Linear Algebra Dictionary

by Andrew Lehmann
Last updated: June 21, 2024
Typeset in $\LaTeX\,2_\varepsilon$.

List of Definitions

Chapter 1: Euclidean Vectors

Definition 1.1 (Euclidean vector or tuple).
A Euclidean vector is a list of n real numbers, also called an n-tuple. We write this list in parentheses, for example $(1, 3, -2, \dots, 0)$, and we say that this object belongs to $\mathbb{R}^n$. An arbitrary tuple can be written $\mathbf{v} = (v_1, v_2, \cdots, v_n)$ where the components $v_i \in \mathbb{R}$ for any index i.

Definition 1.2 (Tuple addition).
Euclidean vectors are added to each other component by component. In symbols

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n).$$

Note: this means you can only add two tuples together of the same size. It makes no sense to add a 3-tuple to a 5-tuple.

Definition 1.3 (Scalar multiplication).
Let $c \in \mathbb{R}$, called a scalar quantity, and $\mathbf{v} \in \mathbb{R}^n$ with components v_i. Then the scalar multiplication $c\mathbf{v}$ gives a vector $\mathbf{w}$ with components $w_i = cv_i$ for every index i. In tuple form

$$c(v_1, v_2, \dots, v_n) = (cv_1, cv_2, \dots, cv_n).$$

Definition 1.4 (Canonical Euclidean unit vectors).
The canonical Euclidean vectors in $\mathbb{R}^n$ are the n vectors of the form

$$\begin{aligned}\mathbf{e}_1 &= (1, 0, \dots, 0)\\ \mathbf{e}_2 &= (0, 1, \dots, 0)\\ &\vdots\\ \mathbf{e}_n &= (0, 0, \dots, 1).\end{aligned}$$

More compactly

$$\mathbf{e}_k = (\alpha_1, \alpha_2, \dots, \alpha_n) \quad \textit{where} \quad \alpha_j = \begin{cases} 1 & \textit{for } j = k, \\ 0 & \textit{for } j \neq k. \end{cases}$$

Definition 1.5 (Dot product).
For two n-tuples $\mathbf{a}$ and $\mathbf{b}$, their dot product, also called scalar product and Euclidean inner product, is the real number given by the addition of component by component multiplication

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i.$$

If the matrix does not satisfy any of these (e.g. if we can find a positive and a negative quadratic form) then the matrix is called indefinite.

Definition 8.6 (Eigenvalues of a definite matrix).
Let A be an $n \times n$ symmetric real matrix. Then all eigenvalues of A are real numbers. Furthermore, A is

- *positive definite if an only if every eigenvalue is strictly positive,*
- *positive semi-definite if an only if every eigenvalue is non-negative,*
- *negative definite if an only if every eigenvalue is strictly negative,*
- *negative semi-definite if an only if every eigenvalue is strictly non-positive.*

Chapter 8: Orthogonal matrices

Definition 8.1 (Orthogonal matrix).
A square matrix A is orthogonal if and only if its inverse is its transpose. That is, if and only if $AA^T = A^TA = I$.

Definition 8.2 (Properties of orthogonal matrices).
If A is an orthogonal matrix of size n, then

- *its columns are pair-wise orthogonal,*
- *its columns are unit length,*
- *its columns (considered as n-tuples) form an orthonormal basis of $\mathbb{R}^n$,*
- *its rows (considered as n-tuples) form an orthonormal basis of $\mathbb{R}^n$,*
- *it has determinant ± 1.*

Definition 8.3 (Orthogonally diagonalizable matrix).
A square matrix A is orthogonally diagonalizable if there exists a diagonal matrix D and orthogonal matrix Q such that

$$A = QDQ^T.$$

Definition 8.4 (Quadratic form).
Let A be an $n \times n$ matrix and $\mathbf{v} \in \mathbf{R}^n$ considered as a column. Then a quadratic form is a multiplication of the form $\mathbf{v}^T A\mathbf{v}$ resulting in a real number.

Definition 8.5 (Definite matrix).
Let A be an $n \times n$ symmetric real matrix. By considering the sign of quadratic forms with A we can define several cases. A is

- *positive definite if an only if $\mathbf{v}^T A\mathbf{v} > 0$ for every $\mathbf{v} \in \mathbb{R}^n$,*
- *positive semi-definite if an only if $\mathbf{v}^T A\mathbf{v} \geq 0$ for every $\mathbf{v} \in \mathbb{R}^n$,*
- *negative definite if an only if $\mathbf{v}^T A\mathbf{v} < 0$ for every $\mathbf{v} \in \mathbb{R}^n$,*
- *negative semi-definite if an only if $\mathbf{v}^T A\mathbf{v} \leq 0$ for every $\mathbf{v} \in \mathbb{R}^n$.*

Definition 1.6 (Euclidean Norm).
The norm of an n-tuple $\mathbf{v}$, denoted $\|\mathbf{v}\|$, is given by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_2^2}.$$

Definition 1.7 (Orthogonal Euclidean vectors).
Two vectors in $\mathbb{R}^n$ are orthogonal if and only if their dot product equals zero.

