LINEAR ALGEBRA SUMMARY SHEET

RADON ROSBOROUGH

https://intuitiveexplanations.com/math/linear-algebra-summary-sheet/

This document is a concise collection of many of the important theorems of linear algebra, organized so that making connections between different concepts and theorems is as easy as possible. The material here is primarily based on the textbook I used in my Linear Algebra class (*Linear Algebra and Its Applications* by David C. Lay, fourth edition).

1. Notation

(1) I denotes an identity matrix, the size of which is dependent on context.
(2) If $\mathcal{B}$ is a basis for $\mathbb{R}^n$, then $P_{\mathcal{B}}$ denotes the change-of-coordinates matrix from $\mathcal{B}$ to $\mathbb{R}^n$, which satisfies $\mathbf{x} = P_{\mathcal{B}} [\mathbf{x}]_{\mathcal{B}}$ for all $\mathbf{x}$ in $\mathbb{R}^n$.
(3) If $\mathcal{B}$ and $\mathcal{C}$ are bases for a vector space V, then $P_{\mathcal{C}\leftarrow\mathcal{B}}$ denotes the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$, which satisfies $[\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}} [\mathbf{x}]_{\mathcal{B}}$ for all $\mathbf{x}$ in V.
(4) If $\mathcal{B}$ is a basis for a vector space V and $T : V \to V$ is a linear transformation, then $[T]_{\mathcal{B}}$ denotes the matrix for T relative to $\mathcal{B}$, or the $\mathcal{B}$-matrix for T, which satisfies $[T(\mathbf{x})]_{\mathcal{B}} = [T]_{\mathcal{B}} [\mathbf{x}]_{\mathcal{B}}$ for all $\mathbf{x}$ in V.
(5) If $\mathcal{B}$ and $\mathcal{C}$ are bases for the vectors spaces V and W, respectively, and $T : V \to W$ is a linear transformation, then $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ denotes the matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$, which satisfies $[T(\mathbf{x})]_{\mathcal{C}} = [T]_{\mathcal{C}\leftarrow\mathcal{B}} [\mathbf{x}]_{\mathcal{C}}$ for all $\mathbf{x}$ in V.

2. Theorems

Theorem 1 (Solution Sets of Linear Systems).

(1) A linear system either is inconsistent, has exactly one solution, or has infinitely many solutions.
(2) Let A be an $m \times n$ matrix and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$. If the equation $A\mathbf{x} = \mathbf{b}$ is consistent, then the mapping $\mathbf{x} \mapsto \mathbf{x} + \mathbf{b}$ is a bijection from the solution set of $A\mathbf{x} = \mathbf{0}$ to the solution set of $A\mathbf{x} = \mathbf{b}$.

Theorem 2 (Consistent Linear Systems). Let $A = \begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_n\end{bmatrix}$ be an $m \times n$ matrix, let $\mathbf{b}$ be a vector in $\mathbb{R}^m$, let M be the augmented matrix $\begin{bmatrix}A & \mathbf{b}\end{bmatrix}$, and let L be the linear system with augmented matrix M. Then the following statements are logically equivalent:

(1) The linear system L is consistent.
(2) The last column of M is not a pivot column.
(3) The echelon form of M has no row of the form $\begin{bmatrix}0 & \cdots & 0 & b\end{bmatrix}$, where $b \neq 0$.
(4) The equation $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ has at least one solution.
(5) The vector $\mathbf{b}$ is a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_n$.
(6) The vector $\mathbf{b}$ is an element of $\operatorname{Span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$.
(7) The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution.

Theorem 3 (Always-Consistent Linear Systems). Let $A = \begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_n\end{bmatrix}$ be an $m \times n$ matrix and let $T : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation given by $\mathbf{x} \mapsto A\mathbf{x}$. Then the following statements are logically equivalent:

(1) For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution.
(2) Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_n$.
(3) The vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ span $\mathbb{R}^m$.
(4) The matrix A has a pivot position in every row.
(5) The linear transformation T is onto (it maps $\mathbb{R}^n$ onto $\mathbb{R}^m$).

