MATH 537, Fall 2020
Ordinary Differential Equations

Lecture #5

A Brief Review of Linear Algebra

Instructor: Dr. Bo-Wen Shen*

Department of Mathematics and Statistics
San Diego State University

Outline

- Fundamental Concepts
  - Eight Axioms; Subset vs. Subspace
- Basic Matrix Operations
  - Addition, Subtraction, Scalar Multiplication, Matrix Multiplication
- Matrix Properties
  - Rank, Range, Kernel, and Nullspace
- Main Types of Matrices
  - Matrix Transpose, Identity, and Inverse;
  - Symmetric, Skew-symmetric, and Orthogonal Matrices
- Linear Dependence/Independence and Linear Systems
- Elimination and LU Decomposition
  - Elementary Row Operations, Augmented Matrix, and Row Echelon Form
- Fundamental Theorem for Linear Systems
- Eigenvalue Problem
  - Similarity Transformation and Diagonalization
- Quadratic Forms and Rayleigh Quotient

References

1. Nagle, Saff, and Snider, Differential Equations, Chapter 9
2. Sadun, Applied Linear Algebra,
3. Erwin Kreyszig, Advanced Engineering Mathematics,, 10th edition.
4. Lloyd N. Trefethen and David Bau, III, Numerical Linear Algebra, Society for Industrial and Applied Mathematics (SIAM), 1997. ISBN 0-89871-361-7.
5. Gene H. Golub and Charles F. Van Loan, 2012, Matrix Computations, 4th Edition, Johns Hopkins University Press, 2012. ISBN-13 978-1421407944.

Linear Algebra (LA): Eight Axioms

Definition. *Let V be a set on which addition and scalar multiplication are defined. (That is, if $\mathbf{x}$ and $\mathbf{y}$ are elements of V, and c is a scalar, then $\mathbf{x}+\mathbf{y}$ and $c\mathbf{x}$ are elements of V.) If the following eight axioms are satisfied by all elements $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ of V and all scalars a and b, then V is called a* vector space *and the elements of V are called* vectors. *If these axioms apply to multiplication by real scalars, then V is called a* real vector space. *If the axioms apply to multiplication by complex scalars, then V is a* complex vector space.

(1) *Commutativity of addition:* $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$.

(2) *Associativity of addition:* $(\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z})$.

LA: Eight Axioms (.continued)

(3) *Additive identity:* *There exists a vector, denoted* $\mathbf{0}$*, such that, for every vector* $\mathbf{x}$, $\mathbf{0} + \mathbf{x} = \mathbf{x}$.

(4) *Additive inverses:* *For every vector* $\mathbf{x}$ *there exists a vector* $(-\mathbf{x})$ *such that* $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.

(5) *First distributive law:* $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$.

(6) *Second distributive law:* $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$.

(7) *Multiplicative identity:* $1(\mathbf{x}) = \mathbf{x}$.

(8) *Relation to ordinary multiplication:* $(ab)\mathbf{x} = a(b\mathbf{x})$.

Subset vs. Subspace

Definition. *A set S is closed under addition if the sum of any two elements of S is in S, and is closed under scalar multiplication if the product of an arbitrary scalar and an arbitrary element of S is in S.*

Frequently we consider a subset W of a vector space V. In this case, addition and scalar multiplication are already defined, and already satisfy the eight axioms. If W is closed under addition and scalar multiplication, then W is a vector space in its own right, and we call W a *subspace* of V.

subspace: closed under (1) addition and (2) scalar multiplication

Preview of Sect 5.1: Subspace

given $V_1, \ldots, V_k \in \mathbb{R}^n$, the set

$$\mathcal{S} = \{\alpha_1 V_1 + \cdots + \alpha_k V_k \mid \alpha_j \in \mathbb{R}\}$$

is a subspace of $\mathbb{R}^n$. In this case we say that $\mathcal{S}$ is *spanned* by $V_1, \ldots, V_k$. Equivalently, it can be shown (see Exercise 12 at the end of this chapter) that a subspace $\mathcal{S}$ is a nonempty subset of $\mathbb{R}^n$ having the following two properties:

1. If $X, Y \in \mathcal{S}$, then $X + Y \in \mathcal{S}$;
2. If $X \in \mathcal{S}$ and $\alpha \in \mathbb{R}$, then $\alpha X \in \mathcal{S}$.

subspace: closed under (1) addition and (2) scalar multiplication

Subset vs. Subspace: Examples

(1) Let $W = \{\mathbf{x} \in \mathbb{R}^2 | x_1 + x_2 = 0\}$. The sum of any two vectors in W is in W, and any scalar multiple of a vector in W is in W. (Check this!) W is a subspace of the vector space $\mathbb{R}^2$.

(2) More generally, let A be any fixed $m \times n$ matrix and let $W = \{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} = 0\}$. W is closed under addition and scalar multiplication (why?), so W is a subspace of $\mathbb{R}^n$.

(3) Let $\mathbb{R}[t]$ be the set of all real-valued polynomials in a fixed variable t. Polynomials are continuous functions on $[0, 1]$, so $\mathbb{R}[t]$ is a subset of $C^0[0, 1]$. The sum of polynomials is a polynomial, as is the product of a scalar and a polynomial, so $\mathbb{R}[t]$ is a subspace of $C^0[0, 1]$.

(4) Let $\mathbb{R}_n[t]$ be the set of real polynomials of degree n or less. For $n < m$, $\mathbb{R}_n[t]$ is a subspace of $\mathbb{R}_m[t]$, and $\mathbb{R}_n[t]$ is always a subspace of $\mathbb{R}[t]$.

(5) Instead of considering polynomials with real coefficients, we could consider polynomials with complex coefficients to get examples of complex vector spaces. The space of polynomials with complex coefficients is usually denoted $\mathbb{C}[t]$.

(6) Let $W' = \{\mathbf{x} \in \mathbb{R}^2 | x_1 + x_2 = 1\}$. W' is not a vector space, as the sum of two elements of W', or a scalar multiple of an element of W', is typically not in W'. (Again, check this!)

(7) Let $\mathbb{Z}^2 = \{\mathbf{x} \in \mathbb{R}^2 | x_1 \text{ and } x_2 \text{ are integers }\}$. $\mathbb{Z}^2$ is closed under addition, but not under scalar multiplication. $\mathbb{Z}^2$ is a sub*set* of $\mathbb{R}^2$, but not a sub*space*.

Subset vs. Subspace: Examples (.continued)