Definition 1.8 (Displacement vector).
Given two Euclidean vectors $\mathbf{a}$ and $\mathbf{b}$, the displacement vector pointing from $\mathbf{a}$ to $\mathbf{b}$ is given by $\mathbf{r} = \mathbf{b} - \mathbf{a}$ as pictured below. Of course we can also create the displacement vector in the other direction, from $\mathbf{b}$ to $\mathbf{a}$, given by $\mathbf{a} - \mathbf{b}$.

Definition 1.9 (Vector form of a straight line).
The set of vectors in $\mathbb{R}^n$ of the form $\mathbf{v} = \mathbf{a} + t\mathbf{r}$ for a parameter $t \in \mathbb{R}$ represents a straight line through the space $\mathbb{R}^n$. That is,

$$\{(x, y) \mid \forall x \in \mathbb{R} \; and \, y = mx + b\} = \{\mathbf{a} + t\mathbf{r} \mid \forall t \in \mathbb{R}\}$$

where $\mathbf{a}$ is an arbitrary pair $(x, mx+b)$ and $\mathbf{r}$ is a displacement vector between any two distinct pairs $(x_1, mx_1 + b)$ and $(x_2, mx_2 + b)$.

Chapter 2: Matrix Algebra

Definition 2.1 (Matrix).
A matrix is a collection of numbers from a field $\mathbb{F}$ (e.g. rational numbers) usually represented by a rectangular array. For example, an $m \times n$ (said m by n) matrix A with coefficients $a_{ij} \in \mathbb{F}$ would be represented by an array with m rows and n columns:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2j} & \cdots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3j} & \cdots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & a_{12} & a_{13} & \cdots & a_{ij} & \cdots & a_{im} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nj} & \cdots & a_{nm} \end{pmatrix} = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

Sometimes it is convenient to refer to the coefficients in the array like so: $a_{ij} = (A)_{ij}$.

Definition 2.2 (Set of all $m \times n$ matrices).
We write the set of all $m \times n$ matrices with coefficients in $\mathbb{F}$ as

$$\mathcal{M}_{m,n}(\mathbb{F})$$

Definition 2.3 (Matrix columns and rows).
For a matrix $A \in \mathcal{M}_{m,n}(\mathbb{F})$ we denote its j^{th} column and i^{th} row

$$A^{(j)} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{ij} \\ \vdots \\ a_{mj} \end{pmatrix}, \qquad A_{(i)} = \begin{pmatrix} a_{i1} & a_{12} & a_{13} & \cdots & a_{ij} & \cdots & a_{in} \end{pmatrix}$$

Definition 2.4 (Transpose of a matrix).
The transpose of an $m \times n$ matrix, A, is an $n \times m$ matrix, denoted A^T, with rows equal to the columns of A. That is, $\left(A^T\right)_{ij} = (A)_{ji}$ for all combinations of i and j.

Definition 7.6 (Orthogonal vectors).
Two vectors $\mathbf{u}$ and $\mathbf{v}$ of an inner product space $(V, \langle,\rangle)$ are orthogonal if and only if their inner product is zero: $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Definition 7.7 (Norm).
A vector $\mathbf{v}$ in an inner product space $(V, \langle,\rangle)$ has norm

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}.$$

The Euclidean norm is therefore

$$\|(v_1, \ldots, v_n)\| = \sqrt{v_1^2 + \cdots + v_n^2}.$$

If a vector has norm equal to 1 we say it is a unit vector or has unit length. If we divide a vector by its norm we say that is has been normalised.

Definition 7.8 (To normalise a vector).
Consider a vector $\mathbf{v}$ in an inner product space $(V, \langle,\rangle)$. We say we "normalise" this vector by dividing it by its norm. That is, $\mathbf{v}'$ is the normalised $\mathbf{v}$ if

$$\mathbf{v}' = \frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

When we normalise a vector we guarantee that it has length 1:

$$\|\frac{\mathbf{v}}{\|\mathbf{v}\|}\| = \frac{\|\mathbf{v}\|}{\|\mathbf{v}\|} = 1$$

Definition 7.9 (Orthonormal basis).
An orthonormal basis of an inner product space $(V, \langle,\rangle)$ is a set of vectors $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ each having norm of 1 and that are pairwise orthogonal:

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & \textit{whenever } i = j, \\ 0 & \textit{whenever } i \neq j. \end{cases}$$

Chapter 7: Inner product spaces

Definition 7.1 (Inner product).
An inner product is a mapping that takes any two vectors of a vector space, V, to a scalar, $f : V \times V \to \mathbb{R}$ but often denoted with angle brackets $f(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u}, \mathbf{v} \rangle$, satisfying the following properties:

$$\forall\, \mathbf{u},\, \mathbf{v},\, \mathbf{w} \in V \quad \text{and} \quad \forall\, k \in \mathbb{R}$$

$$\begin{array}{lll} (IP1) & \langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle & (\textit{commutativity}) \\ (IP2) & \langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle & (\textit{linearity over vector addition}) \\ (IP3) & \langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle & (\textit{linearity over scalar multiplication}) \\ (IP4) & \langle \mathbf{u}, \mathbf{u} \rangle \geq 0 & (\textit{positive definite}) \end{array}$$