Theorem 4 (Linear Systems with Unique Solutions). Let $A = \begin{bmatrix}\mathbf{a}_1 & \cdots & \mathbf{a}_n\end{bmatrix}$ be an $m \times n$ matrix, let $\mathbf{b}$ be a vector in $\mathbb{R}^m$, let L be the linear system with augmented matrix $\begin{bmatrix}A & \mathbf{b}\end{bmatrix}$, and let $T : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation given by $\mathbf{x} \mapsto A\mathbf{x}$. Then the following statements are logically equivalent:

(1) The linear system L has at most one solution.
(2) The linear system L either is inconsistent or has no free variables.
(3) The matrix A has a pivot position in every column.
(4) The equation $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ has at most one solution.

(5) The equation $A\mathbf{x} = \mathbf{b}$ has at most one solution.
(6) The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
(7) The vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent.
(8) The linear transformation T is one-to-one.

Theorem 5 (Square Linear Systems). Let $A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$ be an $n \times n$ matrix, and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation given by $\mathbf{x} \mapsto A\mathbf{x}$. Then the following statements are logically equivalent:

(1) The matrix A is invertible.
(2) The matrix A is row equivalent to the $n \times n$ identity matrix.
(3) The matrix A has n pivot positions.
(4) For each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution.
(5) The vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ span $\mathbb{R}^n$.
(6) The linear transformation T is onto (it maps $\mathbb{R}^n$ onto $\mathbb{R}^n$).
(7) For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has at most one solution.
(8) The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
(9) The vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent.
(10) The linear transformation T is one-to-one.
(11) $\det A \neq 0$.
(12) The vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are a basis for $\mathbb{R}^n$.
(13) $\operatorname{Col} A = \mathbb{R}^n$.
(14) $\operatorname{Row} A = \mathbb{R}^n$.
(15) $\operatorname{rank} A = n$.
(16) $\operatorname{Nul} A = \{\mathbf{0}\}$.
(17) $\operatorname{Nul} A^T = \{\mathbf{0}\}$.
(18) $\dim \operatorname{Nul} A = 0$.
(19) $\dim \operatorname{Nul} A^T = 0$.
(20) 0 is not an eigenvalue of A.

Theorem 6 (Inversion of a Matrix). Let A be an invertible matrix. Then $\begin{bmatrix} A & I \end{bmatrix}$ is row equivalent to $\begin{bmatrix} I & A^{-1} \end{bmatrix}$.

Theorem 7 (LU Factorization). Let A be an $m \times n$ matrix, let U be an echelon form of A obtained only by adding multiples of rows to rows below them, and let L be an $m \times m$ matrix that is reduced to I_m by the same sequence of elementary row operations that reduces A to U. Then $A = LU$, where L is lower triangular and U is upper triangular.

Theorem 8 (Properties of Linear Dependence). Let $\{\mathbf{a}_1, \ldots, \mathbf{a}_p\}$ be a set of vectors in $\mathbb{R}^n$. Then the following statements are true:

(1) Suppose that $p = 1$. Then the set of vectors is linearly dependent if and only if $\mathbf{a}_1 = \mathbf{0}$.
(2) Suppose that $p = 2$. Then the set of vectors is linearly dependent if and only if at least one of the vectors is a scalar multiple of the other.
(3) The set of vectors is linearly dependent if and only if at least one of the vectors is a linear combination of the other vectors.
(4) If the set of vectors is linearly dependent and $\mathbf{a}_1 \neq \mathbf{0}$, then at least one of the vectors $\mathbf{a}_2, \ldots, \mathbf{a}_p$ is a linear combination of the preceding vectors.
(5) If $p > n$, then the set of vectors is linearly dependent.
(6) If at least one of the vectors is $\mathbf{0}$, then the set of vectors is linearly dependent.

Theorem 9 (Properties of Linear Transformations).

(1) A transformation T is linear if and only if both $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and $T(c\mathbf{u}) = cT(\mathbf{u})$ for all vectors $\mathbf{u}, \mathbf{v}$ and all scalars c.

(2) A transformation T is linear if and only if $T(c\mathbf{u} + d\mathbf{v}) = c\,T(\mathbf{u}) + d\,T(\mathbf{v})$ for all vectors $\mathbf{u}, \mathbf{v}$ and all scalars c, d.
(3) If a transformation T is given by $\mathbf{x} \mapsto A\mathbf{x}$ for some $m \times n$ matrix A, then T is a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$.
(4) Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is given by $\mathbf{x} \mapsto A\mathbf{x}$, where $A = \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix}$.

