Advanced Engineering Mathematics Vectors, Matrices, and Vector Calculus by Dennis G. Zill Notes

Chris Doble

June 2023

Contents

1 Vectors

1.1 Vectors in 2-Space

- The zero vector can be assigned any direction
- The vectors **i** and **j** are known as the **standard basis vectors** for $\mathbb{R}^2$

1.2 Vectors in 3-Space

- In $\mathbb{R}^3$ the octant in which all coordinates are positive is known as the **first octant**. There is no agreement for naming the other seven octants.

1.3 Dot Product

- The **dot product** is also known as the **inner product** or the **scalar product** and is denoted $\mathbf{a} \cdot \mathbf{b}$
- Two non-zero vectors are orthogonal iff their dot product is 0
- The zero vector is considered orthogonal to all vectors
- The angles α, β, and γ between a vector and the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, respectively are called the **direction angles** of the vector
- The cosines of a vectors direction angles (the **direction cosines**) can be calculated as

$$\begin{aligned}
\cos\alpha &= \frac{\mathbf{a} \cdot \mathbf{i}}{||\mathbf{a}||||\mathbf{i}||} \\
&= \frac{a_1}{||\mathbf{a}||} \\
\cos\beta &= \frac{\mathbf{a} \cdot \mathbf{j}}{||\mathbf{a}||||\mathbf{j}||} \\
&= \frac{a_2}{||\mathbf{a}||} \\
\cos\gamma &= \frac{\mathbf{a} \cdot \mathbf{k}}{||\mathbf{a}||||\mathbf{k}||} \\
&= \frac{a_3}{||\mathbf{a}||}
\end{aligned}$$

  Equivalently, these can be calculated as the components of the unit vector $\mathbf{a}/||\mathbf{a}||$.

- To find the component of a vector $\mathbf{a}$ in the direction of a vector $\mathbf{b}$

$$\mathrm{comp}_{\mathbf{b}}\mathbf{a} = ||\mathbf{a}|| \cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{||\mathbf{b}||}$$

- To project a vector $\mathbf{a}$ onto a vector $\mathbf{b}$

$$\mathrm{proj}_{\mathbf{b}}\mathbf{a} = (\mathrm{comp}_{\mathbf{b}}\mathbf{a})\frac{\mathbf{b}}{||\mathbf{b}||} = \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}}\right)\mathbf{b}$$

1.4 Cross Product

- The cross product is only defined in $\mathbb{R}^3$
- The **scalar triple product** of vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ is defined as

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

- The area of a parallelogram with sides $\mathbf{a}$ and $\mathbf{b}$ is $||\mathbf{a} \times \mathbf{b}||$
- The area of a triangle with sides $\mathbf{a}$ and $\mathbf{b}$ is $\frac{1}{2}||\mathbf{a} \times \mathbf{b}||$
- The volume of a paralleleipied with sides $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ is $|\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})|$
- $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = 0$ iff $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are coplanar

1.5 Lines and Planes in 3-Space

- There is a unique line between any two points $\mathbf{r_1}$ and $\mathbf{r_2}$ in 3-space. The equation for that line is

$$\mathbf{r} = \mathbf{r_1} + t(\mathbf{r_2} - \mathbf{r_1}) = \mathbf{r_1} + t\mathbf{a}$$

  where t is called a **parameter**, the nonzero vector $\mathbf{a}$ is called a **direction vector**, and its components are called **direction numbers**.

- Equating the components of the equation above we find

$$\begin{aligned} x &= r_1 + ta_1 \\ y &= r_2 + ta_2 \\ z &= r_3 + ta_3. \end{aligned}$$

  These are the **parametric equations** for the line through $\mathbf{r_1}$ and $\mathbf{r_2}$.