Definition 7.2 (Euclidean dot product).
The Euclidean dot product is the canonical inner product defined on the vector space of real n-tuples, $\mathbb{R}^n$. Given two vectors $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$, their dot product is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i.$$

Definition 7.3 (Inner product of functions).
Let $\mathcal{C}([a, b])$ be the vector space of real functions that are continuous on the interval $[a, b]$. We can define an inner product on any functions $f, g \in \mathcal{C}([a, b])$

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx.$$

You should verify that this definition satisfies the 4 properties of inner products.

Definition 7.4 (Inner product space).
An inner product space is a vector space and a definition of an inner product considered as a pair $(V, \langle , \rangle)$. We say that V is equipped with the inner product.

Definition 7.5 (Euclidean inner product space).
A Euclidean inner product space is the vector space of real n-tuples equipped with the euclidean dot product: $(\mathbb{R}^n, \cdot)$.

Definition 2.5 (Diagonal matrix).
A square matrix A is said to be diagonal if all its non-diagonal elements are zero, e.g. $(A)_{ij} = 0$ whenever $i \neq j$.

Definition 2.6 (Symmetric matrix).
A matrix A is symmetric if it is equal to its transpose, $A = A^T$.

Definition 2.7 (Matrix addition).
Matrix addition is done coefficient by coefficient, that is, for two matrices A and B we define the i,j^{th} coefficient of the addition as the addition of the i,j^{th} coefficients of each matrix:

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij}\,.$$

Definition 2.8 (Scalar multiplication).
Given a number $k \in \mathbb{R}$ (called a scalar) and a matrix $A \in \mathcal{M}_{m,n}$, we define matrix scalar multiplication, kA, to be a matrix $B \in \mathcal{M}_{m,n}$ with coefficients given by:

$$b_{ij} = k a_{ij},$$

that is, we multiply every coefficient by the scalar.

Definition 2.9 (Zero matrix).
The zero matrix of any shape is a matrix $M_0 \in \mathcal{M}_{m,n}$ consisting entirely of zeros as coefficients.

Definition 2.10 (Additive inverse).
Given a matrix $A \in \mathcal{M}_{ij}$, its additive inverse is the same matrix multiplied by the scalar -1. We denote the additive inverse of A as $-A$.

Definition 2.11 (Multiplication of a matrix by a column).
Consider a matrix $A \in \mathcal{M}_{m,n}$ and a column $X \in \mathcal{M}_{n,1}$. We define the product AX to result in the column $Y \in \mathcal{M}_{m,1}$ with coefficients

$$(Y)_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{im}x_m = \sum_{k=1}^{m} a_{ik} x_k$$

Visually

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \cdots + x_m \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix}$$

$$\implies Y = x_1 A^{(1)} + x_2 A^{(2)} + \cdots + x_m A^{(m)}$$

Definition 2.12 (Rotation matrix - arbtirary angle anti-clockwise)**.**
By using a column $X \in \mathcal{M}_{2,1}$ to represent a Euclidean vector, the following matrix allows the operation of rotataion, anti-clockwise, of X by an angle θ:

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

where the rotated vector is represented by a column $X' \in \mathcal{M}_{2,1}$ obtained by matrix multiplication $X' = R_\theta X$.

Definition 2.13 (Multiplication of two matrices)**.**
Consider two matrices $A \in \mathcal{M}_{n,m}$ and $B \in \mathcal{M}_{m,q}$. We define the product AB to be the matrix $C \in \mathcal{M}_{n,q}$ with coefficients

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} = \sum_{k=1}^{m} a_{ik}b_{kj}$$

$$\implies \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mq} \end{pmatrix}$$

$$= \left(\underbrace{b_{11} \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + \cdots + b_{m1} \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix}}_{\textit{first column}} \quad \cdots \quad \underbrace{b_{1q} \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + \cdots + b_{mq} \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix}}_{n^{th} \ \textit{column}} \right)$$

Additionally, for the product

$$\overbrace{(n \,.\, m)}^{A}\overbrace{(m \,.\, q)}^{B}$$

we will call the indices for the columns of A and rows of B the inner indices (blue), whereas the indices for the rows of A and columns of B will be called the outer indices (red).

Definition 6.10 (Diagonalizable matrix)**.**
A square matrix A is diagonalizable if and only if there exists an invertible matrix P and diagonal matrix D such that

$$A = PDP^{-1}.$$

Alternative: A square matrix A is diagonalizable if and only if it is similar to a diagonal matrix D.