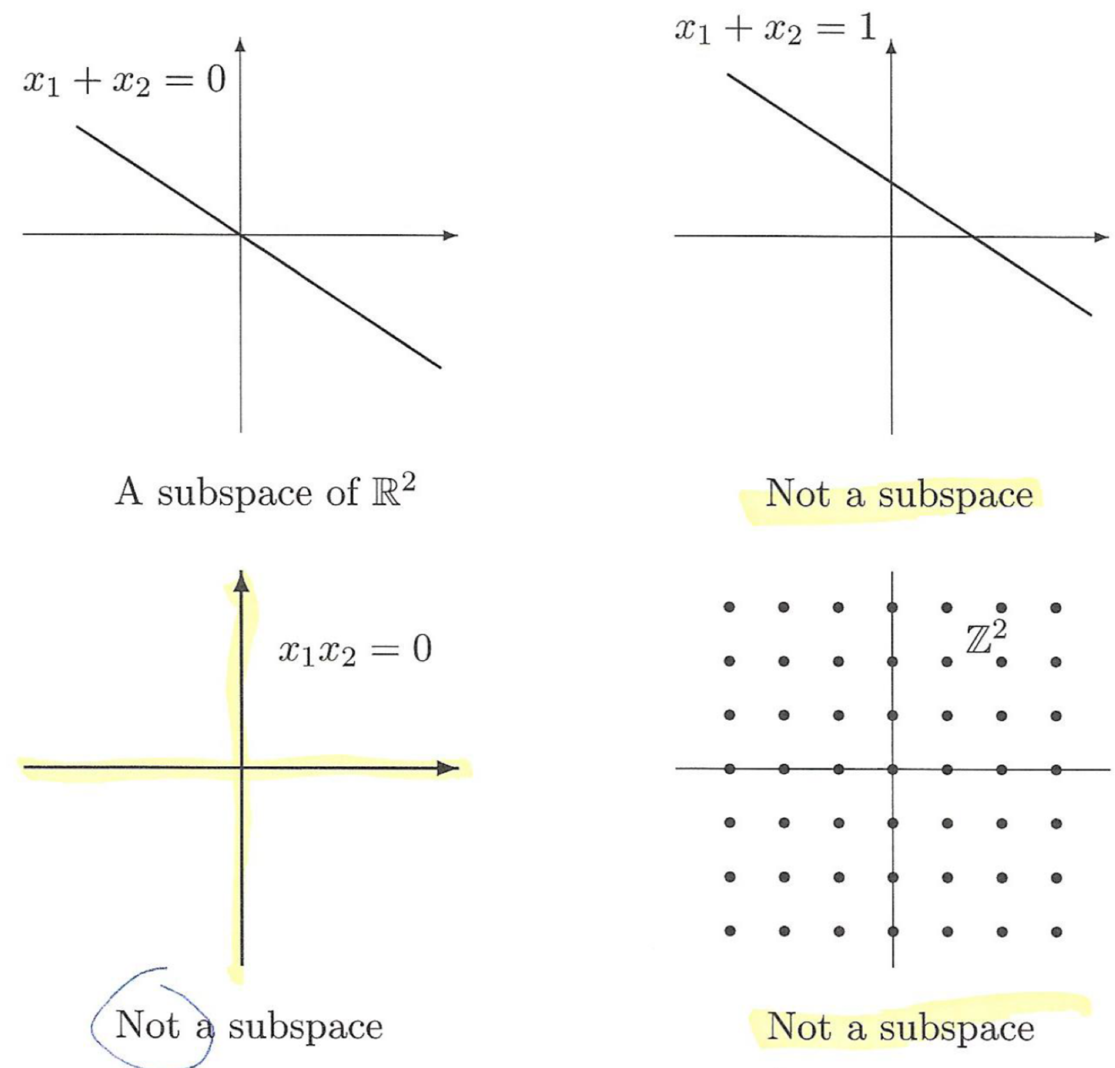

Figure 2.1. Four subsets of $\mathbb{R}^2$

Outline

- Fundamental Concepts
  - Eight Axioms; Subset vs. Subspace
- Basic Matrix Operations
  - Addition, Subtraction, Scalar Multiplication, Matrix Multiplication
- Matrix Properties
  - Rank, Range, Kernel, and Nullspace
- Main Types of Matrices
  - Matrix Transpose, Identity, and Inverse;
  - Symmetric, Skew-symmetric, and Orthogonal Matrices
- Linear Dependence/Independence and Linear Systems
- Elimination and LU Decomposition
  - Elementary Row Operations, Augmented Matrix, and Row Echelon Form
- Fundamental Theorem for Linear Systems
- Eigenvalue Problem
  - Similarity Transformation and Diagonalization
- Quadratic Forms and Rayleigh Quotient

1. Addition/Subtraction

Two Matrices have the same dimension

$$\underset{3\times 2}{\mathbf{A}} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \qquad \underset{3\times 2}{\mathbf{B}} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$$

3 rows
2 columns

$$\underset{3\times 2}{\mathbf{A} + \mathbf{B}} = \begin{bmatrix} 1+1 & 4+2 \\ 2+2 & 5+3 \\ 3+3 & 6+4 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ 4 & 8 \\ 6 & 10 \end{bmatrix}$$

2. Scalar Multiplication

To multiply a matrix by a scalar (number), we simply multiply each element in the matrix by the number:

$$3\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{bmatrix}.$$

In other words, $r\mathbf{A} = r[a_{ij}] = [ra_{ij}]$. The notation $-\mathbf{A}$ stands for $(-1)\mathbf{A}$.

Properties of Matrix Addition and Scalar Multiplication

Properties of Matrix Addition and Scalar Multiplication. Matrix addition and scalar multiplication are nothing more than mere bookkeeping, and the usual algebraic properties hold. If **A, B,** and **C** are $m \times n$ matrices and r, s are scalars, then

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}\ , \qquad \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}\ ,$$

$$\mathbf{A} + \mathbf{0} = \mathbf{A}\ , \qquad \mathbf{A} + (-\mathbf{A}) = \mathbf{0}\ ,$$

$$r(\mathbf{A} + \mathbf{B}) = r\mathbf{A} + r\mathbf{B}\ , \qquad (r + s)\mathbf{A} = r\mathbf{A} + s\mathbf{A}\ ,$$

$$r(s\mathbf{A}) = (rs)\mathbf{A} = s(r\mathbf{A})\ .$$

3. Matrix Multiplication

$$\begin{array}{ll} & \mathbf{A} \\ \begin{array}{l}\text{Row 1} \\ \text{Row 2}\end{array} & \begin{bmatrix} \boxed{2 \quad 5} \\ 4 \quad 1 \end{bmatrix} \end{array} \qquad \begin{array}{c} \mathbf{B} \\ \begin{bmatrix} \boxed{\begin{matrix}4\\5\end{matrix}} & \begin{matrix}6\\8\end{matrix} \end{bmatrix} \\ \text{Col. 1 Col. 2} \end{array} \qquad \begin{array}{ll} & \mathbf{AB} \\ \boxed{\text{Row 1}} & \begin{bmatrix} 33 & \\ & \end{bmatrix} \\ & \boxed{\text{Col. 1}} \end{array}$$

We take the cross products and sum:

$$2(4) + 5(5) = 33$$

$$\begin{array}{ll} & \mathbf{A} \\ \begin{array}{l}\text{Row 1} \\ \text{Row 2}\end{array} & \begin{bmatrix} \boxed{2 \quad 5} \\ 4 \quad 1 \end{bmatrix} \end{array} \qquad \begin{array}{c} \mathbf{B} \\ \begin{bmatrix} \begin{matrix}4\\5\end{matrix} & \boxed{\begin{matrix}6\\8\end{matrix}} \end{bmatrix} \\ \text{Col. 1 Col. 2} \end{array} \qquad \begin{array}{ll} & \mathbf{AB} \\ \text{Row 1} & \begin{bmatrix} 33 & 52 \\ & \end{bmatrix} \\ & \text{Col. 1 Col. 2} \end{array}$$

3. Matrix Multiplication

More generally, the product of two matrices **A** and **B** is formed by taking the array of dot products of the *rows* of the first "factor" **A** with the *columns* of the second factor **B**; the dot product of the ith row of **A** with the jth column of **B** is written as the ijth entry of the product **AB**:

$$\begin{array}{l} \text{row} \\ \end{array} \begin{bmatrix} 1 & 0 & 1 \\ 3 & -1 & 2 \end{bmatrix} \begin{array}{c} \begin{array}{ccc} & \text{column} & \end{array} \\ \begin{bmatrix} 1 & 2 & x \\ -1 & -1 & y \\ 4 & 1 & z \end{bmatrix} \end{array} = \begin{bmatrix} 1+0+4 & 2+0+1 & x+0+z \\ 3+1+8 & 6+1+2 & 3x-y+2z \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 3 & x+z \\ 12 & 9 & 3x-y+2z \end{bmatrix}.$$

Note that **AB** is only defined when the number of columns of **A** matches the number of rows of **B**. A useful formula for the product of an $m \times n$ matrix **A** and an $n \times p$ matrix **B** is

$$\mathbf{AB} := [c_{ij}], \qquad \text{where} \quad c_{ij} := \sum_{k=1}^{n} a_{ik} b_{kj} \ .$$

Matrix Multiplication

Equal

$$\underset{2\times 2}{\mathbf{A}} \quad \underset{2\times 2}{\mathbf{B}} = \underset{2\times 2}{\mathbf{AB}}$$

Dimension of product

Equal

$$\underset{2\times 3}{\mathbf{A}} \quad \underset{3\times 1}{\mathbf{B}} = \underset{2\times 1}{\mathbf{AB}}$$

Dimension of product

$$\underset{2\times 2}{\mathbf{A}} \quad \underset{2\times 1}{\mathbf{X}} = \underset{2\times 1}{\mathbf{AX}}$$

$$\underset{3\times 3}{\mathbf{A}} \quad \underset{3\times 1}{\mathbf{X}} = \underset{3\times 1}{\mathbf{AX}}$$

In a Matrix Form

Consider the following two lines in R^2

$$ax + by = \alpha$$
$$cx + dy = \beta,$$

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $\quad X = \begin{pmatrix} x \\ y \end{pmatrix}$

$$AX = \begin{pmatrix} \boxed{a \quad b} \\ c \quad d \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} x \\ y \end{matrix}} \end{pmatrix} = \begin{pmatrix} ax + by \\ \end{pmatrix}$$

$$AX = \begin{pmatrix} a \quad b \\ \boxed{c \quad d} \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} x \\ y \end{matrix}} \end{pmatrix} = \begin{pmatrix} \\ cx + dy \end{pmatrix}$$

Thus, $AX = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$

$$AX = r \qquad r = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

Properties of Matrix Multiplication

Properties of Matrix Multiplication

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) \qquad \text{(Associativity)}$$

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC} \qquad \text{(Distributivity)}$$

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \qquad \text{(Distributivity)}$$

$$(r\mathbf{A})\mathbf{B} = r(\mathbf{AB}) = \mathbf{A}(r\mathbf{B}) \qquad \text{(Associativity)}$$

Outline

- Fundamental Concepts
  - Eight Axioms; Subset vs. Subspace
- Basic Matrix Operations
  - Addition, Subtraction, Scalar Multiplication, Matrix Multiplication
- Matrix Properties
  - Rank, Range, Kernel, and Nullspace
- Main Types of Matrices
  - Matrix Transpose, Identity, and Inverse;
  - Symmetric, Skew-symmetric, and Orthogonal Matrices
- Linear Dependence/Independence and Linear Systems
- Elimination and LU Decomposition
  - Elementary Row Operations, Augmented Matrix, and Row Echelon Form
- Fundamental Theorem for Linear Systems
- Eigenvalue Problem
  - Similarity Transformation and Diagonalization
- Quadratic Forms and Rayleigh Quotient

Column Space

Example 1 Multiply A times $\boldsymbol{x}$ using the three rows of A. Then use the two columns:

By rows $\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ 2x_1 + 4x_2 \\ 3x_1 + 7x_2 \end{bmatrix}$ = inner products of the rows with $\boldsymbol{x} = (x_1, x_2)$

By columns $\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}$ = combination of the columns $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$

Thus $A\boldsymbol{x}$ is a linear combination of the columns of A. This is fundamental.

