a norm↑ is defined.

Linear Combinations

- **Linear combination**: an expression constructed from a set of terms by multiplying each term by a scalar and adding the results, i.e.,

$$a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + a_3\boldsymbol{v}_3 + a_n\boldsymbol{v}_n$$

- For example, in a 3-D vector space $\mathbb{R}^3$, then any vector in the space is can be made by a linear combination of the following three vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ (unit vectors) multiplied by some scalar:

$$\lambda\begin{bmatrix}1\\0\\0\end{bmatrix} + \lambda\begin{bmatrix}0\\1\\0\end{bmatrix} + \lambda\begin{bmatrix}0\\0\\1\end{bmatrix}$$

Subspace

- **Linear (vector) subspace**: a vector space that is a subset of some larger vector space.
- Algebraically, a subspace is the set of all vectors that can be created by taking a linear combination of some vector or a set of vectors.
- A vector subspace must be closed under addition and scalar multiplication while also containing the zero vector, i.e.,

$$\forall \boldsymbol{v}, \boldsymbol{w} \in L; \quad \forall \lambda, \alpha \in \mathbb{R}; \quad \lambda\boldsymbol{v} + \alpha\boldsymbol{w} \in L$$

  - For all ($\forall$) vectors $\boldsymbol{v}$, $\boldsymbol{w}$ in ($\in$) the linear subspace (V), and for all scalars λ, α in the set of real numbers ($\mathbb{R}$), then any linear combination of the two are in the same subspace.
  - The above equation also implies the inclusion of zero vector, which is a trivial subspace of the vector space.
- The geometric interpretation is best described through some examples:
  - A $\mathbb{R}^1$ vector and all the scaled possibilities added together describes a line stretches infinity in both directions, while a $\mathbb{R}^2$ vector would create a 2-D plane.
  - Both of the previous subspaces exist in the higher dimensional vector spaces, i.e., a 2-D plane and a 1-D line both exist in the 3-D vector space.
  - All subspaces also pass through the origin, including the O-D subspace, which is just the origin.

Subsets

- **Subset** $\subseteq$: a set A is a subset of a set B if all elements of A are also elements of B; B is then b a **superset** $\supseteq$ of A.
- Not all subsets of vector spaces are subsets; subsets don't need to include the origin, don't need to be closed, or could have boundaries.

Span

- **Linear span (hull)** $\text{span}(S)$: a set S of vectors (from a vector space) that is the smallest linear subspace that contains the set.
- The span is typically infinite, but it also can be defined as the set of all finite linear combinations↑ of vectors of S given an arbitrary field K:

$$\text{span}(S) = \left\{ \sum_{i=i}^{k} \lambda_i \boldsymbol{v}_i \,\middle|\, k \in \mathbb{N}, \boldsymbol{v}_i \in S, \lambda_i \in K \right\}$$

- A frequent question is asked whether a vector is in a span or not, e.g., are vectors $v, w \in \text{span}(S)$:

$$\boldsymbol{v} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} w = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 7 \\ 0 \end{bmatrix} \right\}$$

- $\boldsymbol{v} \in \text{span}(S)$ since both vectors in S can be scaled then combined in some way to form $\boldsymbol{v}$, i.e.,

$$\boldsymbol{v} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \frac{5}{6} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{6} \begin{bmatrix} 1 \\ 7 \\ 0 \end{bmatrix}$$

  - Determining the weights will be discussed later using matrix computations↓.
- However, it's clear that $\boldsymbol{w} \notin \text{span}(S)$ since 0 cannot be scaled to equal 1.
- Geometrically, a useful intuitive example is a 2-D span in 3-D space; the 2-D plane can be moved anywhere in the 3-D space as long as the two vectors describing the 2-D plane are linearly independent↓.