Definition 6.11 (Eigenvalue diagonalization)**.**
For a diagonalizable matrix A of size n, we can sometimes find a diagonal matrix consisting of the eigenvalues of A, $\lambda_1, \ldots, \lambda_n$. In this case we can write

$$A = PDP^{-1}$$

where P consists of eigenvectors of A as columns. The matrix P is the transition matrix from the eigenbasis, $\mathcal{E}$, to the canonical basis of $\mathbb{R}^n$: $P_{\mathcal{E}\to\mathcal{C}_n}$.

eigenvectors corresponding to λ_k. This can be written as the set of linear combinations of linearly independent eigenvectors corresponding to λ_k:

$$E_{\lambda_k} = \{\alpha_1\mathbf{v}_1 + \cdots + \alpha_m\mathbf{v}_m \,|\, \forall j\; A\mathbf{v}_j = \lambda_k\mathbf{v}_j,\; \alpha_j \in \mathbb{R}\} = SPAN(\mathbf{v}_1, \cdots, \mathbf{v}_m)$$
$$\text{for maximum number of eigenvectors such that}$$
$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_m\mathbf{v}_m = \mathbf{0} \implies \alpha_1 = \alpha_2 = \cdots = \alpha_m = 0.$$

As the set $\{\mathbf{v}_1, \cdots, \mathbf{v}_m\}$ generates E_{λ_k} and the vectors are linearly independent, the set forms a basis and therefore gives the dimension E_{λ_k}.

The eigenspace can also be written like a kernel

$$E_{\lambda_k} = \{\mathbf{v} \in \mathbb{R}^n \,|\, (A - \lambda_k I)\,\mathbf{v} = \mathbf{0}\}.$$

Definition 6.6 (Algebraic and geometric multiplicity of an eigenvalue).
For an $n \times n$ matrix with characteristic polynomial

$$P(\lambda) = C(\lambda - \lambda_1)^{m_1} \times \cdot \times (\lambda - \lambda_k)^{m_k} \times \cdot \times (\lambda - \lambda_p)^{m_p}$$

for some constant C. There can be up to n distinct eigenvalues ($p \leq n$). The exponent m_k is called the algebraic multiplicity of the eigenvalue λ_k. The dimension of the eigenspace corresponding to λ_k is its geometric multiplicity.

Definition 6.7 (Eigenbasis).
Consider a square matrix A of size n. If the dimensions of its eigenspaces add up to n, then there exist n linearly independent eigenvectors of A. These eigenvectors form a basis of $\mathbb{R}^n$ called an eigenbasis.

Definition 6.8 (Similar matrices).
Two matrices A and B are similar if there exists an invertible matrix P such that

$$B = PAP^{-1}.$$

Definition 6.9 (Diagonalizable linear map).
Let $f : V \to V$ be an endomorphism. f is called diagonalizable if there exists a basis, $\mathcal{B}$, of V such that the matrix representation of f in $\mathcal{B}$ is diagonal:

$$(\mathcal{M}(f, \mathcal{B}))_{ij} = 0 \quad \text{whenever } i \neq j.$$

Definition 2.14 (Identity matrix).
The n-dimensional identity matrix I is a square matrix of size $n \times n$ with 1s along the diagonal and 0s elsewhere, that is,

$$(I)_{ij} = \begin{cases} 1 & \text{whenever } i = j, \\ 0 & \text{whenever } i \neq j. \end{cases}$$

Definition 2.15 (Invertible matrix).
A matrix A is invertible if and only if there exists a matrix B such that

$$AB = BA = I$$

This matrix B is called the inverse of A and is denoted A^{-1}. As we have commutative matrices, $AB = BA$, recall that this can only happen if A is square. So, only square matrices can have inverses.

Definition 2.16 (Determinant of a 1 by 1 matrix).
The determinant of any 1 by 1 matrix is given by its only coefficient:

$$\det\big((a)\big) = a$$

Definition 2.17 (Submatrix).
From a matrix A we generate the submatrix A_{ij} by deleting the ith row and jth column:

$$\text{For } A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,j-1} & a_{i,j} & a_{i,j+1} & \cdots & a_{i,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j} & a_{m,j+1} & \cdots & a_{m,n} \end{pmatrix}$$
$$\text{The submatrix } A_{ij} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{m,n} \end{pmatrix}$$

Note: we generally have to specify in words that we create a submatrix. The notation A_{ij} is a little ambiguous without being explicit.

Definition 2.18 (Determinant of an $n \times n$ matrix)**.**
For any square matrix $A \in \mathcal{M}_{n,n}$, its determinant is given by

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

where the a_{ij} are coefficients of A, A_{ij} is the i, j^{th} submatrix of A and for any $1 \leq j \leq n$. We can also sum over the j index for any $1 \leq i \leq n$

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

and we will show that the answer is the same.

Definition 2.19 (Cramer system)**.**
Suppose we have the following linear system of equations (with unknowns equal to equations)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = y_n \end{cases} \qquad (S)$$

with the associated matrix form

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{Y}$$

We say that (S) is a Cramer system if $\det(A) \neq 0$.