Definition **The combinations of the columns** fill out the **column space** of A.

Column Space: An Illustration

Definition **The combinations of the columns** fill out the **column space** of A.

Example 3 What are the column spaces of $A_2 = \begin{bmatrix} 2 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 10 \end{bmatrix}$ and $A_3 = \begin{bmatrix} 2 & 3 & 1 \\ 2 & 4 & 1 \\ 3 & 7 & 1 \end{bmatrix}$?

Solution. The column space of A_2 is the same plane as before. The new column $(5, 6, 10)$ is the sum of column 1 + column 2. So $\boldsymbol{a}_3$ = column 3 is already in the plane and adds nothing new. By including this "*dependent*" column we don't go beyond the original plane.

The column space of A_3 is the whole 3D space $\mathbf{R}^3$. Example 2 showed us that the new third column $(1, 1, 1)$ is not in the plane $\mathbf{C}(A)$. Our column space $\mathbf{C}(A_3)$ has grown bigger. But there is nowhere to stop between a plane and the full 3D space. Visualize the $x - y$ plane and a third vector (x_3, y_3, z_3) out of the plane (meaning that $z_3 \neq 0$). They combine to give **every vector in** $\mathbf{R}^3$.

Column Space and Row Space

A **basis** for a subspace is a full set of independent vectors: **All vectors in the space are combinations of the basis vectors.** Examples will make the point.

Definition **The combinations of the columns** fill out the **column space** of A.

The rank of a matrix is the dimension of its column space.

The number of *independent columns* equals the number of *independent rows*

Rank

Rank

The *column rank* of a matrix is the dimension of its column space. Similarly, the *row rank* of a matrix is the dimension of the space spanned by its rows. Row rank always equals column rank (among other proofs, this is a corollary of the singular value decomposition, discussed in Lectures 4 and 5), so we refer to this number simply as the *rank* of a matrix.

An $m \times n$ matrix of *full rank* is one that has the maximal possible rank (the lesser of m and n). This means that a matrix of full rank with $m \geq n$ must have n linearly independent columns. Such a matrix can also be characterized by the property that the map it defines is one-to-one.

- Rank: the number of independent vectors
- Column rank = Row rank
- Full rank: e.g., rank =n for a n x n Matrix

Row Space and Column Space

Row Space and Column Space

The row space and the column space of a matrix $\mathbf{A}$ *have the same dimension, equal to rank* $\mathbf{A}$.

Null Space (Ax=0)

Finally, for a given matrix $\mathbf{A}$ the solution set of the homogeneous system $\mathbf{Ax} = 0$ is a vector space, called the **null space** of $\mathbf{A}$, and its dimension is called the **nullity** of $\mathbf{A}$. In the next section we motivate and prove the basic relation

$$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = \text{Number of columns of } \mathbf{A}. \tag{6}$$

Kreyszig (2011)

Range and Nullspace

The *range of a matrix* A, written range(A), is the set of vectors that can be expressed as Ax for some x. The formula (1.2) leads naturally to the following characterization of range(A).

Theorem 1.1. range(A) *is the space spanned by the columns of* A.

Proof. By (1.2), any Ax is a linear combination of the columns of A. Conversely, any vector y in the space spanned by the columns of A can be written as a linear combination of the columns, $y = \sum_{j=1}^{n} x_j a_j$. Forming a vector x out of the coefficients x_j, we have $y = Ax$, and thus y is in the range of A. □

Nullspace

The *nullspace* of $A \in \mathbb{C}^{m\times n}$, written null$(A)$, is the set of vectors x that satisfy $Ax = 0$, where 0 is the 0-vector in $\mathbb{C}^m$. The entries of each vector $x \in$ null(A) give the coefficients of an expansion of zero as a linear combination of columns of A: $0 = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$.

Special case with $|A| \neq 0$

$$Given\ \boldsymbol{Ax} = 0$$

$$\boldsymbol{x} = 0\ when\ |A| \neq 0\ (trivial\ solutions\ only)$$

$$null\ (A) = \{0\}$$

Kernel vs. Nullspace

The terminology "kernel" and "nullspace" refer to the same concept, in the context of vector spaces and linear transformations. It is more common in the literature to use the word nullspace when referring to a matrix and the word kernel when referring to an abstract linear transformation. However, using either word is valid. Note that a matrix is a linear transformation from one coordinate vector space to another. Additionally, the terminology "kernel" is used extensively to denote the analogous concept as for linear transformations for morphisms of various other algebraic structures, e.g. groups, rings, modules and in fact we have a definition of kernel in the very abstract context of abelian categories.

Rank A + Nullity A = n

The solution space of (4) is also called the **null space** of $\mathbf{A}$ because $\mathbf{Ax} = \mathbf{0}$ for every $\mathbf{x}$ in the solution space of (4). Its dimension is called the **nullity** of $\mathbf{A}$. Hence Theorem 2 states that

$$\text{rank } \mathbf{A} + \text{nullity } \mathbf{A} = n \tag{5}$$

where n is the number of unknowns (number of columns of $\mathbf{A}$).

Furthermore, by the definition of rank we have rank $\mathbf{A} \leqq m$ in (4). Hence if $m < n$, then rank $\mathbf{A} < n$. By Theorem 2 this gives the practically important

Kreyszig (2011)
Section 5.4

Determinants. For a 2×2 matrix $\mathbf{A}$, the **determinant of A**, denoted $\det \mathbf{A}$ or $|\mathbf{A}|$, is defined by

$$\det \mathbf{A} := \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \ .$$

We can define the determinant of a 3×3 matrix $\mathbf{A}$ in terms of its cofactor expansion about the first row; that is,

$$\det \mathbf{A} := \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \ .$$

For example,

$$\begin{vmatrix} 1 & 2 & 1 \\ 0 & 3 & 5 \\ 2 & 1 & -1 \end{vmatrix} = 1 \begin{vmatrix} 3 & 5 \\ 1 & -1 \end{vmatrix} - 2 \begin{vmatrix} 0 & 5 \\ 2 & -1 \end{vmatrix} + 1 \begin{vmatrix} 0 & 3 \\ 2 & 1 \end{vmatrix}$$

$$= 1(-3 - 5) - 2(0 - 10) + 1(0 - 6) = 6 \ .$$

Third-Order Determinants

A **determinant of third order** can be defined by

$$(4) \quad D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}.$$

Rank in terms of Determinants

Rank in Terms of Determinants

Consider an $m \times n$ *matrix* $\mathbf{A} = [a_{jk}]$:

(1) $\mathbf{A}$ *has rank* $r \geqq 1$ *if and only if* $\mathbf{A}$ *has an* $r \times r$ *submatrix with a nonzero determinant.*

(2) *The determinant of any square submatrix with more than r rows, contained in* $\mathbf{A}$ *(if such a matrix exists!) has a value equal to zero.*

Furthermore, if $m = n$, *we have:*

(3) *An* $n \times n$ *square matrix* $\mathbf{A}$ *has rank n if and only if*

$$\det \mathbf{A} \neq 0.$$

full rank

Cramer's Rule using Determinants

Cramer's Rule for Linear Systems of Three Equations

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3
\end{aligned}
\tag{5}
$$

is

$$
x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad x_3 = \frac{D_3}{D} \qquad (D \neq 0) \tag{6}
$$

with the *determinant D of the system* given by (4) and

$$
D_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad D_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \quad D_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}.
$$

Note that D_1, D_2, D_3 are obtained by replacing Columns 1, 2, 3, respectively, by the column of the right sides of (5).

Outline

- Fundamental Concepts
  - Eight Axioms; Subset vs. Subspace
- Basic Matrix Operations
  - Addition, Subtraction, Scalar Multiplication, Matrix Multiplication
- Matrix Properties
  - Rank, Range, Kernel, and Nullspace
- Main Types of Matrices
  - Matrix Transpose, Identity, and Inverse;
  - Symmetric, Skew-symmetric, and Orthogonal Matrices
- Linear Dependence/Independence and Linear Systems
- Elimination and LU Decomposition
  - Elementary Row Operations, Augmented Matrix, and Row Echelon Form
- Fundamental Theorem for Linear Systems
- Eigenvalue Problem
  - Similarity Transformation and Diagonalization
- Quadratic Forms and Rayleigh Quotient

Matrix Transpose: A^T

Matrix Transpose. The matrix obtained from **A** by interchanging its rows and columns is called the **transpose** of **A** and is denoted by $\mathbf{A}^T$. For example, if

$$\mathbf{A} = \begin{bmatrix} \boxed{1 \quad 2 \quad 6} \\ -1 \quad 2 \quad -1 \end{bmatrix}, \quad \text{then}$$

$$\mathbf{A}^T = \begin{bmatrix} 1 & -1 \\ 2 & 2 \\ 6 & -1 \end{bmatrix}.$$

In general, we have $[a_{ij}]^T = [b_{ij}]$, where $b_{ij} = a_{ji}$. Properties of the transpose are explored in Problem 7.