- By solving the parametric equations for t and equating the results we find the **symmetric equations** for the line

$$t = \frac{x - r_1}{a_1} = \frac{y - r_2}{a_2} = \frac{z - r_3}{a_3}.$$

- Given a point P_1 and a vector $\mathbf{n}$, there exists only one plane containing P_1 with $\mathbf{n}$ normal. The vector from P_1 to another point P on that plane will be perpendicular to $\mathbf{n}$, so the equation for the plane is

$$\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_1) = 0$$

  where $\mathbf{r} = \overrightarrow{OP}$ and $\mathbf{r_1} = \overrightarrow{OP_1}$. If

$$\mathbf{n} = a\hat{\mathbf{i}} + b\hat{\mathbf{j}} + c\hat{\mathbf{k}}$$

the cartesian form of this equation is

$$a(x - x_1) + b(y - y_1) + c(z - z_1) = 0$$

and is called the **point-normal form**.

- The graph of any equation $ax + by + cz + d = 0$, where a, b, and c are not all zero, is a plane with the normal vector $\mathbf{n} = a\hat{\mathbf{i}} + b\hat{\mathbf{j}} + c\hat{\mathbf{k}}$.
- Given three noncollinear points, a normal vector can be found by forming two vectors from two pairs of points and take their cross product.
- A line and a plane that aren't parellel intersect at a single point.
- Two planes that aren't parallel must intersect in a line.

1.6 Vector Spaces

- The length of a vector is called its **norm**
- The process of multipying a vector by the reciprocal of its norm is called **normalizing** the vector
- Two nonzero vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^n$ are said to be orthogonal if $\mathbf{a} \cdot \mathbf{b} = 0$

Definition 7.6.1 Vector Space

Let V be a set of elements on which two operations called **vector addition** and **scalar multiplication** are defined. Then V is said to be a **vector space** if the following 10 properties are satisfied.

Axioms for Vector Addition:

(*i*) If $\mathbf{x}$ and $\mathbf{y}$ are in V, then $\mathbf{x} + \mathbf{y}$ is in V.

(*ii*) For all $\mathbf{x}, \mathbf{y}$ in V, $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$. ← commutative law

(*iii*) For all $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in V, $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$. ← associative law

(*iv*) There is a unique vector $\mathbf{0}$ in V such that
$\mathbf{0} + \mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{x}$. ← zero vector

(*v*) For each $\mathbf{x}$ in V, there exists a vector $-\mathbf{x}$ such that
$\mathbf{x} + (-\mathbf{x}) = (-\mathbf{x}) + \mathbf{x} = \mathbf{0}$. ← negative of a vector

Axioms for Scalar Multiplication:

(*vi*) If k is any scalar and $\mathbf{x}$ is in V, then $k\mathbf{x}$ is in V.

(*vii*) $k(\mathbf{x} + \mathbf{y}) = k\mathbf{x} + k\mathbf{y}$ ← distributive law

(*viii*) $(k_1 + k_2)\mathbf{x} = k_1\mathbf{x} + k_2\mathbf{x}$ ← distributive law

(*ix*) $k_1(k_2\mathbf{x}) = (k_1k_2)\mathbf{x}$

(*x*) $1\mathbf{x} = \mathbf{x}$

- If a subset W of a vector space V is itself a vector space under the operations of vector addition and scalar multiplication defined on V, then W is called a **subspace** of V
- Every vector space has at least two subspaces: itself and the zero subspace $\{\mathbf{0}\}$

- A set of vectors $\{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ is said to be **linearly independent** if the only constants satisfying the equation

$$k_1\mathbf{x_1} + k_2\mathbf{x_2} + \cdots + k_n\mathbf{x_n} = \mathbf{0}$$

  are $k_1 = k_2 = \cdots = k_n = 0$. If the set of vectors is not linearly independent it is said to be **linearly dependent**.

- If a set of vectors $B = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ in a vector space V is linearly independent and every vector in V can be expressed as a linear combination of vectors in B then B is said to be a **basis** for V.

- The number of vectors in a basis B for a vector space V is said to be the **dimension** of the space.

- If the basis of a vector space contains a finite number of vectors, then the space is **finite dimensional**; otherwise it is **infinite dimensional**.

- If S denotes any set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ in a vector space V, then the set of all linear combinations of the vectors in S

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n$$

  is called the **span** of the vectors and is denoted $\mathrm{Span}(S)$.