Definition 2.20 (Cofactor matrix)**.**
From a matrix A we generate its cofactor matrix C_A which has entries given by determinants of submatrices of A with the same plus/minus pattern as in a determinant calculation. That is, the entries of C_A are $c_{ij} = (-1)^{i+j} \det(A_{ij})$:

$$C_A = \begin{pmatrix} |A_{11}| & -|A_{12}| & |A_{13}| & \cdots \\ -|A_{21}| & |A_{22}| & -|A_{23}| & \cdots \\ |A_{31}| & -|A_{32}| & |A_{33}| & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Chapter 6: Eigenvalues and Eigenvectors

Definition 6.1 (Eigenvalues and eigenvectors of a linear map)**.**
Consider an endomorphism $f : V \to V$ and a vector $\mathbf{v} \in V$. We call a number $\lambda \neq 0$ an eigenvalue of f if there exists a non-zero vector $\mathbf{v}$ satisfying the relation

$$f(\mathbf{v}) = \lambda \mathbf{v}.$$

We say that $\mathbf{v}$ is an eigenvector of f corresponding or associated to the eigenvalue λ.

Definition 6.2 (Eigenvectors and eigenvalues of a matrix)**.**
For a square matrix A, an eigenvector of A is a non-zero vector, $\mathbf{v}$, that satisfies the matrix equation

$$A\mathbf{v} = \lambda \mathbf{v}$$

where λ is called an eigenvalue of A. We say that $\mathbf{v}$ is an eigenvector of A corresponding or associated to the eigenvalue λ.

Definition 6.3 (Characteristic polynomial of a matrix)**.**
For a square matrix A, the characteristic polynomial is the given by

$$P(\lambda) = \det(A - \lambda I)$$

where I is the identity matrix with the same size as A and λ is the variable of the polynomial. The degree of this polynomial is always the same as the size of the matrix A.

Definition 6.4 (Eigenspectrum)**.**
The eigenvalues of a square matrix A are roots of the characteristic polynomial of A. That is, eigenvalues are solutions of

$$\det(A - \lambda I) = 0.$$

There can be multiple distinct eigenvalues of A, and are conventionally denoted λ_1, λ_2, $\ldots$ etc. The set of these eigenvalues, $\{\lambda_1, \lambda_2, \ldots\}$, is called the eigenspectrum of A.

Definition 6.5 (Eigenspace)**.**
For any eigenvalue λ_k of an $n \times n$ matrix A, the eigenspace corresponding to λ_k, denoted E_{λ_k}, is the set of all

Definition 5.11 (Transition matrix (change-of-basis matrix)).
The transition matrix changes the representation of the coordinates of a vector from one basis into another. Let $\mathcal{A}$ and $\mathcal{B}$ be two bases of the same vector space, V, and let $\mathbf{v} \in V$. The transition matrix from $\mathcal{A}$ to $\mathcal{B}$, denoted $P_{\mathcal{A}\to\mathcal{B}}$ is defined by the relation

$$P_{\mathcal{A}\to\mathcal{B}}[\mathbf{v}]_{\mathcal{A}} = [\mathbf{v}]_{\mathcal{B}}.$$

If we let the bases $\mathcal{A} = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ and $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ then the transition matrix can be calculated by

$$P_{\mathcal{A}\to\mathcal{B}} = \begin{pmatrix} | & | & & | \\ [\mathbf{a}_1]_{\mathcal{B}} & [\mathbf{a}_2]_{\mathcal{B}} & \dots & [\mathbf{a}_n]_{\mathcal{B}} \\ | & | & & | \end{pmatrix}$$

where the vertical lines are reminders that the coordinates of the A basis vectors are columns.

Definition 5.12 (Changing the bases of a matrix representation).
Let $f : U \to V$ be a linear map, $\mathcal{B}_U$ and $\mathcal{B}'_U$ be two bases of U, $\mathcal{B}_V$ and $\mathcal{B}'_V$ be two bases of V, and $F = \mathcal{M}(f, \mathcal{B}_U \to \mathcal{B}_V)$ be the matrix representation of f from basis $\mathcal{B}_U$ to basis $\mathcal{B}_V$.
Then $F' = \mathcal{M}(f, \mathcal{B}'_U \to \mathcal{B}'_V)$, the matrix representation of f from basis $\mathcal{B}'_U$ to basis $\mathcal{B}'_V$, is given by

$$F' = P_{\mathcal{B}_V\to\mathcal{B}'_V}\, F\, P_{\mathcal{B}'_U\to\mathcal{B}_U}$$

The following diagram may help visualise this relation

$$\begin{tikzcd}
\mathcal{B}_U \arrow[r, "F"] & \mathcal{B}_V \arrow[d, "P_{\mathcal{B}_V\to\mathcal{B}'_V}"] \\
\mathcal{B}'_U \arrow[u, "P_{\mathcal{B}'_U\to\mathcal{B}_U}"] \arrow[r, "F'"'] & \mathcal{B}'_V
\end{tikzcd}$$

To read this schematic, consider that the arrow for F' has the same input and output as following the other three arrows to go up, then right (through F) then down again. This ordered path is the matrix multiplication given above.