Matrix Identity: I

$$IA = A = AI$$

Matrix Identity. There is a "multiplicative identity" in matrix algebra, namely, a square diagonal matrix **I** with ones down the main diagonal. Multiplying **I** on the right or left by any other matrix (with compatible dimensions) reproduces the latter matrix:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

(The notation $\mathbf{I}_n$ is used if it is convenient to specify the dimensions, $n \times n$, of the identity matrix.)

Matrix Inverse: $AA^{-1}=I$

Matrix Inverse. Some *square* matrices $\mathbf{A}$ can be paired with other (square) matrices $\mathbf{B}$ having the property that $\mathbf{BA} = \mathbf{I}$:

(2) $$\begin{bmatrix} \frac{3}{2} & -1 & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{3}{2} & 1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

When this happens, it can be shown that

(i) $\mathbf{B}$ is the *unique* matrix satisfying $\mathbf{BA} = \mathbf{I}$, and
(ii) $\mathbf{B}$ also satisfies $\mathbf{AB} = \mathbf{I}$.

In such a case, we say that $\mathbf{B}$ is the **inverse** of $\mathbf{A}$ and write $\mathbf{B} = \mathbf{A}^{-1}$.

Not every matrix possesses an inverse; the zero matrix $\mathbf{0}$, for example, can never satisfy the equation $\mathbf{0B} = \mathbf{I}$. A matrix that has no inverse is said to be **singular.**

Matrix Inverse: AA^{-1}=I

Inverse

A *nonsingular* or *invertible* matrix is a square matrix of full rank. Note that the m columns of a nonsingular $m \times m$ matrix A form a basis for the whole space $\mathbb{C}^m$. Therefore, we can uniquely express any vector as a linear combination of them. In particular, the canonical unit vector with 1 in the jth entry and zeros elsewhere, written e_j, can be expanded:

$$e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

AZ=I

where I is the $m \times m$ matrix known as the *identity*. The matrix Z is the *inverse* of A. Any square nonsingular matrix A has a unique inverse, written A^{-1}, that satisfies $AA^{-1} = A^{-1}A = I$.

Inverse of a Matrix by Determinants

Inverse of a Matrix by Determinants

The inverse of a nonsingular $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ is given by

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}}[C_{jk}]^{\mathsf{T}} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \cdot & \cdot & \cdots & \cdot \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}, \tag{4}$$

where C_{jk} is the cofactor of a_{jk} in $\det \mathbf{A}$ (see Sec. 7.7). (CAUTION! Note well that in $\mathbf{A}^{-1}$, the cofactor C_{jk} occupies the same place as a_{kj} (not a_{jk}) does in $\mathbf{A}$.)

In particular, the inverse of

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \textit{is} \quad \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}. \tag{4*}$$

Symmetric

Real Symmetric Matrices

Definition 4. A **real symmetric matrix A** is a matrix with real entries that satisfies $\mathbf{A}^T = \mathbf{A}$.

Symmetric, Skew-symmetric, and Orthogonal Matrices

Symmetric, Skew-Symmetric, and Orthogonal Matrices

A ***real*** square matrix $\mathbf{A} = [a_{jk}]$ is called

symmetric if transposition leaves it unchanged,

(1) $$\boxed{\mathbf{A}^\mathsf{T} = \mathbf{A},} \qquad \text{thus} \qquad a_{kj} = a_{jk},$$

skew-symmetric if transposition gives the negative of $\mathbf{A}$,

(2) $$\boxed{\mathbf{A}^\mathsf{T} = -\mathbf{A},} \qquad \text{thus} \qquad a_{kj} = -a_{jk},$$

orthogonal if transposition gives the inverse of $\mathbf{A}$,

(3) $$\boxed{\mathbf{A}^\mathsf{T} = \mathbf{A}^{-1}.}$$

Outline

- Fundamental Concepts
  - Eight Axioms; Subset vs. Subspace
- Basic Matrix Operations
  - Addition, Subtraction, Scalar Multiplication, Matrix Multiplication
- Matrix Properties
  - Rank, Range, Kernel, and Nullspace
- Main Types of Matrices
  - Matrix Transpose, Identity, and Inverse;
  - Symmetric, Skew-symmetric, and Orthogonal Matrices
- Linear Dependence/Independence and Linear Systems
- Elimination and LU Decomposition
  - Elementary Row Operations, Augmented Matrix, and Row Echelon Form
- Fundamental Theorem for Linear Systems
- Eigenvalue Problem
  - Similarity Transformation and Diagonalization
- Quadratic Forms and Rayleigh Quotient

Linear Dependence

Linear Dependence of Vector Functions

Definition 1. The m vector functions $\mathbf{x}_1, \ldots, \mathbf{x}_m$ are said to be **linearly dependent on an interval** I if there exist constants $c_1, \ldots, c_m$, not all zero, such that

$$c_1\mathbf{x}_1(t) + \cdots + c_m\mathbf{x}_m(t) = \mathbf{0} \tag{3}$$

for all t in I. If the vectors are not linearly dependent, they are said to be **linearly independent on** I.

- at least one of the vectors can be defined as a linear combination of the others

Example 1 Show that the vector functions $\mathbf{x}_1(t) = \text{col}(e^t, 0, e^t)$, $\mathbf{x}_2(t) = \text{col}(3e^t, 0, 3e^t)$, and $\mathbf{x}_3(t) = \text{col}(t, 1, 0)$ are linearly dependent on $(-\infty, \infty)$.

Solution Notice that $\mathbf{x}_2$ is just 3 times $\mathbf{x}_1$ and therefore $3\mathbf{x}_1(t) - \mathbf{x}_2(t) + 0 \cdot \mathbf{x}_3(t) = \mathbf{0}$ for all t. Hence, $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ are linearly dependent on $(-\infty, \infty)$. ◆

Linear Independence

Example 3 Show that the vector functions $\mathbf{x}_1(t) = \text{col}(e^{2t}, 0, e^{2t})$, $\mathbf{x}_2(t) = \text{col}(e^{2t}, e^{2t}, -e^{2t})$, and $\mathbf{x}_3(t) = \text{col}(e^t, 2e^t, e^t)$ are linearly independent on $(-\infty, \infty)$.

Solution To prove independence, we *assume* c_1, c_2, and c_3 are constants for which

$$c_1\mathbf{x}_1(t) + c_2\mathbf{x}_2(t) + c_3\mathbf{x}_3(t) = \mathbf{0}$$

holds at every t in $(-\infty, \infty)$ and show that this forces $c_1 = c_2 = c_3 = 0$. In particular, when $t = 0$ we obtain

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \mathbf{0} \ ,$$

which is equivalent to the system of linear equations

$$\begin{aligned} c_1 + c_2 + c_3 &= 0 \ , \\ c_2 + 2c_3 &= 0 \ , \\ c_1 - c_2 + c_3 &= 0 \ . \end{aligned} \tag{4}$$

Either by solving (4) or by checking that the determinant of its coefficients is nonzero (recall Theorem 1 on page 513), we can verify that (4) has only the trivial solution $c_1 = c_2 = c_3 = 0$. Therefore the vector functions $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ are linearly independent on $(-\infty, \infty)$ (in fact, on any interval containing $t = 0$). ◆

Wronskian: $|A|$

is not zero. Because of the analogy with scalar equations, we call this determinant the **Wronskian**.

Wronskian

Definition 2. The **Wronskian** of n vector functions $\mathbf{x}_1(t) = \text{col}(x_{1,1}, \ldots, x_{n,1}), \ldots, \mathbf{x}_n(t) = \text{col}(x_{1,n}, \ldots, x_{n,n})$ is defined to be the real-valued function

$$W[\mathbf{x}_1, \ldots, \mathbf{x}_n](t) := \begin{vmatrix} x_{1,1}(t) & x_{1,2}(t) & \cdots & x_{1,n}(t) \\ x_{2,1}(t) & x_{2,2}(t) & \cdots & x_{2,n}(t) \\ \vdots & \vdots & & \vdots \\ x_{n,1}(t) & x_{n,2}(t) & \cdots & x_{n,n}(t) \end{vmatrix} .$$

Wronskian for n Functions

DEFINITION 4.1.2 Wronskian

Suppose each of the functions $f_1(x), f_2(x), \ldots, f_n(x)$ possesses at least $\boxed{n-1}$ derivatives. The determinant

$$W(f_1, f_2, \ldots, f_n) = \begin{vmatrix} f_1 & f_2 & \cdots & f_n \\ f_1' & f_2' & \cdots & f_n' \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)} & f_2^{(n-1)} & \cdots & f_n^{(n-1)} \end{vmatrix},$$

where the primes denote derivatives, is called the **Wronskian** of the functions.

Example: Are y_1 and y_2 Linearly Independent?