Linear Independence

- **Linearly dependent**: a set of V vectors where at least one vector in the set can be defined as a linear combination↑ of the others.
- **Linearly independent**: when no vector in the set can be written in the above way.
- Algebraically, a sequence of vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ that a vector space V are linearly independent if there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$ (not all zero) such that they form the zero vector 0, i.e.,

$$\lambda_1 \boldsymbol{v}_1, \lambda_2 \boldsymbol{v}_2, \ldots, \lambda_k \boldsymbol{v}_k = 0 \qquad \lambda \in \mathbb{R}$$

- Geometrically, a set of V vectors in independent if each vectors points in a geometric dimension not reachable using other vectors in the set.
  - E.g., if two vectors are lie along the same line, then they are really just the same vector; same concept with a 2-D plane containing three vectors in $\mathbb{R}^3$.

Basis

- **Basis**: the combination of span↑ and linear independence↑, i.e., the set B of vectors in vector space V if every element of V may be written as a unique finite linear combination↑ of elements of B.
  - More concisely, a basis is simply a linearly independent spanning set.
  - **Components (coordinates)**: the coefficients of the linear combination and are referred to as of the vector with respect to B, or **basis vectors**.
- Examples of standard basis vectors:

$$\mathbb{R}^2 \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \qquad \mathbb{R}^3 \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

- The standard basis vectors of $\mathbb{R}^3$ was used in the example of linear combinations↑ to demonstrate how these unit vectors could all be scaled so that they could describe any vector in the space.
- (fix link) so it's better to think of them as the "rulers" used to describe any other vector in the space.

Matrices

Matrix Terminology

- **Matrix** $M_{r,c}$: a rectangular array of elements arranged in rows $\leftrightarrow$ and columns $\updownarrow$, e.g.,

$$M = \begin{bmatrix} 1 & 0 & 3 \\ 5 & 4 & 2 \\ 7 & 6 & 9 \end{bmatrix} \qquad M_{3,2} = 6$$

- **Block (partitioned) matrix**: a matrix that is interpreted as having been broken into sections called blocks or submatrices, e.g.,

$$M = \begin{bmatrix} D & N \\ Y & D \end{bmatrix} = \begin{bmatrix} 4 & 2 & 0 & 0 \\ 6 & 9 & 0 & 0 \\ 1 & 1 & 4 & 2 \\ 1 & 1 & 6 & 9 \end{bmatrix}$$

$$D = \begin{bmatrix} 4 & 2 \\ 6 & 9 \end{bmatrix} \quad N = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad Y = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

  - Can be used for large matrices with high level structure, offering convenient notation, and sometimes providing computational benefits.
- **Diagonal**: the elements of matrix starting from the top left $\searrow$ lower right.
  - **Off-diagonal**: elements not along the diagonal (0s and 1s in example below)
  - Works for both square and rectangular matrices↓, e.g.,

$$\begin{bmatrix} 4 & 1 & 0 & 1 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 6 & 0 \\ 1 & 1 & 0 & 9 \end{bmatrix} \qquad \begin{bmatrix} 4 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 6 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 9 & 1 & 0 & 1 \end{bmatrix}$$

- **Matrix size**: a matrix with m rows and n columns is called an $m \times n$ matrix, or m-by-n matrix, while m and n are the dimensions.
  - Order matters—the convention is rows then columns, i.e., $m \times n \neq n \times m$.
  - **MR**. **NiC**e guy: a useful mnemonic to remember typical conventions.
- **Matrix dimensionality**:
  - $\mathbb{R}^{mn}$: describes the total number of elements, here the multiplication of the dimensions is commutative (order doesn't matter).
  - $\mathbb{R}^{m \times n}$: the specific matrix size using rows and columns as described above.
  - $C(M) \in \mathbb{R}^m$: a collection of column vectors, i.e., a matrix spanned by set column vectors with m elements.
  - $R(M) \in \mathbb{R}^n$: a collection of row vectors, the inverse of above.

Square and Rectangular Matrices

- **Square matrix**: a matrix with the same number of rows and columns.
  - An $n \times n$ matrix is known as a square matrix of order n.
  - Any two square matrices of the same order can be added and multiplied.
- **Rectangular matrix**: a matrix with an unequal number of rows and columns, i.e., $m \neq n$.
- Both square and rectangular matrices have a diagonal↑, as described above.