Chapter 3: Linear Systems

Definition 3.1 (Linear system of equations).
A system of m linear equations with n unknowns, denoted (S), has the general form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_m \end{cases} \qquad (S)$$

where the x_j are the unknowns we want to find, a_{ij} are the coefficients and the y_i are the constant terms.

Definition 3.2 (Homogeneous linear system).
For any system of linear equations, (S), given by $AX = Y$, we associate the ***homogeneous system****, denoted (H):*

$$AX = 0_m$$

for the column

$$0_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathcal{M}_{m,1}$$

We will denote the solution set of (H) by $\mathcal{H}$.
Note: the homogeneous system always admits at least one solution, the trivial solution $X = 0_n$.

Definition 3.3 (Equivalent systems).
Two systems of linear equations are ***equivalent*** *if they share the same set of solutions.*

Definition 3.4 (Elementary operations).

There are ***elementary operations*** *that we can do to systems of equations that give new systems that remain equivalent to the old.*

$$\underline{\textit{Exchanging two equations}}$$
$$(S_1)\begin{cases} I_1 + I_2 - I_3 = 0 \\ 13I_1 - 6I_2 = 20 \end{cases} \equiv (S_2)\begin{cases} 13I_1 - 6I_2 = 20 \\ I_1 + I_2 - I_3 = 0 \end{cases}$$

$$\underline{\textit{Multiplying one equation by a non-zero constant}}$$

$$(S_1)\begin{cases} I_1 + I_2 - I_3 = 0 \\ 13I_1 - 6I_2 = 20 \end{cases} \equiv (S_2)\begin{cases} I_1 + I_2 - I_3 = 0 \\ I_1 - (6/13)I_2 = (20/13) \end{cases}$$

$$\underline{\textit{Adding a multiple of one equation to another equation}}$$

$$(S_1)\begin{cases} I_1 + I_2 - I_3 = 0 \\ 13I_1 - 6I_2 = 20 \end{cases} \equiv (S_2)\begin{cases} I_1 + I_2 - I_3 = 0 \\ 15I_1 - 4I_2 - 2I_3 = 20 \end{cases}$$

Definition 3.5 (Overdetermined system).
An overdetermined system has more equations than unknowns. We say "there are too many equations". Such a system allows solutions only if certain conditions are met.

Definition 3.6 (Underdetermined system).

*An **underdetermined system** has less equations than unknowns. We say "there are not enough equations". Such a system has either no solutions, or infinitely many.*

Definition 3.7 (Cramer system).
Suppose we have the following linear system of equations (with unknowns equal to equations)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = y_n \end{cases} \qquad (S)$$

with the associated matrix form

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{Y}$$

We say that (S) is a Cramer system if $\det(A) \neq 0$.

Definition 5.7 (Surjectivity).
Let $f : V \to W$ be a linear map. We say that f is surjective if every vector in the output space has a corresponding input vector. In symbols

$$\forall\, \mathbf{w} \in W \quad \exists \mathbf{v} \in V \text{ such that } f(\mathbf{v}) = \mathbf{w}.$$

Definition 5.8 (Categories of linear maps).
Let $f : V \to W$ be a linear map.

- *If $W = V$ we call f an endomorphism.*
- *If f is both injective and surjective then we say it is bijective and we call it an isomorphism.*
- *If f is both an isomorphism and an endomorphism we call it an automorphism.*

Definition 5.9 (Composition of linear maps).
Composition of linear maps works exactly as you would expect if you remember the composition of regular functions. We must have a coherence between the output of one linear map and the input of another. So, two linear maps $f : A \to B$ and $g : U \to V$ can be composed as a well defined linear map $g \circ f$ ("g of f") if and only if the output space of f is the input space of g: $U = B$. For any $\mathbf{u} \in A$ the composition is written

$$g \circ f : A \to V \quad \text{and} \quad (g \circ f)(\mathbf{u}) = g(f(\mathbf{u})).$$

Definition 5.10 (Matrix representation of a linear map).
Let $f : V \to W$ be a linear map, $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ be a basis of V, $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ be a basis of W. Let $\mathbf{v}$ be any vector in V and $\mathbf{w} = f(\mathbf{v}) \in W$. Then the matrix representation of f in bases $\mathcal{A}$ and $\mathcal{B}$, defined by the unique $m \times n$ matrix, denoted $\mathcal{M}(f, \mathcal{A} \to \mathcal{B})$, which takes the coordinates of $\mathbf{v}$ to the coordinates of $\mathbf{w}$ in their respective bases:

$$\mathcal{M}(f, \mathcal{A} \to \mathcal{B})[\mathbf{v}]_{\mathcal{A}} = [\mathbf{w}]_{\mathcal{B}}.$$