Linear Dependence and Independence of Solutions

Let the ODE (1) *have continuous coefficients* $p(x)$ *and* $q(x)$ *on an open interval* I. *Then two solutions* y_1 *and* y_2 *of* (1) *on* I *are linearly dependent on* I *if and only if their* "**Wronskian**"

(6) $$W(y_1, y_2) = y_1 y_2' - y_2 y_1'$$

is 0 *at some* x_0 *in* I. *Furthermore, if* $W = 0$ *at an* $x = x_0$ *in* I, *then* $W = 0$ *on* I; *hence, if there is an* x_1 *in* I *at which* W *is not* 0, *then* y_1, y_2 *are linearly independent on* I.

$$W(y_1, y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}$$

Linear Systems of Equations

A **linear system of *m* equations in *n* unknowns** $x_1, \cdots, x_n$ is a set of equations of the form

$$
\text{(1)} \qquad
\begin{aligned}
a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots&\cdots \\
a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}
$$

Kreyszig (2011)

Matrix Form of the Linear System

Matrix Form of the Linear System (1). From the definition of matrix multiplication we see that the m equations of (1) may be written as a single vector equation

$$\mathbf{Ax} = \mathbf{b} \tag{2}$$

where the **coefficient matrix** $\mathbf{A} = [a_{jk}]$ is the $m \times n$ matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

are column vectors. We assume that the coefficients a_{jk} are not all zero, so that $\mathbf{A}$ is not a zero matrix. Note that $\mathbf{x}$ has n components, whereas $\mathbf{b}$ has m components. The

Kreyszig (2011)

LA Systems Ax=b with a Square Matrix

In general, we express the linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \ , \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \ , \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

in matrix notation as $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is the *coefficient matrix*, $\mathbf{x}$ is the vector of unknowns, and $\mathbf{b}$ is the vector of constants occurring on the right-hand side:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

If $\mathbf{b} = \mathbf{0}$, the system $\mathbf{Ax} = \mathbf{b}$ is said to be *homogeneous* (analogous to the nomenclature of Section 4.2).

The Matrix Formulation of LA Systems: An Example

The Matrix Formulation of Linear Algebraic Systems. Matrix algebra was developed to provide a convenient tool for expressing and analyzing linear algebraic systems. Note that the set of equations

$$
\begin{aligned}
x_1 + 2x_2 + x_3 &= 1 \ , \\
x_1 + 3x_2 + 2x_3 &= -1 \ , \\
x_1 \phantom{{}+ 3x_2} + x_3 &= 0
\end{aligned}
$$

can be written using the matrix product

$$
\textbf{(1)} \qquad \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} .
$$

A x = b

Outline

- Fundamental Concepts
  - Eight Axioms; Subset vs. Subspace
- Basic Matrix Operations
  - Addition, Subtraction, Scalar Multiplication, Matrix Multiplication
- Matrix Properties
  - Rank, Range, Kernel, and Nullspace
- Main Types of Matrices
  - Matrix Transpose, Identity, and Inverse;
  - Symmetric, Skew-symmetric, and Orthogonal Matrices
- Linear Dependence/Independence and Linear Systems
- Elimination and LU Decomposition
  - Elementary Row Operations, Augmented Matrix, and Row Echelon Form
- Fundamental Theorem for Linear Systems
- Eigenvalue Problem
  - Similarity Transformation and Diagonalization
- Quadratic Forms and Rayleigh Quotient

Elementary Row Operations. Row-Equivalent Systems

Elementary Row Operations for Matrices:

Interchange of two rows

Addition of a constant multiple of one row to another row

*Multiplication of a row by a **nonzero** constant c*

CAUTION! These operations are for rows, ***not for columns***! They correspond to the following

- elementary row operations on the augmented matrix
- $Ax=\lambda x$

Kreyszig (2011)

Elementary Row Operations

Behavior of an *n*th-Order Determinant under Elementary Row Operations

(a) *Interchange of two rows multiplies the value of the determinant by* -1.

(b) *Addition of a multiple of a row to another row does not alter the value of the determinant.*

(c) *Multiplication of a row by a nonzero constant c multiplies the value of the determinant by c.* (This holds also when $c = 0$, but no longer gives an elementary row operation.)

Gaussian Elimination (wiki)

- In linear algebra, **Gaussian elimination** (also known as **row reduction**) is an algorithm for solving systems of linear equations.
- It is usually understood as a sequence of operations performed on the corresponding matrix of coefficients.
- This method can also be used to find the rank of a matrix, to calculate the determinant of a matrix, and to calculate the inverse of an invertible square matrix.
- The method is named after Carl Friedrich Gauss (1777–1855), although it was known to Chinese mathematicians as early as 179 CE (see History section).

Using row operations to convert a matrix into reduced row echelon form is sometimes called Gauss–Jordan elimination. Some authors use the term Gaussian elimination to refer to the process until it has reached its upper triangular, or (non-reduced) row echelon form.

Gauss-Jordan Algorithm: (Gaussian Elimination)

Example 1 Solve the system

coupled

$$\begin{aligned} 2x_1 + 6x_2 + 8x_3 &= 16 , \\ 4x_1 + 15x_2 + 19x_3 &= 38 , \\ 2x_1 \qquad\quad + 3x_3 &= 6 . \end{aligned}$$

eliminate x_1 in the 2nd and 3rd Eqs

Solution By subtracting 2 times the first equation from the second, we eliminate x_1 from the latter. Similarly, x_1 is eliminated from the third equation by subtracting 1 times the first equation from it:

$$\begin{aligned} 2x_1 + 6x_2 + 8x_3 &= 16 , \\ 3x_2 + 3x_3 &= 6 , \\ -6x_2 - 5x_3 &= -10 . \end{aligned}$$

eliminate x_2 in the 1st and 3rd Eqs

Next we subtract multiples of the second equation from the first and third to eliminate x_2 in them; the appropriate multiples are 2 and -2, respectively:

(1) – 2*(2)

(3) + 2*(2)

$$\begin{aligned} 2x_1 \qquad + 2x_3 &= 4 , \\ 3x_2 + 3x_3 &= 6 , \\ x_3 &= 2 . \end{aligned}$$

eliminate x_3 in the 1st and 2nd Eqs

Finally, we eliminate x_3 from the first two equations by subtracting multiples (2 and 3, respectively) of the third equation:

uncoupled

$$\begin{aligned} 2x_1 &= 0 , \\ 3x_2 &= 0 , \\ x_3 &= 2 . \end{aligned}$$

The system is now uncoupled; i.e., we can solve each equation separately:

$$x_1 = 0 , \quad x_2 = 0 , \quad x_3 = 2 . \ \blacklozenge$$

Nagle, et al. (2011)

(see details below)

Back Substitution

Gauss Elimination and Back Substitution

The Gauss elimination method can be motivated as follows. Consider a linear system that is in ***triangular form*** (in full, $\boxed{\textit{upper}\text{ triangular form}}$) such as

$$\begin{aligned} 2x_1 + 5x_2 &= 2 \\ 13x_2 &= -26 \end{aligned} \qquad\qquad A = \begin{bmatrix} 2 & 5 \\ 0 & 13 \end{bmatrix}$$

(*Triangular* means that all the nonzero entries of the corresponding coefficient matrix lie above the diagonal and form an upside-down 90° triangle.) Then we can solve the system by **back substitution**, that is, we solve the last equation for the variable, $x_2 = -26/13 = -2$, and then work backward, substituting $x_2 = -2$ into the first equation and solving it for x_1, obtaining $x_1 = \frac{1}{2}(2 - 5x_2) = \frac{1}{2}(2 - 5 \cdot (-2)) = 6$. This gives us the idea of first reducing a general system to triangular form. For instance, let the given system be

Kreyszig (2011)

Augmented Matrix

not a zero matrix. Note that $\mathbf{x}$ has n components, whereas $\mathbf{b}$ has m components. The matrix

$$\tilde{\mathbf{A}} = \left[\begin{array}{cccc} a_{11} & \cdots & a_{1n} & b_1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

is called the **augmented matrix** of the system (1). The dashed vertical line could be omitted, as we shall do later. It is merely a reminder that the last column of $\tilde{\mathbf{A}}$ did not come from matrix $\mathbf{A}$ but came from vector $\mathbf{b}$. Thus, we *augmented* the matrix $\mathbf{A}$.

Kreyszig (2011)

Gaussian Elimination

a general system to triangular form. For instance, let the given system be

$$\begin{aligned} 2x_1 + 5x_2 &= 2 \\ -4x_1 + 3x_2 &= -30. \end{aligned} \qquad \text{Its augmented matrix is} \qquad \begin{bmatrix} 2 & 5 & 2 \\ -4 & 3 & -30 \end{bmatrix}.$$

We leave the first equation as it is. We eliminate x_1 from the second equation, to get a triangular system. For this we add twice the first equation to the second, and we do the same

operation on the ***rows*** of the augmented matrix. This gives $-4x_1 + 4x_1 + 3x_2 + 10x_2 = -30 + 2 \cdot 2$, that is,

$$\begin{aligned} 2x_1 + 5x_2 &= 2 \\ 13x_2 &= -26 \end{aligned} \qquad \text{Row 2} + 2\ \text{Row 1} \qquad \begin{bmatrix} 2 & 5 & 2 \\ 0 & 13 & -26 \end{bmatrix}$$

where Row 2 + 2 Row 1 means "Add twice Row 1 to Row 2" in the original matrix. This is the **Gauss elimination** (for 2 equations in 2 unknowns) giving the triangular form, from which back substitution now yields $x_2 = -2$ and $x_1 = 6$, as before.

Since a linear system is completely determined by its augmented matrix, ***Gauss elimination can be done by merely considering the matrices***, as we have just indicated.