Symmetric and Skew-Symmetric Matrices

- **Symmetric matrix**: a square matrix that can be mirrored across the diagonal, e.g.,

$$\begin{bmatrix} 4 & -6 & -1 \\ -6 & 2 & 9 \\ -1 & 9 & 0 \end{bmatrix}$$

  - Algebraically, it's a square matrix $\boldsymbol{A}$ that is equal to its transpose↓, i.e., $\boldsymbol{A} = \boldsymbol{A}^T$.
  - It does not matter what is on the diagonal, as any number is equal to itself.
- **Skew-symmetric matrix**: a square matrix that is still symmetric, but all elements mirrored across the diagonal and inverted, e.g.,

$$\begin{bmatrix} 0 & +6 & +1 \\ -6 & 0 & -9 \\ -1 & 9 & 0 \end{bmatrix}$$

  - Algebraically, it's a square matrix $\boldsymbol{A}$ that is equal to its negative transpose, i.e., $\boldsymbol{A} = -\boldsymbol{A}^T$
  - Here all elements on the diagonal must be zero, as zero is the only number that can be equal its inverse.

Identity and Zero Matrices

- **Identity matrix $\boldsymbol{I}_n$**: a matrix with size $n \times n$ with all elements along the diagonal = 1 and all other elements = 0, e.g, $\boldsymbol{I}_3$ and $\boldsymbol{I}_n$ ($\cdots$ indicate continuation of pattern):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

  - Essentially, the identity matrix is the equivalent of the number 1 in linear algebra.
- **Zero matrix $\mathbf{0}$**: a matrix of all zeros.

Diagonal and Triangular Matrices

- **Diagonal matrix**: when all elements outside the main diagonal are zero, i.e.,

$$\begin{bmatrix} e_{1,1} & 0 & \cdots & 0 \\ 0 & e_{2,2} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & e_{i,i} \end{bmatrix}$$

  - Elements along the diagonal don't have to be the same, and they can be zero (meaning rectangular matrices still can be diagonal, technically).
  - When all the elements along the main diagonal are the same, then it's a scaled version of the identity matrix↑, i.e., $\lambda \boldsymbol{I}$.
- **Triangular matrices**: when all elements above or below the diagonal are zero, but not on both sides.
  - Upper triangular matrix: when all the elements below the diagonal are zero.
  - Lower triangular matrix: when all the elements above the diagonal are zero.

$$\text{Upper} = \begin{bmatrix} e_{1,1} & e_{1,2} & e_{1,3} \\ 0 & e_{2,2} & e_{2,3} \\ 0 & 0 & e_{3,3} \end{bmatrix} \qquad \text{Lower} = \begin{bmatrix} e_{1,1} & 0 & 0 \\ e_{2,1} & e_{2,2} & 0 \\ e_{3,1} & e_{3,2} & e_{3,3} \end{bmatrix}$$

Augmented and Complex Matrices

- **Augmented (concatenated) matrix $A \mid B$**: a matrix obtained by appending the columns of two given matrices, e.g.,

$$\begin{bmatrix} 4 & 2 & 0 \\ 3 & 7 & 6 \\ 1 & 6 & 9 \end{bmatrix} \Bigg| \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix} = \left[\begin{array}{ccc|c} 4 & 2 & 0 & 4 \\ 3 & 7 & 6 & 2 \\ 1 & 6 & 9 & 0 \end{array}\right]$$

  - Used typically for the purpose of performing the same elementary row operations on each of the given matrices.
  - Matrices must have the same number of rows (or columns for vertical augmentation) for the concatenation to be applied.
- **Complex matrix**: A matrix whose elements may contain complex numbers.
- The conjugate transpose↑ discussed previously in vectors can is used here as well.
- Transposition↓ will be discussed shortly, but for now, the complex conjugate still behaves the same (just imaginary numbers change sign), e.g.,

$$\begin{bmatrix} 1 & -1+5i & 0 \\ 1 & -2 & -4 \\ 6i & -4 & 5-2i \end{bmatrix}^H = \begin{bmatrix} 1 & 1 & -6i \\ -1-5i & -2 & -4 \\ 0 & -4 & 5+2i \end{bmatrix}$$