can be calculated by expressing the coordinates of the linear map acting on the basis vectors of the input space

$$\mathcal{M}(f, \mathcal{A} \to \mathcal{B}) = \begin{pmatrix} | & & | \\ [f(\mathbf{a}_1)]_{\mathcal{B}} & \ldots & [f(\mathbf{a}_n)]_{\mathcal{B}} \\ | & & | \end{pmatrix}$$

where the vertical lines are reminders that the coordinates of the $f(\mathbf{a}_k)$ vectors are columns. We often shorten "matrix representation of f" to just "matrix of f". If the input and output vector spaces are the same, i.e. if f is an endomorphism, we can use the same basis for both spaces and we may shorten the notation $\mathcal{M}(f, \mathcal{A} \to \mathcal{A}) = \mathcal{M}(f, \mathcal{A})$.

Chapter 5: Linear Maps

Definition 5.1 (Linear map).
A mapping, f, from a vector space V to a vector space W, denoted $f : V \to W$, is called a linear map if it satisfies the following property:

$$\forall \mathbf{u},\, \mathbf{v} \in V, \forall \alpha, \beta \in \mathbb{R}$$
$$f(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha f(\mathbf{u}) + \beta f(\mathbf{v}).$$

We say that a linear map preserves linear combinations.

Definition 5.2 (Image).
The image of a linear map $f : V \to W$, denoted $im(f)$, is the set of all possible "output" vectors of the map:

$$im(f) = \{\mathbf{w} \in W \mid \exists \mathbf{v} \in V\; f(\mathbf{v}) = \mathbf{w}\} \subseteq W.$$

Definition 5.3 (Rank).
The rank of a linear map is the dimension of its image: $rank(f) = \dim(im(f))$.

Definition 5.4 (Kernel).
The kernel of a linear map $f : V \to W$, denoted $\ker(f)$, is the set of vectors that f maps to the zero vector, $\mathbf{0}_W$, of W. That is,

$$\ker(f) = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0}_W\}.$$

Definition 5.5 (Nullity).
The nullity of a linear map is the dimension of its kernel: $nullity(f) = \dim(\ker(f))$.

Definition 5.6 (Injectivity).
Let $f : V \to W$ be a linear map. We say f is injective if no two vectors of V are mapped to the same vector of W. In symbols we have two equivalent expressions

$$\forall \mathbf{x}, \mathbf{y} \in V, \quad (f(\mathbf{x}) = f(\mathbf{y}) \implies \mathbf{x} = \mathbf{y})$$
$$\text{or}$$
$$\forall \mathbf{x}, \mathbf{y} \in V, \quad (\mathbf{x} \neq \mathbf{y} \implies f(\mathbf{x}) \neq f(\mathbf{y}))$$

Chapter 4: Vector Spaces

Definition 4.1 (Vector space).
A vector space over a field $\mathbb{F}$ is a set, call it V, with elements called vectors supplied with definitions of two operations, vector addition (VA) and scalar multiplication (SM), that satisfy the following vector space axioms:

$$\forall\, \mathbf{u}, \mathbf{v}, \mathbf{w} \in V \quad \text{and} \quad \forall\, k, l \in \mathbb{F}$$

(VA1)	$\mathbf{u} + \mathbf{v} \in V$	*(closure under vector addition)*
(VA2)	$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$	*(associativity of vector addition)*
(VA3)	$\exists\, \mathbf{0} \in V$, *such that* $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$	*(additive identity)*
(VA4)	$\exists - \mathbf{u} \in V$ *such that* $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$	*(additive inverse)*
(VA5)	$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	*(commutativity of vector addition)*
(SM1)	$k\mathbf{u} \in V$	*(closure under scalar multiplication)*
(SM2)	$k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$	*(distributivity over vector addition)*
(SM3)	$(k + l)\mathbf{u} = k\mathbf{u} + l\mathbf{u}$	*(distributivity over field addition)*
(SM4)	$k(l\mathbf{u}) = (kl)\mathbf{u}$	*(compatibility of scalar and field multiplication)*
(SM5)	$1\mathbf{u} = \mathbf{u}$	*(multiplicative identity)*

Definition 4.2 (Linear Combination).
Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of vectors in a vector space V. A linear combination of these vectors is a new vector, $\mathbf{w} \in V$, of the form

$$\mathbf{w} = \alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n$$

where the α_k are real numbers.

Definition 4.3 (Vector subspace).
Suppose that V is a vector space and W is a subset of V. We call W a vector subspace if it satisfies the vector space axioms for the same definition of vector addition and scalar multiplication defined for V.

Definition 4.4 (Span).
Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of vectors from a vector space V. The span of these vectors is the set of all linear combinations of those vectors:

$$SPAN(\mathcal{B}) = SPAN(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \{\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n \,|\, \forall \alpha_1, \ldots, \alpha_n \in \mathbb{R}^n\}.$$

This set forms a vector subspace of V. It is obviously non-empty because it at least contains the vectors of $\mathcal{B}$. It is also automatically closed under vector addition and scalar multiplication because those are exactly the operations we used to create all the vectors in the span! Therefore $SPAN(\mathcal{B})$ is a vector subspace of V.