Row Echelon Form and Information From It

At the end of the Gauss elimination the form of the coefficient matrix, the augmented matrix, and the system itself are called the **row echelon form**. In it, rows of zeros, if present, are the last rows, and, in each nonzero row, the leftmost nonzero entry is farther to the right than in the previous row. For instance, in Example 4 the coefficient matrix and its augmented in row echelon form are

$$(8) \qquad \begin{bmatrix} 3 & 2 & 1 \\ 0 & -\frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 0 & -\frac{1}{3} & \frac{1}{3} & -2 \\ 0 & 0 & 0 & 12 \end{array}\right].$$

Kreyszig (2011)

Row Echelon Form: A_{mxn}

At the end of the Gauss elimination (before the back substitution), the row echelon form of the augmented matrix will be

rank: r

no solution if one of f_{r+1}~f_m is not zero.

$$
\left[\begin{array}{ccccc|c}
r_{11} & r_{12} & \cdots\cdots & & r_{1n} & f_1 \\
 & r_{22} & \cdots\cdots & & r_{2n} & f_2 \\
 & & \ddots & & \vdots & \vdots \\
 & & & r_{rr} \quad \cdots & r_{rn} & f_r \\
 & & & & & f_{r+1} \\
 & & & & & \vdots \\
 & & & & & f_m
\end{array}\right]. \tag{9}
$$

Here, $r \leqq m$, $r_{11} \neq 0$, and all entries in the blue triangle and blue rectangle are zero.

The number of nonzero rows, r, in the row-reduced coefficient matrix **R** is called the **rank of R** and also the **rank of A**. Here is the method for determining whether $\mathbf{Ax} = \mathbf{b}$ has solutions and what they are:

Kreyszig (2011)

Three Scenarios for Solutions

(a) No solution. If r is less than m (meaning that $\mathbf{R}$ actually has at least one row of all 0s) *and* at least one of the numbers $f_{r+1}, f_{r+2}, \cdots, f_m$ is not zero, then the system

$\mathbf{Rx} = \mathbf{f}$ is inconsistent: No solution is possible. Therefore the system $\mathbf{Ax} = \mathbf{b}$ is inconsistent as well. See Example 4, where $r = 2 < m = 3$ and $f_{r+1} = f_3 = 12$.

If the system is consistent (either $r = m$, or $r < m$ *and* all the numbers $f_{r+1}, f_{r+2}, \cdots, f_m$ are zero), then there are solutions.

(b) Unique solution. If the system is consistent and $r = n$, there is exactly one solution, which can be found by back substitution. See Example 2, where $r = n = 3$ and $m = 4$.

$\rightarrow$ nxn matrix

(c) Infinitely many solutions. To obtain any of these solutions, choose values of $x_{r+1}, \cdots, x_n$ arbitrarily. Then solve the rth equation for x_r (in terms of those arbitrary values), then the $(r-1)$st equation for x_{r-1}, and so on up the line. See Example 3.

Kreyszig (2011)

LU Decomposition (wiki)

Let A be a square matrix. An **LU factorization** refers to the factorization of A, with proper row and/or column orderings or permutations, into two factors, $\boxed{\text{a lower triangular matrix } L \text{ and an upper triangular matrix } U}$,

$$A = LU,$$

In the lower triangular matrix all elements above the diagonal are zero, in the upper triangular matrix, all the elements below the diagonal are zero. For example, for a 3-by-3 matrix A, its LU decomposition looks like this:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}.$$

Outline

- Fundamental Concepts
  - Eight Axioms; Subset vs. Subspace
- Basic Matrix Operations
  - Addition, Subtraction, Scalar Multiplication, Matrix Multiplication
- Matrix Properties
  - Rank, Range, Kernel, and Nullspace
- Main Types of Matrices
  - Matrix Transpose, Identity, and Inverse;
  - Symmetric, Skew-symmetric, and Orthogonal Matrices
- Linear Dependence/Independence and Linear Systems
- Elimination and LU Decomposition
  - Elementary Row Operations, Augmented Matrix, and Row Echelon Form
- Fundamental Theorem for Linear Systems
- Eigenvalue Problem
  - Similarity Transformation and Diagonalization
- Quadratic Forms and Rayleigh Quotient

Fundamental Theorem for Linear Systems

Fundamental Theorem for Linear Systems

$rank\,(A) = rank(\tilde{A})$

$\tilde{A}$: *augmented matrix*

$\boxed{\textbf{(a) Existence.}}$ *A linear system of m equations in n unknowns* $x_1, \cdots, x_n$

$$\text{(1)} \qquad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots&\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

is **consistent**, *that is, has solutions, if and only if the coefficient matrix* **A** *and the augmented matrix* $\tilde{\mathbf{A}}$ *have the same rank. Here,*

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{A}} = \left[\begin{array}{ccc:c} a_{11} & \cdots & a_{1n} & b_1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

$\boxed{\textbf{(b) Uniqueness.}}$ *The system (1) has precisely one solution if and only if this common rank r of* **A** *and* $\tilde{\mathbf{A}}$ *equals n.*

$rank\,(A) = rank(\tilde{A}) = \boldsymbol{n}$

Homogeneous Linear System

Homogeneous Linear System

A homogeneous linear system

$$
(4) \qquad \begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
\cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; &\cdot \; \cdot \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}
$$

A x=0

always has the **trivial solution** $x_1 = 0, \cdots, x_n = 0$. *Nontrivial solutions exist if and only if rank* $\mathbf{A} < n$. *If* rank $\mathbf{A} = r < n$, *these solutions, together with* $\mathbf{x} = \mathbf{0}$, *form a vector space* (see Sec. 7.4) *of dimension* $n - r$ called the **solution space** of (4).

In particular, if $\mathbf{x}_{(1)}$ *and* $\mathbf{x}_{(2)}$ *are solution vectors of* (4), *then* $\mathbf{x} = c_1\mathbf{x}_{(1)} + c_2\mathbf{x}_{(2)}$ *with any scalars* c_1 *and* c_2 *is a solution vector of* (4). (This ***does not hold*** for nonhomogeneous systems. Also, the term *solution space is used for homogeneous* systems only.)

Kreyszig (2011)

Row-Equivalent Systems

We now call a linear system S_1 **row-equivalent** to a linear system S_2 if S_1 can be obtained from S_2 by (finitely many!) row operations. This justifies Gauss elimination and establishes the following result.

Row-Equivalent Systems

Row-equivalent linear systems have the same set of solutions.

Because of this theorem, systems having the same solution sets are often called *equivalent systems*. But note well that we are dealing with ***row operations***. No column operations on the augmented matrix are permitted in this context because they would generally alter the solution set.

Kreyszig (2011)

Equivalent Statements for a Non-singular Matrix

Square Matrix

Theorem 1.3. *For $A \in \mathbb{C}^{m \times m}$, the following conditions are equivalent:*

(a) *A has an inverse A^{-1},*

$\boxed{\text{(b) } \operatorname{rank}(A) = m,}$

(c) $\operatorname{range}(A) = \mathbb{C}^m$,

(d) $\operatorname{null}(A) = \{0\}$,

(e) *0 is not an eigenvalue of A,*

(f) *0 is not a singular value of A,*

$\boxed{\text{(g) } \det(A) \neq 0.}$

dim range(A) = m

Ax=0 ➔ x=0; “null space”=”kernel”

Equivalent Statements for a Non-singular Matrix

Square Matrix

Theorem 2.5. *Let A be an $n \times n$ matrix. Then the following statements are equivalent:*

(1) *The columns of A are linearly independent.*

(2) *The columns of A span $\mathbb{R}^n$.*

(3) *The columns of A form a basis for $\mathbb{R}^n$.*

(4) *The equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbb{R}^n$.*

(5) *A is an invertible matrix.*

(6) *The determinant of A is nonzero.*

(7) *A is row equivalent to the identity matrix.*

Equivalent Statements for a Singular Matrix

Square Matrix

Matrices and Systems of Equations

Theorem 1. Let $\mathbf{A}$ be an $n \times n$ matrix. The following statements are equivalent:

(a) $\mathbf{A}$ is singular (does not have an inverse).

(b) The determinant of $\mathbf{A}$ is zero.

(c) $\mathbf{Ax} = \mathbf{0}$ has nontrivial solutions ($\mathbf{x} \neq \mathbf{0}$).

(d) The columns (rows) of $\mathbf{A}$ form a linearly dependent set.

Overdetermined vs. Underdetermined

A_{mn} x =b. (m, #of equations; n, # of variables)

- overdetermined, m>n
- determined, m=n
- underdetermined, m<n

- consistent, at least one solution
- inconsistent, no solution,

A linear system (1) is called **overdetermined** if it has more equations than unknowns, as in Example 2, **determined** if $m = n$, as in Example 1, and **underdetermined** if it has fewer equations than unknowns.

Furthermore, a system (1) is called **consistent** if it has at least one solution (thus, one solution or infinitely many solutions), but **inconsistent** if it has no solutions at all, as $x_1 + x_2 = 1, x_1 + x_2 = 0$ in Example 1, Case (c).