Basic Matrix Operations

Matrix Addition and Subtraction

- **Matrix addition (subtraction)**: the operation of adding (subtracting) two matrices of equal dimensions $m \times n$ by adding (subtracting) the corresponding elements together, e.g.,

$$\begin{bmatrix} 1 & 2 & 5 \\ 0 & 6 & 8 \\ 9 & 6 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 3 & 5 \\ 1 & -6 & 9 \\ -5 & -4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 10 \\ 1 & 0 & 17 \\ 4 & 2 & 4 \end{bmatrix}$$

- Note: there are other operations which could also be considered addition for matrices, such as the direct sum and the Kronecker sum (not discussed as of now).
- ✓ Commutative: $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$
- ✓ Associative: $\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C})$

Matrix Scalar Multiplication

- **Matrix scalar multiplication**: the same as vector scalar multiplication↑ or simply, scalar multiplication, as vectors are $m \times 1$ (or $1 \times n$) matrices.
- Scalar multiplication is true when both left scalar and right scalar are equal, i.e.,

$$\lambda(\boldsymbol{M})_{ij} = (\lambda \boldsymbol{M})_{ij} = (\boldsymbol{M}\lambda)_{ij} = (\boldsymbol{M})_{ij}\lambda$$

- More explicitly:

$$\begin{bmatrix} \lambda e & \lambda e & \cdots & \lambda e \\ \lambda e & \lambda e & \cdots & \lambda e \\ \vdots & \vdots & \ddots & \vdots \\ \lambda e & \lambda e & \cdots & \lambda e \end{bmatrix} = \left(\lambda \text{ or } \begin{bmatrix} e & e & \cdots & e \\ e & e & \cdots & e \\ \vdots & \vdots & \ddots & \vdots \\ e & e & \cdots & e \end{bmatrix} \text{ or } \lambda \right) = \begin{bmatrix} e\lambda & e\lambda & \cdots & e\lambda \\ e\lambda & e\lambda & \cdots & e\lambda \\ \vdots & \vdots & \ddots & \vdots \\ e\lambda & e\lambda & \cdots & e\lambda \end{bmatrix}$$

- The above is true only where the underlying ring (algebraic structure that generalize fields...I need to learn more about this...) is commutative. This fact is essential for later proofs.

Transposition

- **Transpose** ^{T}: an operation where a matrix is flipped over its diagonal↑, i.e., it switches the row and column indices of the matrix, e.g.,

$$\begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 0 \\ 3 & 7 & 1 \\ 4 & 8 & 2 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 0 & 1 & 2 \end{bmatrix}$$

- Formally, the element of the i-th row, j-th column of matrix $\boldsymbol{M}$ when transposed becomes the element of the j-th row, i-th column of matrix $\boldsymbol{M}^T$, i.e.,

$$\boldsymbol{M}_{i,j} = \boldsymbol{M}^T_{j,i}$$

- Alternatively, with regard to dimensionality, if $\boldsymbol{M}$ is an $m \times n$ matrix, then $\boldsymbol{M}^T$ is an $n \times m$ matrix.
  - Thus, a transposed matrix that is transposed again will produce the original matrix, i.e.,

$$\left(\boldsymbol{M}^T_{j,i}\right)^T = \boldsymbol{M}_{i,j}$$

- Revisiting complex matrices↑:
  - **Hermitian matrix**: a square complex matrix whose transpose is equal to every entry being replaced with its complex conjugate↑.
    - Denoted: $\boldsymbol{M}^T = \overline{\boldsymbol{M}}$
  - **Skew-Hermitian matrix**: a Hermitian matrix whose transpose is equal to the negation of its complex conjugate.
    - Denoted: $\boldsymbol{M}^T = -\overline{\boldsymbol{M}}$

Diagonal and Trace

- The main diagonal↑ of a matrix can be extracted and turned into a vector.
  - Not to be confused with diagonalization↓ of a matrix, which is a result of matrix decomposition↓ resulting from eigendecomposition↓.
- **Trace** $\text{tr}(\boldsymbol{M})$: the sum of all diagonal elements, defined only for square matrices.