Theorem 10 (Properties of Matrix Operations).

(1) Let A, B, and C be $m \times p$, $p \times q$, and $q \times n$ matrices respectively. Then:
   (a) $A(BC) = (AB)C$.
   (b) $A(B + C) = AB + AC$.
   (c) $(A + B)C = AC + BC$.
(2) Let A and B be $m \times n$ matrices. Then:
   (a) $(A^T)^T = A$.
   (b) $(A + B)^T = A^T + B^T$.
(3) Let $A_1, \dots, A_n$ be matrices such that the product $A_1 \cdots A_n$ is defined. Then $(A_1 \cdots A_n)^T = A_n^T \cdots A_1^T$.
(4) Let $A_1, \dots, A_n$ be $n \times n$ matrices. Then $A_1 \cdots A_n$ is invertible if and only if each of $A_1, \dots, A_n$ are invertible. If $A_1 \cdots A_n$ is invertible, then $(A_1 \cdots A_n)^{-1} = A_n^{-1} \cdots A_1^{-1}$.
(5) Let A be an $n \times n$ matrix. Then A^T is invertible if and only if A is invertible. If A^T is invertible, then $(A^T)^{-1} = (A^{-1})^T$.

Theorem 11 (Properties of Determinants).

(1) If A is a triangular matrix, then $\det A$ is the product of the entries on the main diagonal of A.
(2) If A is a square matrix, and the matrix B is obtained from A by performing one elementary row operation[1], then:
   (a) If a multiple of one row of A is added to another row, then $\det B = \det A$.
   (b) If two rows of A are interchanged, then $\det B = -\det A$.
   (c) If one row of A is multiplied by k, then $\det B = k \det A$.
(3) If A and B are square matrices, then $\det AB = (\det A)(\det B)$.
(4) If A is invertible, then $\det A$ is $(-1)^n$ multiplied by the product of the pivots in any echelon form of A, where n is the number of row interchanges necessary to reduce A to echelon form.
(5) If A is an $n \times n$ matrix, $\mathbf{b}$ is a vector in $\mathbb{R}^n$, $\det A \neq 0$, and $A\mathbf{x} = \mathbf{b}$, then

$$\mathbf{x} = \frac{1}{\det A} \begin{bmatrix} \det A_1(\mathbf{b}) \\ \vdots \\ \det A_n(\mathbf{b}) \end{bmatrix},$$

   where $A_i(\mathbf{b})$ denotes the matrix obtained by replacing column i of A with $\mathbf{b}$.
(6) If A is an $n \times n$ matrix and $\det A \neq 0$, then

$$A^{-1} = \frac{\operatorname{adj} A}{\det A},$$

[1] The same applies to elementary column operations.

where

$$\operatorname{adj} A = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}^T,$$

$$C_{ij} = (-1)^{i+j} \det A_{ij},$$

and A_{ij} denotes the matrix obtained by removing row i and column j from A.

(7) If $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix}$ is a 2×2 matrix, then $|\det A|$ is the area of the parallelogram determined by $\mathbf{a}_1$ and $\mathbf{a}_2$. If $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix}$ is a 3×3 matrix, then $|\det A|$ is the volume of the parallelepiped determined by $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$.

(8) If A is a 2×2 matrix, $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear transformation defined by $\mathbf{x} \mapsto A\mathbf{x}$, and S is a region in $\mathbb{R}^2$ with finite area, then the area of $T(S)$ is the area of S multiplied by $|\det A|$. If A is a 3×3 matrix, $T : \mathbb{R}^3 \to \mathbb{R}^3$ is the linear transformation defined by $\mathbf{x} \mapsto A\mathbf{x}$, and S is a region in $\mathbb{R}^3$ with finite volume, then the volume of $T(S)$ is the volume of S multiplied by $|\det A|$.

Theorem 12 (Properties of Subspaces).

(1) If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are vectors in a vector space V, then $\operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a subspace of V.
(2) If A is an $m \times n$ matrix, then $\operatorname{Nul} A$ is a subspace of $\mathbb{R}^n$.
(3) If A is an $m \times n$ matrix, then $\operatorname{Col} A$ is a subspace of $\mathbb{R}^m$.
(4) If A is an $m \times n$ matrix, then $\operatorname{Row} A$ is a subspace of $\mathbb{R}^n$.