Kreyszig (2011)

Outline

- Fundamental Concepts
  - Eight Axioms; Subset vs. Subspace
- Basic Matrix Operations
  - Addition, Subtraction, Scalar Multiplication, Matrix Multiplication
- Matrix Properties
  - Rank, Range, Kernel, and Nullspace
- Main Types of Matrices
  - Matrix Transpose, Identity, and Inverse;
  - Symmetric, Skew-symmetric, and Orthogonal Matrices
- Linear Dependence/Independence and Linear Systems
- Elimination and LU Decomposition
  - Elementary Row Operations, Augmented Matrix, and Row Echelon Form
- Fundamental Theorem for Linear Systems
- Eigenvalue Problem
  - Similarity Transformation and Diagonalization
- Quadratic Forms and Rayleigh Quotient

Eigenvalues and Eigenvectors

$$\boldsymbol{Au} = r\mathbf{u} \qquad (\boldsymbol{A} - r\boldsymbol{I})\boldsymbol{u} = 0$$

r: *eigenvalue*

$\boldsymbol{u}$: *eigenvector*

If the above system has at least one non-trivial solution, $|\boldsymbol{A} - r\boldsymbol{I}| = 0$.

Eigenvalues and Eigenvectors

Definition 3. Let $\mathbf{A} = [a_{ij}]$ be an $n \times n$ constant matrix. The **eigenvalues** of $\mathbf{A}$ are those (real or complex) numbers r for which $(\mathbf{A} - r\mathbf{I})\mathbf{u} = \mathbf{0}$ has at least one nontrivial solution† $\mathbf{u}$. The corresponding nontrivial solutions $\mathbf{u}$ are called the **eigenvectors** of $\mathbf{A}$ associated with r.

Nagle, et al. (2011)

Eigenvalues and Eigenvectors: An Example

Example 2 Find the eigenvalues and eigenvectors of the matrix

$$\mathbf{A} := \begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{bmatrix}.$$

Solution The characteristic equation for **A** is

$$|\mathbf{A} - r\mathbf{I}| = \begin{vmatrix} 1-r & 2 & -1 \\ 1 & -r & 1 \\ 4 & -4 & 5-r \end{vmatrix} = 0\,,$$

which simplifies to $(r-1)(r-2)(r-3) = 0$. Hence, the eigenvalues of **A** are $r_1 = 1$, $r_2 = 2$, and $r_3 = 3$.

Eigenvectors: An Example (cont.)

$r_1 = 1$, $r_2 = 2$, and $r_3 = 3$. To find the eigenvectors corresponding to $r_1 = 1$, we set $r = 1$ in $(\mathbf{A} - r\mathbf{I})\mathbf{u} = \mathbf{0}$. This gives

$$\text{(8)} \qquad \begin{bmatrix} 0 & 2 & -1 \\ 1 & -1 & 1 \\ 4 & -4 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

Using elementary row operations (Gaussian elimination), we see that (8) is equivalent to the two equations

$$\begin{aligned} u_1 - u_2 + u_3 &= 0 \ , \\ 2u_2 - u_3 &= 0 \ . \end{aligned}$$

Thus, we can obtain the solutions to (8) by assigning an arbitrary value to u_2 (say, $u_2 = s$), solving $2u_2 - u_3 = 0$ for u_3 to get $u_3 = 2s$, and then solving $u_1 - u_2 + u_3 = 0$ for u_1 to get $u_1 = -s$. Hence, the eigenvectors associated with $r_1 = 1$ are

$$\text{(9)} \qquad \mathbf{u}_1 = s \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} .$$

Eigenvectors: An Example (cont.)

For $r_2 = 2$, we solve

$$\begin{bmatrix} -1 & 2 & -1 \\ 1 & -2 & 1 \\ 4 & -4 & 3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

in a similar fashion to obtain the eigenvectors

$$\mathbf{u}_2 = s \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix} . \tag{10}$$

Finally, for $r_3 = 3$, we solve

$$\begin{bmatrix} -2 & 2 & -1 \\ 1 & -3 & 1 \\ 4 & -4 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and get the eigenvectors

$$\mathbf{u}_3 = s \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix} . \tag{11}$$

◆

Characteristic Polynomial: $|A - zI|$

The *characteristic polynomial* of $A \in \mathbb{C}^{m\times m}$, denoted by p_A or simply p, is the degree m polynomial defined by

$$p_A(z) = \det(zI - A). \tag{24.5}$$

Thanks to the placement of the minus sign, p is *monic:* the coefficient of its degree m term is 1.

Theorem 24.1. *λ is an eigenvalue of A if and only if $p_A(\lambda) = 0$.*

Proof. This follows from the definition of an eigenvalue:

$$\begin{aligned}\lambda \text{ is an eigenvalue} &\iff \text{there is a nonzero vector } x \text{ such that } \lambda x - Ax = 0\\ &\iff \lambda I - A \text{ is singular}\\ &\iff \det(\lambda I - A) = 0. \qquad \square\end{aligned}$$

Theorem 24.1 has an important consequence. *Even if a matrix is real, some of its eigenvalues may be complex.* Physically, this is related to the phenomenon that real dynamical systems can have motions that oscillate as well as grow or decay. Algorithmically, it means that even if the input to a matrix eigenvalue problem is real, the output may have to be complex.

Eigenvectors and Eigenspace

Eigenvalues

The eigenvalues of a square matrix $\mathbf{A}$ *are the roots of the characteristic equation* (4) *of* $\mathbf{A}$.

Hence an $n \times n$ *matrix has at least one eigenvalue and at most* n *numerically different eigenvalues.*

Eigenvectors, Eigenspace

If $\mathbf{w}$ *and* $\mathbf{x}$ *are eigenvectors of a matrix* $\mathbf{A}$ *corresponding to* ***the same*** *eigenvalue* λ, *so are* $\mathbf{w} + \mathbf{x}$ (provided $\mathbf{x} \neq -\mathbf{w}$) *and* $k\mathbf{x}$ *for any* $k \neq 0$.

Hence the eigenvectors corresponding to one and the same eigenvalue λ *of* $\mathbf{A}$, *together with* $\mathbf{0}$, *form a vector space* (cf. Sec. 7.4), *called the* **eigenspace** *of* $\mathbf{A}$ *corresponding to that* λ.

Review: Symmetric and Orthogonal Matrices

Symmetric, Skew-Symmetric, and Orthogonal Matrices

A ***real*** square matrix $\mathbf{A} = [a_{jk}]$ is called

symmetric if transposition leaves it unchanged,

(1) $$\mathbf{A}^{\mathsf{T}} = \mathbf{A}, \qquad \text{thus} \qquad a_{kj} = a_{jk},$$

skew-symmetric if transposition gives the negative of $\mathbf{A}$,

(2) $$\mathbf{A}^{\mathsf{T}} = -\mathbf{A}, \qquad \text{thus} \qquad a_{kj} = -a_{jk},$$

orthogonal if transposition gives the inverse of $\mathbf{A}$,

(3) $$\mathbf{A}^{\mathsf{T}} = \mathbf{A}^{-1}.$$

Eigenvalues

Eigenvalues of the Transpose

The transpose $\mathbf{A}^\mathsf{T}$ of a square matrix $\mathbf{A}$ has the same eigenvalues as $\mathbf{A}$.

Eigenvalues of Symmetric and Skew-Symmetric Matrices

(a) *The eigenvalues of a symmetric matrix are real.*

(b) *The eigenvalues of a skew-symmetric matrix are pure imaginary or zero.*

Inner Product

Invariance of Inner Product

An orthogonal transformation preserves the value of the **inner product** *of vectors* **a** *and* **b** *in* R^n, *defined by*

$$\mathbf{a} \bullet \mathbf{b} = \mathbf{a}^{\mathsf{T}}\mathbf{b} = [a_1 \quad \cdots \quad a_n] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}. \tag{7}$$

That is, for any **a** *and* **b** *in* R^n, *orthogonal* $n \times n$ *matrix* $\mathbf{A}$, *and* $\mathbf{u} = \mathbf{A}\mathbf{a}, \mathbf{v} = \mathbf{A}\mathbf{b}$ *we have* $\mathbf{u} \bullet \mathbf{v} = \mathbf{a} \bullet \mathbf{b}$.

Hence the transformation also preserves the **length** *or* **norm** *of any vector* **a** *in* R^n *given by*

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \bullet \mathbf{a}} = \sqrt{\mathbf{a}^{\mathsf{T}}\mathbf{a}}. \tag{8}$$

Orthogonality of Eigenvectors

Orthonormality of Column and Row Vectors

A real square matrix is orthogonal if and only if its column vectors $\mathbf{a}_1, \cdots, \mathbf{a}_n$ *(and also its row vectors) form an* **orthonormal system**, *that is,*

$$\mathbf{a}_j \bullet \mathbf{a}_k = \mathbf{a}_j^{\mathsf{T}} \mathbf{a}_k = \begin{cases} 0 & \text{if} \quad j \neq k \\ 1 & \text{if} \quad j = k. \end{cases} \tag{10}$$

Determinant of an Orthogonal Matrix

The determinant of an orthogonal matrix has the value $+1$ *or* -1.

Symmetric Matrices

A symmetric matrix has an orthonormal basis of eigenvectors for R^n.