$$\text{tr}(\boldsymbol{M}) = \sum_{i=1}^{n} e_{i,i} = e_{1,1} + e_{2,2} + \cdots + e_{n,n}$$

- The trace is a linear mapping (two vector spaces that preserves the operations of vector addition and scalar multiplication.), i.e.,

$$\text{tr}(\boldsymbol{A} + \boldsymbol{B}) = \text{tr}(\boldsymbol{A}) + \text{tr}(\boldsymbol{B}) \qquad \text{tr}(\lambda \boldsymbol{A}) = \lambda\,\text{tr}(\boldsymbol{A})$$

- Additionally, a matrix and its transpose have the same trace, as elements along the main diagonal are not affected, i.e., $\text{tr}(\boldsymbol{M}) = \text{tr}(\boldsymbol{M}^T)$

Broadcasting

- **Broadcasting**: duplication of a vector so that the dimensionality matches a larger matrix, allowing for simplification of element wise addition or multiplication.
  - Technically not a valid operation in linear algebra at face value, but is used commonly in applied linear algebra and machine learning.

Matrix Multiplication

- **Standard matrix multiplication** AB: a binary operation that produces a matrix from two matrices whose inner dimensions match.
  - Multiplication of matrices can be thought of as going from right to left ($\boldsymbol{A} \leftarrow \boldsymbol{B}$), or rather, $\boldsymbol{B}$ post-multiplies $\boldsymbol{A}$.
  - The number of columns (n) in the first matrix must be equal to the number of rows (m) in the second matrix.
  - **Matrix product**: the product, whose size is equal to rows of the first matrix and the number of columns of the second matrix.

$$(m_1 \times n_1)(m_2 \times n_2) = m_1 \times n_2$$

- Transposing↑ a matrix switches the dimensions, so it can enable computation in some cases, e.g,

$$\underbrace{\boldsymbol{A}}_{5\times 7} = \underbrace{\boldsymbol{A}^T}_{7\times 5} \qquad \underbrace{\boldsymbol{A}^T}_{7\times 5}\underbrace{\boldsymbol{B}}_{5\times 2} = \underbrace{\boldsymbol{C}}_{7\times 2}$$

- Revisiting the dot (inner) product↑ and the outer product↑ of vectors:
  - The dot product must have equal-length vectors since the transpose of left vector makes inner dimensions much match, e.g.,

$$\underbrace{\boldsymbol{v}}_{5\times 1}\underbrace{\boldsymbol{w}}_{5\times 1} \rightarrow \underbrace{\boldsymbol{v}^T}_{1\times 5}\underbrace{\boldsymbol{w}}_{5\times 1} = 1 \times 1 \text{ (scalar)}$$

  - However, the transpose of the right vector makes the 1×1 dimensions match, making the outer dimensions irrelevant to the validity of the computation, e.g.,

$$\underbrace{\boldsymbol{v}}_{6\times 1}\underbrace{\boldsymbol{w}}_{9\times 1} \rightarrow \underbrace{\boldsymbol{v}}_{6\times 1}\underbrace{\boldsymbol{w}^T}_{1\times 9} = 6 \times 9 \text{ (matrix)}$$

Standard Matrix Multiplication Perspectives

- There are four ways to compute and conceptualize the process of multiplication of two matrices, all of which give the same result.

1. **The element perspective**:
2. **The layer perspective**:
3. **The column perspective**:
4. **The row perspective**:

Diagonal Matrix Multiplication

-

Order of Operations

-

Matrix Vector Multiplication

-

Transformation Matrix

-

Additive and Multiplicative Matrices

-

Hadamard Multiplication

-

Multiplication of Symmetric Matrices

-

Frobenius Dot Product

-

Matrix Norms

-

Matrix Rank

-

Matrix Spaces

-

The Determinant

-

Matrix Inverse

-

To Be Defined

- Diagonalization
- Matrix decomposition
- Eigendecomposition