Theorem 13 (Properties of Bases).

(1) Let S be a finite subset of a vector space and let $H = \operatorname{Span} S$. If one of the vectors, $\mathbf{v}$, in S is a linear combination of the others, then the set formed by removing $\mathbf{v}$ from S still spans H.
(2) Let $S \neq \{\mathbf{0}\}$ be a subset of a vector space V. If S spans V, then some subset of S is a basis for V, and if S is linearly independent, then some superset of S is a basis for V.
(3) If a vector space V has a basis containing n vectors, then every basis of V must contain exactly n vectors. Furthermore, any subset of V containing more than n vectors must be linearly dependent, and any subset of V containing more than n elements cannot span V.
(4) If V is an n-dimensional vector space, then any subset of V containing exactly n vectors that either spans V or is linearly independent must be a basis for V.
(5) If H is a subspace of a finite-dimensional vector space V, then $\dim H \leq \dim V$.

Theorem 14 (Properties of Coordinates).

(1) Let $\mathcal{B}$ be a basis with n vectors for a vector space V. Then the transformation $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is an isomorphism from V to $\mathbb{R}^n$.
(2) Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for $\mathbb{R}^n$. Then $P_{\mathcal{B}} = \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{bmatrix}$.
(3) Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C}$ be bases for a vector space V. Then $P_{\mathcal{C} \leftarrow \mathcal{B}} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} \cdots [\mathbf{b}_n]_{\mathcal{C}} \end{bmatrix}$.
(4) Let $\mathcal{B}$ and $\mathcal{C}$ be bases for a vector space V. Then $P_{\mathcal{C} \leftarrow \mathcal{B}} = (P_{\mathcal{B} \leftarrow \mathcal{C}})^{-1}$.
(5) Let $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ be bases for a vector space V. Then $P_{\mathcal{D} \leftarrow \mathcal{C}} P_{\mathcal{C} \leftarrow \mathcal{B}} = P_{\mathcal{D} \leftarrow \mathcal{B}}$.
(6) Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be bases for $\mathbb{R}^n$. Then $\begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n & \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{bmatrix}$ is row equivalent to $\begin{bmatrix} I & P_{\mathcal{C} \leftarrow \mathcal{B}} \end{bmatrix}$.
(7) Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be bases for a subspace H of $\mathbb{R}^m$. Then the matrix $\begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_n & \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{bmatrix}$ is row equivalent to

$$\begin{bmatrix} I_n & P_{\mathcal{C} \leftarrow \mathcal{B}} \\ 0 & 0 \end{bmatrix}$$

(8) Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for a vector space V, and let $T : V \to V$ be a linear transformation. Then

$$[T]_{\mathcal{B}} = \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{B}} & \cdots & [T(\mathbf{b}_n)]_{\mathcal{B}} \end{bmatrix}.$$

(9) Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be bases for the vector spaces V and W, respectively, and let $T : V \to W$ be a linear transformation. Then

$$[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & \cdots & [T(\mathbf{b}_n)]_{\mathcal{C}} \end{bmatrix}.$$

(10) Let A, P, and B be $n \times n$ matrices such that $A = PBP^{-1}$, let $\mathcal{B}$ be the basis for $\mathbb{R}^n$ formed from the columns of P, and let T be the transformation given by $\mathbf{x} \mapsto A\mathbf{x}$. Then $[T]_{\mathcal{B}} = B$. In particular, if B is a diagonal matrix, then $\mathcal{B}$ is the eigenvector basis and the diagonal entries of B are the corresponding eigenvalues.

Theorem 15 (Properties of the Null Space, Row Space, and Column Space). Let A be an $n \times m$ matrix.