- $\mathrm{Span}(S)$ is a subspace of V and is said to be a subspace spanned by its vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.

- If $V = \mathrm{Span}(S)$ then S is said to be a **spanning set** for the vector space V or that S **spans** V.

1.7 Gram–Schmidt Orthogonalization Process

- An **orthonormal basis** is a basis whose vectors are mutually orthogonal and are unit vectors.

- If $B = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ is an orthonormal basis for $\mathbb{R}^n$ then an arbitrary vector $\mathbf{u}$ can be expressed as

$$\mathbf{u} = (\mathbf{u} \cdot \mathbf{w}_1)\mathbf{w}_1 + (\mathbf{u} \cdot \mathbf{w}_2)\mathbf{w}_2 + \cdots + (\mathbf{u} \cdot \mathbf{w}_n)\mathbf{w}_n$$

- The **Gram-Schmidt Orthogonalization Process** is a process for converting any basis of a vector space into an orthonormal basis. First the basis vectors are made orthogonal to each other, then they are normalized. More specifically, to convert a basis $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ into an orthogonal basis $B' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$

  1. Let $\mathbf{v}_1 = \mathbf{u}_1$
  2. Let $\mathbf{v}_2 = \mathbf{u}_2 - \mathrm{proj}_{\mathbf{v}_1}\mathbf{u}_2$

3. ...
4. Let $\mathbf{v}_n = \mathbf{u}_n - \text{proj}_{\mathbf{v}_1}\mathbf{u}_n - \text{proj}_{\mathbf{v}_2}\mathbf{u}_n - \cdots - \text{proj}_{\mathbf{v}_{n-1}}\mathbf{u}_n$

and to convert B' into an orthonormal basis $B'' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$, normalize each $\mathbf{v}_i$, $i = 1, 2, \ldots, n$.

2 Matrices

2.1 Matrix Algebra

- Vectors can be written as horizontal or vertical arrays of numbers
- A **matrix** is any rectangular array of numbers or functions

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

- The numbers or functions in the array are called the **elements** or **entries** of the matrix
- If a matrix has m rows and n columns we say that its **size** is m by n or $m \times n$
- An $n \times n$ matrix is called a **square** matrix of **order** n
- The entry in the ith row and the jth column of an $m \times n$ matrix $\mathbf{A}$ is written a_{ij}
- An $m \times 1$ matrix

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

  is called a **column vector**
- A $1 \times n$ matrix

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

  is called a **row vector**

Definition 8.1.6 Matrix Multiplication

Let $\mathbf{A}$ be a matrix having m rows and p columns, and let $\mathbf{B}$ be a matrix having p rows and n columns. The **product AB** is the $m \times n$ matrix

$$\mathbf{AB} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1p}b_{p1} & \cdots & a_{11}b_{1n} + a_{12}b_{2n} + \cdots + a_{1p}b_{pn} \\ a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2p}b_{p1} & \cdots & a_{21}b_{1n} + a_{22}b_{2n} + \cdots + a_{2p}b_{pn} \\ \vdots & & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \cdots + a_{mp}b_{p1} & \cdots & a_{m1}b_{1n} + a_{m2}b_{2n} + \cdots + a_{mp}b_{pn} \end{pmatrix}$$

$$= \left(\sum_{k=1}^{p} a_{ik}b_{kj} \right)_{m \times n}.$$

- Matrix multiplication is associative, i.e. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
- Matrix multiplication is distributive, i.e. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ and $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$
- The **transpose** of an $m \times n$ matrix $\mathbf{A}$ is an $n \times m$ matrix $\mathbf{A}^T$

$$\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

  i.e. the matrix is flipped along the main diagonal

Theorem 8.1.2 Properties of Transpose

Suppose $\mathbf{A}$ and $\mathbf{B}$ are matrices and k a scalar. Then

(*i*) $(\mathbf{A}^T)^T = \mathbf{A}$ ← transpose of a transpose
(*ii*) $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$ ← transpose of a sum
(*iii*) $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ ← transpose of a product
(*iv*) $(k\mathbf{A})^T = k\mathbf{A}^T$ ← transpose of a scalar multiple

- A matrix that consists of all zero entries is called a **zero matrix**
- A square matrix is said to be a **triangular matrix** if all of its entries above or below the main diagonal are zeroes. More specifically they are called **lower triangular** and **upper triangular** matrices, respectively.
- A square matrix is called a **diagonal matrix** if all entries not on the main diagonal are 0.