Similarity and Similarity Transformation

Similar Matrices. Similarity Transformation

An $n \times n$ matrix $\hat{\mathbf{A}}$ is called **similar** to an $n \times n$ matrix $\mathbf{A}$ if

$$\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \tag{4}$$

for some (nonsingular!) $n \times n$ matrix $\mathbf{P}$. This transformation, which gives $\hat{\mathbf{A}}$ from $\mathbf{A}$, is called a **similarity transformation**.

Eigenvalues and Eigenvectors of Similar Matrices

If $\hat{\mathbf{A}}$ is similar to $\mathbf{A}$, then $\hat{\mathbf{A}}$ has the same eigenvalues as $\mathbf{A}$.

Furthermore, if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, then $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ is an eigenvector of $\hat{\mathbf{A}}$ corresponding to the same eigenvalue.

A and $\hat{A}$ are similar when they have the same eigenvalues.

Similarity and Linearly Conjugate

Similar Matrices. Similarity Transformation

An $n \times n$ matrix $\hat{\mathbf{A}}$ is called **similar** to an $n \times n$ matrix $\mathbf{A}$ if

$$\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \tag{4}$$

for some (nonsingular!) $n \times n$ matrix $\mathbf{P}$. This transformation, which gives $\hat{\mathbf{A}}$ from $\mathbf{A}$, is called a **similarity transformation**.

A and $\hat{A}$ are similar when they have the same eigenvalues.

Proposition: If L_1 and L_2 are linearly conjugate,

$$L_1(x) = A_1 x \text{ and } L_2(x) = A_2 x$$

Then, A_1 and A_2 have the same eigenvalues.

Diagonalization of a Matrix

- X is constructed using the eigenvectors of the matrix A.
- D is diagonal and its elements consist of the eigenvalues of the matrix A.

Diagonalization of a Matrix

If an $n \times n$ matrix $\mathbf{A}$ has a basis of eigenvectors, then

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X} \tag{5}$$

is diagonal, with the eigenvalues of $\mathbf{A}$ as the entries on the main diagonal. Here $\mathbf{X}$ is the matrix with these eigenvectors as column vectors. Also,

$$\mathbf{D}^m = \mathbf{X}^{-1}\mathbf{A}^m\mathbf{X} \qquad (m = 2, 3, \cdots). \tag{5*}$$

Outline

- Fundamental Concepts
  - Eight Axioms; Subset vs. Subspace
- Basic Matrix Operations
  - Addition, Subtraction, Scalar Multiplication, Matrix Multiplication
- Matrix Properties
  - Rank, Range, Kernel, and Nullspace
- Main Types of Matrices
  - Matrix Transpose, Identity, and Inverse;
  - Symmetric, Skew-symmetric, and Orthogonal Matrices
- Linear Dependence/Independence and Linear Systems
- Elimination and LU Decomposition
  - Elementary Row Operations, Augmented Matrix, and Row Echelon Form
- Fundamental Theorem for Linear Systems
- Eigenvalue Problem
  - Similarity Transformation and Diagonalization
- Quadratic Forms and Rayleigh Quotient

Quadratic Forms. Transformation to Principal Axes

By definition, a **quadratic form** Q in the components $x_1, \cdots, x_n$ of a vector $\mathbf{x}$ is a sum of n^2 terms, namely,

$$
\begin{aligned}
Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} &= \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k \\
&= \quad a_{11}x_1^2 \quad + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\
&\quad + a_{21}x_2x_1 + a_{22}x_2^2 \quad + \cdots + a_{2n}x_2x_n \\
&\quad + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\quad + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2.
\end{aligned} \tag{7}
$$

$\mathbf{A} = [a_{jk}]$ is called the **coefficient matrix** of the form. We may assume that $\mathbf{A}$ is ***symmetric***, because we can take off-diagonal terms together in pairs and write the result as a sum of two equal terms; see the following example.

Quadratic Form with a Constraint

$$minimize \quad Q(x) = x^T A x$$

$$P: \; x^T x = constant$$

$$\nabla Q(x) = \lambda \nabla P(x)$$

$$\boldsymbol{Ax} = \lambda \boldsymbol{x}$$

Rayleigh Quotient Optimization

$$Rayleigh\ Quotient \quad r(x) = \frac{x^T A x}{x^T x}$$

$$optimize \quad r(x) = \frac{x^T A x}{x^T x}$$

$$\nabla r(x) = 0$$

$$(\boldsymbol{A} - r\boldsymbol{I})\boldsymbol{x} = 0$$

- If $\boldsymbol{x}$ is an eigenvector of the matrix A, r represents the corresponding eigenvalue

Rayleigh Quotient

The *Rayleigh quotient* of a vector $x \in \mathbb{R}^m$ is the scalar

$$r(x) = \frac{\boxed{x^T A x}}{x^T x}. \qquad (27.1)$$

Quadratic Form

Notice that if x is an eigenvector, then $r(x) = \lambda$ is the corresponding eigenvalue. One way to motivate this formula is to ask: given x, what scalar α "acts most like an eigenvalue" for x in the sense of minimizing $\|Ax - \alpha x\|_2$? This is an $m \times 1$ least squares problem of the form $x\alpha \approx Ax$ (x is the matrix, α is the unknown vector, Ax is the right-hand side). By writing the normal equations (11.9) for this system, we obtain the answer: $\alpha = r(x)$. Thus $r(x)$ is a natural eigenvalue estimate to consider if x is close to, but not necessarily equal to, an eigenvector.

If A is an identity matrix, the quadratic form represents the inner product.

Rayleigh Quotient

To make these ideas quantitative, it is fruitful to view $x \in \mathbb{R}^m$ as a variable, so that r is a function $\mathbb{R}^m \to \mathbb{R}$. We are interested in the local behavior of $r(x)$ when x is near an eigenvector. One way to approach this question is to calculate the partial derivatives of $r(x)$ with respect to the coordinates x_j:

$$\frac{\partial r(x)}{\partial x_j} = \frac{\frac{\partial}{\partial x_j}(x^T A x)}{x^T x} - \frac{(x^T A x)\frac{\partial}{\partial x_j}(x^T x)}{(x^T x)^2}$$

Here, we only consider real and symmetric matrices.

$$= \frac{2(Ax)_j}{x^T x} - \frac{(x^T A x) 2x_j}{(x^T x)^2} = \frac{2}{x^T x}(Ax - r(x)x)_j .$$

If we collect these partial derivatives into an m-vector, we find we have calculated the *gradient* of $r(x)$, denoted by $\nabla r(x)$. We have shown:

$$\nabla r(x) = \frac{2}{x^T x}(Ax - r(x)x). \tag{27.2}$$

From this formula we see that at an eigenvector x of A, the gradient of $r(x)$ is the zero vector. Conversely, if $\nabla r(x) = 0$ with $x \neq 0$, then x is an eigenvector

Lagrange Multipliers and Eigenvalue Problem

$$x' = ax + by \quad (= P(x, y)) \qquad (1)$$

$$y' = cx + dy \quad (= Q(x, y)\,) \qquad (2)$$

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$X' = AX$$

$$F =' RQ' = X^t AX = xP + \mathrm{y}\mathrm{Q} = \frac{1}{2}\frac{d(x^2 + y^2)}{dt}$$

$$G = x^2 + y^2 = r^2$$

$$\nabla F = (P + ax + cy, \mathrm{Q} + bx + dy) = (\mathrm{P}, \mathrm{Q}) + \ (ax + cy, bx + dy)$$

$$= AX + A^t X \quad = 2AX \qquad (if\ A = A^t, \mathrm{i.e, symmetric}) \qquad b = c$$

$$\nabla G = (2x, 2\mathrm{y}\,) = 2X$$

$\boxed{\nabla F = \lambda\, \nabla G \text{ (a lagrange multiplier)}}$ $\Longrightarrow$ $\boxed{AX = \lambda X \text{ (Eigenvalue Problem)}}$

Lagrange Multipliers and Eigenvalue Problem

$$x' = ax + by \quad (= P(x,y)) \qquad (1)$$

$$y' = cx + dy \quad (= Q(x,y)\ ) \qquad (2)$$

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad X = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$X' = AX$$

$$F = 'RQ' = X^t AX = xP + \mathrm{y}\mathrm{Q}$$

$$x(1) + \mathrm{y}(2) \Rightarrow \quad \boxed{xx' + yy' = \frac{1}{2}\frac{d(x^2+y^2)}{dt} = F}$$

$$G = x^2 + y^2 = r^2$$

$$RQ = \frac{X^t AX}{x^2+y^2} = \frac{F}{G} = \frac{1}{2r^2}\frac{d(r^2)}{dt}$$

- An eigenvalue problem is to optimize the RQ
- ~ Growth rate of r^2

Rayleigh Quotients for A and A^t

$$RQ = X^t AX/x^2 + y^2$$

$$x' = ax + by \qquad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad X' = AX$$
$$y' = cx + dy$$

$$X^t AX = x(ax + by\,) + \mathrm{y}(cx + dy\,)$$

$$x' = ax + cy \qquad A^t = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \qquad X' = A^t X$$
$$y' = bx + dy$$

$$X^t A^t X = x(ax + cy\,) + \mathrm{y}(bx + dy\,) = x(ax + by\,) + \mathrm{y}(cx + dy\,) = X^t AX$$