Definition 4.5 (Cartesian form of Euclidean vector sub spaces)**.**
Euclidean vector sub spaces can always be written as a set with some defining equations:

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n \,|\, \textit{equations relating the } x_k\}.$$

For example, the general form of planar vector subspaces of $\mathbb{R}^3$ *is*

$$V_P = \left\{(x, y, z) \in \mathbb{R}^3 \,|\, ax + by + cz = 0\right\}$$

where a*,* b *and* c *are some given constants. This set is read aloud as "all the triples* (x, y, z) *such that* $ax + by + cz = 0$*".*

Definition 4.6 (Sum of subspaces (sum space))**.**
Suppose we have a vector space V *with vector subspaces* F *and* G*. We define the* ***sum of subspaces*** *(or sum space) as a new set denoted*

$$F + G = \{\boldsymbol{f} + \boldsymbol{g} \,|\, \boldsymbol{f} \in F,\ \boldsymbol{g} \in G\}$$

Note: The sum space is a subset of the parent vector space: $F + G \subset V$*.*

Definition 4.7 (Direct sum)**.**
Let F *and* G *be two vector subspaces of a vector space* V *and let* $E = F + G$ *be the sum space. We say* E *is a* ***direct sum*** *of* F *and* G *if each element of* E *has a* ***unique*** *decomposition as a sum of vectors in* F *and vectors in* G*. That is, for every* $\boldsymbol{v} \in E$*, there exists unique vectors* $\boldsymbol{f} \in F$ *and* $\boldsymbol{g} \in G$ *such that* $\boldsymbol{v} = \boldsymbol{f} + \boldsymbol{g}$*. We denote this direct sum with a new symbol*

$$E = F \oplus G$$

Definition 4.8 (Complementary vector subspaces)**.**
Let F *and* G *be two vector subspaces of* V*.* F *and* G *are called* ***complementary*** *if* V *is a direct sum of* F *and* G*. That is, if and only if*

- $V = F + G$*, and*
- $F \cap G = \{\boldsymbol{0}_V\}$

Definition 4.9 (Linear dependence)**.**
A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *from a vector space* V *is said to be linearly dependent if there exists a set of constants* $\{\alpha_1, \dots, \alpha_n\}$ *not all zero such that*

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}_V.$$

Note: the right hand side of the equation is the zero vector, not the real number 0*.*

Definition 4.10 (Linear independence)**.**
A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *from* V *is said to be linearly independent if they are not linearly dependent. That is, the equation*

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}_V.$$

implies that the constants $\alpha_1, \dots, \alpha_n$ *are all zero.*

Definition 4.11 (Basis)**.**
A basis of a vector space V *is a minimal set of vectors which spans the vector space. Formally, the set of vectors* $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ *in a vector space* V *is a basis of* V *if it is a set of linearly independent vectors and* $SPAN(\mathbf{v}_1, \dots, \mathbf{v}_n) = V$*. Note: bases are not unique, but they always contain the same number of vectors.*

Definition 4.12 (Dimension)**.**
The dimension of a vector space is the number of elements in a basis for that vector space.

Definition 4.13 (Canonical basis of $\mathbb{R}^n$)**.**
The canonical basis of the vector space of real n*-tuples,* $\mathbb{R}^n$*, is the ordered set of* n n*-tuples with* k^{th} *element,* $\mathbf{c}_k = (\alpha_1, \dots, \alpha_n)$ *such that*

$$\alpha_j = \begin{cases} 1 & \textit{for } j = k, \\ 0 & \textit{for } j \neq k. \end{cases}$$

That is, as a set the canonical basis is

$$\mathcal{C}_n = \{(1, 0, \dots, 0),\ (0, 1, \dots, 0),\ \dots,\ \underbrace{(0, 0, \dots, 0, \overbrace{\ 1\ }^{k^{th}\ place}, 0, \dots, 0)}_{k^{th}\ tuple},\ \dots,\ (0, 0, \dots, 1)\}.$$

Definition 4.14 (Canonical basis of $\mathcal{P}_n$)**.**
The canonical basis of the vector space of polynomials with degree up to n*,* $\mathcal{P}_n$*, is the ordered set of* n *polynomials with* k^{th} *element,* $\mathbf{c}_k = x^k$*. That is, as a set the canonical basis is*

$$\mathcal{C}_n = \{1,\ x,\ x^2,\ \dots,\ x^n\}.$$

Definition 4.15 (Coordinates of a vector)**.**
Let $\mathbf{v}$ *be a vector in a vector space* V*. The coordinates of* $\mathbf{v}$ *with respect to a given basis* $\mathcal{B}$*, denoted* $[\mathbf{v}]_{\mathcal{B}}$*, is a column of the unique set of coefficients in the linear combination of* $\mathbf{v}$ *in terms of the basis vectors.*