- A square matrix whose entries on the main diagonal are all equal is called a **scalar matrix**
- A square matrix that has the property $\mathbf{A} = \mathbf{A}^T$ is called a **symmetric matrix**

2.2 Systems of Linear Algebraic Equations

- In a linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_n \end{aligned}$$

  the values a_{ij} are called the **coefficients** and the values b_n are called the **constants**

- If all the constants are zero the system is said to be **homogeneous**, otherwise it is **nonhomogeneous**
- A linear system is said to be **consistent** if it has at least one solution, otherwise it's **inconsistent**
- A linear system can be transformed into an equivalent system (i.e. one that has the same solutions) via three elementary operations:
  1. Multiply an equation by a nonzero constant
  2. Interchange the positions of equations in the system
  3. Add a multiple of one equation to any other equation
- A linear system can be represented by an **augmented matrix**, e.g.

$$\left(\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array}\right)$$

- We say that two matrices are **row equivalent** if one can be obtained from the other via a series of elementary row operations
- **Gaussian elimination** is the process of applying elementary row operations to a matrix to put it into **row-echelon form** where:
  1. The first nonzero entry in a row is a 1

2. In subsequent rows, the first 1 entry appears to the right of the 1 entry in earlier rows
3. Rows consisting of all zeroes are at the bottom of the matrix

- **Gauss-Jordan elimination** is the same as Gaussian elimination with an additional constraint that puts the matrix into **reduced row-echelon form** where a column containing a first entry 1 has zeroes everywhere else
- A homogeneous linear system always has a trivial solution where all variables are equal to zero and will have an infinite number of nontrivial solutions if the number of equations m is less than the number of variables n, i.e. $m < n$
- If $\mathbf{X}_1$ is a solution to $\mathbf{AX} = \mathbf{0}$, then so is $c\mathbf{X}_1$ for any constant c
- If $\mathbf{X}_1$ and $\mathbf{X}_2$ are solutions of $\mathbf{AX} = \mathbf{0}$, then so is $\mathbf{X}_1 + \mathbf{X}_2$
- If a linear system contains more equations than variables it is said to be **overdetermined**; if it contains fewer equations than variables it is said to be **underdetermined**

2.3 Rank of a Matrix

- The **rank** of a matrix $\mathbf{A}$ denoted rank($\mathbf{A}$) is the number of linearly independent row vectors in $\mathbf{A}$
- The row vectors of an $m \times n$ matrix $\mathbf{A}$ span a subspace of $\mathbb{R}^n$. This is called the **row space** of $\mathbf{A}$. The set of linearly independent row vectors in $\mathbf{A}$ are a basis for that subspace

Theorem 8.3.1 Rank of a Matrix by Row Reduction

If a matrix $\mathbf{A}$ is row equivalent to a row-echelon form $\mathbf{B}$, then

(*i*) the row space of $\mathbf{A}$ = the row space of $\mathbf{B}$,
(*ii*) the nonzero rows of $\mathbf{B}$ form a basis for the row space of $\mathbf{A}$, and
(*iii*) rank($\mathbf{A}$) = the number of nonzero rows in $\mathbf{B}$.