(1) The pivot columns of A are a basis for $\operatorname{Col} A$, while the pivot rows of any echelon form of A are a basis for $\operatorname{Row} A$.
(2) The equalities $\operatorname{Col} A = \operatorname{Row} A^T$ and $\operatorname{Row} A = \operatorname{Col} A^T$ hold.
(3) The dimension of $\operatorname{Nul} A$ is equal to the number of free variables in the equation $A\mathbf{x} = \mathbf{0}$ (or equivalently the number of nonpivot columns of A), while the dimension of $\operatorname{Nul} A^T$ is equal to the number of free variables in the equation $A^T\mathbf{x} = \mathbf{0}$ (or equivalently the number of nonpivot columns of A^T).
(4) The dimension of $\operatorname{Col} A$, the dimension of $\operatorname{Row} A$, the number of pivot positions in A, and the number of pivot positions in A^T are all equal.
(5) The equalities $\operatorname{rank} A + \dim \operatorname{Nul} A = n$ and $\operatorname{rank} A + \dim \operatorname{Nul} A^T = m$ hold.
(6) $\operatorname{Row} A$ and $\operatorname{Nul} A$ are orthogonal complements, and $\operatorname{Col} A$ and $\operatorname{Nul} A^T$ are orthogonal complements.

Theorem 16 (Properties of Eigenvectors and Eigenvalues).

(1) A number λ is an eigenvalue of a matrix A if and only if $\det(A - \lambda I) = 0$.
(2) The eigenvalues of a triangular matrix are the entries on its main diagonal, with the number of times an entry is repeated being equal to its multiplicity as an eigenvalue.
(3) Any set of n eigenvectors corresponding to n unique eigenvalues of a matrix is linearly independent.
(4) Let A be a matrix with at least n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. If S_k is a linearly independent set of eigenvectors corresponding to λ_k, for $1 \le k \le n$, then the union of $S_1, \ldots, S_n$ is also linearly independent.
(5) The eigenspace for any given eigenvalue of a matrix is a subspace of $\mathbb{R}^n$.
(6) The dimension of the eigenspace for any given eigenvalue of a matrix cannot exceed the multiplicity of the eigenvalue.
(7) The eigenspaces for any two distinct eigenvalues of a matrix are disjoint except at $\mathbf{0}$.
(8) If an $n \times n$ matrix is diagonalizable and $\mathcal{B}_1, \ldots, \mathcal{B}_m$ are bases for the m distinct eigenvalues of A, then the union of $\mathcal{B}_1, \ldots, \mathcal{B}_m$ is a basis for $\mathbb{R}^n$.
(9) If two matrices are similar, then they have the same eigenvalues with the same multiplicities.
(10) Let A be an $n \times n$ matrix. Then the following statements are logically equivalent:
  (a) The matrix A is diagonalizable.
  (b) The matrix A has n linearly independent eigenvectors.
  (c) The sum of the dimensions of the eigenspaces of A equals n.
  (d) The dimension of the eigenspace for each eigenvalue of A equals the multiplicity of the eigenvalue.

(11) If an $n \times n$ matrix A has n distinct eigenvalues, then it is diagonalizable.
(12) If a matrix A is diagonalizable, then $A = PDP^{-1}$, where the columns of P are linearly independent eigenvectors and the diagonal entries of D are the corresponding eigenvalues.
(13) Let A be a real 2×2 matrix with a complex eigenvalue $a - bi$ and an associated complex eigenvector $\mathbf{v}$. Then $A = PCP^{-1}$, where $P = \begin{bmatrix} \operatorname{Re} \mathbf{v} & \operatorname{Im} \mathbf{v} \end{bmatrix}$,

$$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix},$$

and r and ϕ are the magnitude and complex argument of $a + bi$.

Theorem 17 (Properties of Orthogonal Coordinates).

(1) Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be an orthogonal basis for an inner product space V, and let $\mathbf{y}$ be a vector in V. Then

$$\mathbf{y} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_n}{\mathbf{u}_n \cdot \mathbf{u}_n}\mathbf{u}_n.$$

(2) Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be an orthonormal basis for an inner product space V, and let $\mathbf{y}$ be a vector in V. Then

$$\mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_n)\mathbf{u}_n.$$

Theorem 18 (Properties of Orthogonal Projections). Let W be a subspace of $\mathbb{R}^m$, let $\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be a basis for W, let $U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix}$ be an $m \times n$ matrix, and let $\mathbf{y}$ be a vector in $\mathbb{R}^m$. Then:

(1) There is a unique representation of $\mathbf{y}$ in the form $\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}$, where $\widehat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. Typically $\widehat{\mathbf{y}}$ is denoted by $\operatorname{proj}_W \mathbf{y}$.
(2) The closest point in W to $\mathbf{y}$ is $\widehat{\mathbf{y}}$, i.e. $||\mathbf{y} - \widehat{\mathbf{y}}|| < ||\mathbf{y} - \mathbf{v}||$ if $\mathbf{v} \neq \widehat{\mathbf{y}}$.
(3) The equation $U^T U \mathbf{x} = U^T \mathbf{y}$ has a unique solution, and $\widehat{\mathbf{y}} = \mathbf{x}$.
(4) If $\mathcal{U}$ is an orthogonal basis for W, then

$$\widehat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_n}{\mathbf{u}_n \cdot \mathbf{u}_n}\mathbf{u}_n.$$

(5) If $\mathcal{U}$ is an orthonormal basis for W, i.e. U is an orthogonal matrix, then

$$\widehat{\mathbf{y}} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{y} \cdot \mathbf{u}_n)\mathbf{u}_n = UU^T\mathbf{y}.$$

(6) If $\mathcal{U}$ is an orthonormal basis for $\mathbb{R}^m$, i.e. U is a square orthogonal matrix, then $\widehat{\mathbf{y}} = \mathbf{y}$.
(7) Let

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{u}_1, \\ \mathbf{v}_2 &= \mathbf{u}_2 - \operatorname{proj}_{\operatorname{Span}\{\mathbf{v}_1\}} u_2, \\ &\vdots \\ \mathbf{v}_n &= \mathbf{u}_n - \operatorname{proj}_{\operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}} u_n. \end{aligned}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is an orthogonal basis for W, and

$$\begin{aligned} \operatorname{Span}\{\mathbf{v}_1\} &= \operatorname{Span}\{\mathbf{u}_1\} \\ \operatorname{Span}\{\mathbf{v}_1, \mathbf{v}_2\} &= \operatorname{Span}\{\mathbf{u}_1, \mathbf{u}_2\} \\ &\vdots \\ \operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} &= \operatorname{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}. \end{aligned}$$

This method of obtaining an orthogonal basis for W is called the Gram-Schmidt process.

Theorem 19 (Properties of Orthogonal Matrices).

(1) Let U be a matrix. Then U has orthonormal columns if and only if $UU^T = I$, and U has orthonormal rows if and only if $U^T U = I$.
(2) Let U be a square matrix. Then U has orthonormal columns if and only if it has orthonormal rows.
(3) Let U be an $m \times n$ orthogonal matrix, and let $\mathbf{x}$ and $\mathbf{y}$ be vectors in $\mathbb{R}^m$. Then U preserves lengths, angles, and orthogonality. That is:
  (a) $||U\mathbf{x}|| = ||\mathbf{x}||$.
  (b) $(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$.
  (c) $(U\mathbf{x}) \cdot (U\mathbf{y}) = 0$ if and only if $\mathbf{x} \cdot \mathbf{y} = 0$.
(4) Let A be an $m \times n$ matrix with linearly independent columns, and let $U = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix}$ be the $m \times n$ orthogonal matrix obtained by finding an orthogonal basis for $\operatorname{Col} A$ using the Gram-Schmidt process, normalizing the vectors, and using them as the columns of U. Let R be the $n \times n$ matrix equal to $Q^T A$. Also, for $1 \leq i \leq n$, if the (i, i) entry of R is negative, multiply the ith row of R and the ith column of U by -1. Then $A = PQ$, and R is an upper triangular invertible matrix with positive entries on its diagonal.

Theorem 20 (Properties of Inner Products). Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in an inner product space. Then:

(1) $||\mathbf{u} + \mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2$ if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.
(2) $|\mathbf{u} \cdot \mathbf{v}| \leq ||\mathbf{u}||\,||\mathbf{v}||$.
(3) $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$.

Theorem 21 (Applications).

(1) Let C be a square matrix whose entries are nonnegative and whose column sums are each less than 1. Then $I + C + C^2 + C^3 + \cdots = (I - C)^{-1}$.
(2) Let $\{u_k^{(1)}\}, \{u_k^{(2)}\}, \ldots, \{u_k^{(n)}\}$ be vectors in $\mathbb{S}$. If the matrix

$$\begin{bmatrix} \{u_k^{(1)}\} & \{u_k^{(2)}\} & \cdots & \{u_k^{(n)}\} \\ \{u_{k+1}^{(1)}\} & \{u_{k+1}^{(2)}\} & \cdots & \{u_{k+1}^{(n)}\} \\ \vdots & \vdots & \ddots & \vdots \\ \{u_{k+n-1}^{(1)}\} & \{u_{k+n-1}^{(2)}\} & \cdots & \{u_{k+n-1}^{(n)}\} \end{bmatrix}$$

is invertible for at least one value of k, then the vectors are linearly independent. Furthermore, if the vectors are solutions to the same homogeneous linear difference equation and the matrix is not invertible for any value of k, then the vectors are linearly dependent.
(3) The set of all solutions to the homogeneous linear difference equation

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0,$$

is an n-dimensional subspace of $\mathbb{S}$.
(4) Let $\{z_k\}$ be a vector in $\mathbb{S}$. If the nonhomogeneous linear difference equation

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k$$

is consistent, then the mapping $\{y_k\} \mapsto \{y_k + z_k\}$ is a bijection from the solution set of the homogeneous linear difference equation

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0$$

to the solution set of the nonhomogeneous equation.
(5) The solutions of the homogeneous linear difference equation

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0$$

are given by the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$, where

$$\mathbf{x}_k = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+n-1} \end{bmatrix}, A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & a_{n-2} & \cdots & -a_1 \end{bmatrix}.$$

(6) Call a vector a **probability vector** if its entries are nonnegative and sum to 1. Let P be a square matrix whose columns are probability vectors. Then there is a unique probability vector $\mathbf{q}$ such that $P\mathbf{q} = \mathbf{q}$. Furthermore, $\lim_{k\to\infty} A^k\mathbf{x} = \mathbf{q}$ for any probability vector $\mathbf{x}$.

(7) Let A be an $n\times n$ matrix with n linearly independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues, possibly with repetition, and let $c_1, \ldots, c_n$ be the coordinates of a vector $\mathbf{x}$ in the eigenvector basis. Then

$$A^k\mathbf{x} = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_n\lambda_n^k\mathbf{v}_n.$$

(8) Let A be a matrix. If $A = PDP^{-1}$, with D a diagonal matrix, then $A^k = PD^kP^{-1}$.

(9) If A is an $n\times n$ diagonalizable matrix with n linearly independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$, then the general solution to the system of differential equations given by $\mathbf{x}'(t) = A\mathbf{x}(t)$ is $\mathbf{x}(t) = c_1\mathbf{v}_1e^{\lambda_1 t} + \cdots + c_n\mathbf{v}_ne^{\lambda_n t}$.

(10) Suppose A is an $n\times n$ matrix, the general solution to the system of differential equations given by $\mathbf{x}'(t) = A\mathbf{x}(t)$ is an arbitrary linear combination of a collection of linearly independent eigenfunctions $\mathbf{y}_1(t), \ldots, \mathbf{y}_n(t)$, and two of those solutions, say $\mathbf{y}_1(t)$ and $\mathbf{y}_2(t)$, are complex conjugates. If $\mathbf{y}_1(t)$ and $\mathbf{y}_2(t)$ are replaced by either $\operatorname{Re}\mathbf{y}_1(t)$ and $\operatorname{Im}\mathbf{y}_1(t)$ or $\operatorname{Re}\mathbf{y}_2(t)$ and $\operatorname{Im}\mathbf{y}_2(t)$, the collection of eigenfunctions will still be linearly independent and an arbitrary linear combination of them will still give the general solution to the system.

(11) Let A be an $m \times n$ matrix, and let $\mathbf{b}$ be a vector in $\mathbb{R}^m$. Then $\widehat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if $A^TA\widehat{\mathbf{x}} = A^T\mathbf{b}$.

(12) Let A be an $m \times n$ matrix. Then the following statements are logically equivalent:
   (a) The equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b}$ in $\mathbb{R}^m$.
   (b) The columns of A are linearly independent.
   (c) The matrix A^TA is invertible.

   When these statements are true, A has a QR factorization and the least-squares solution is given by $\widehat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b} = R^{-1}Q^T\mathbf{b}$.