- A linear system of equations $\mathbf{AX} = \mathbf{B}$ is consistent iff the rank of the coefficient matrix $\mathbf{A}$ is equal to the rank of the augmented matrix of the system $(\mathbf{A}|\mathbf{B})$
- Suppose a linear system $\mathbf{AX} = \mathbf{B}$ with m equations and n variables is consistent. If rank($\mathbf{A}$) $= r$ then the solution of the system contains $n - r$ variables

2.4 Determinants

- Suppose $\mathbf{A}$ is an $n \times n$ matrix. Associated with $\mathbf{A}$ is a number called the **determinant of A** and is denoted by

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

- A determinant of an $n \times n$ matrix is called a **determinant of order** n

- The determinant of a 1×1 matrix is the element of the matrix

- Each element in an $n \times n$ matrix has an associated **cofactor** defined as

$$a_{ij} = (-1)^{i+j} M_{ij}$$

  where M_{ij} is the determinant of the $(n-1) \times (n-1)$ matrix produced by deleting row i and column j from $\mathbf{A}$

- The determinant of an arbitrary $n \times n$ matrix $\mathbf{A}$ can be calculated by choosing an arbitrary row or column and summing the products of each element in that column/row with their cofactors, e.g. if we choose the first row of a 3×3 matrix then

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

2.5 Properties of Determinants

- The determinant of a matrix and its transpose are the same

- If any two rows/columns of a matrix are the same its determinant is zero

- If all the entries in a row/column of a matrix are zero, then its determinant is zero

- Interchanging any two rows/columns of a matrix negates its determinant

- Multiplying a row/column of a matrix by a nonzero real number k also multiplies the determinant by k

- If $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ matrices, then $\det \mathbf{AB} = \det \mathbf{A} \cdot \det \mathbf{B}$
- Adding a multiply of one row/column to another doesn't change the determinant
- The determinant of a triangular matrix is the product of the entries along the main diagonal
- Sometimes it's faster to calculate a matrix's determinant by reducing it to row-echelon form and multiplying the elements along the main diagonal than performing cofactor expansion
- Multiplying the entries of a row/column with the cofactors of another row/colum and summing the results always equals zero

2.6 Inverse of a Matrix

- Given an $n \times n$ matrix $\mathbf{A}$, if there exists another $n \times n$ matrix $\mathbf{B}$ such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$ then $\mathbf{A}$ is said to be **nonsingular** or **invertible** and $\mathbf{B}$ is said to be the unique **inverse** of $\mathbf{A}$, i.e. $\mathbf{B} = \mathbf{A}^{-1}$
- Some $n \times n$ matrices don't have an inverse and are called **singular**
- The **adjoint** of an $n \times n$ matrix $\mathbf{A}$ is the transpose of the matrix of cofactors corresponding to the entries of $\mathbf{A}$

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

- If $\mathbf{A}$ is an $n \times n$ matrix and $\det \mathbf{A} \neq 0$ then

$$\mathbf{A}^{-1} = \left(\frac{1}{\det \mathbf{A}}\right) \operatorname{adj} \mathbf{A}$$

- From the above, the inverse of a 2×2 matrix $\mathbf{A}$ is

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

- An $n \times n$ matrix $\mathbf{A}$ is nonsingular (has an inverse) if $\det \mathbf{A} \neq 0$

Theorem 8.6.4 Finding the Inverse

If an $n \times n$ matrix $\mathbf{A}$ can be transformed into the $n \times n$ identity $\mathbf{I}$ by a sequence of elementary row operations, then $\mathbf{A}$ is nonsingular. The same sequence of operations that transforms $\mathbf{A}$ into the identity $\mathbf{I}$ will also transform $\mathbf{I}$ into $\mathbf{A}^{-1}$.

- Inverse matrices can be used to solve linear systems. If $\mathbf{AX} = \mathbf{B}$ and $\mathbf{A}$ is invertible, then

$$\mathbf{A}^{-1}\mathbf{AX} = \mathbf{A}^{-1}\mathbf{B} \Rightarrow \mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$$

- When $\det \mathbf{A} \neq 0$ the solution of the system $\mathbf{AX} = \mathbf{B}$ is unique

- A homogeneous system of linear equations $\mathbf{AX} = \mathbf{0}$ has only the trivial solution iff $\mathbf{A}$ is nonsingular and an infinite number of solutions iff it is singular

2.7 Cramer's Rule

- If $\mathbf{A}$ is the coefficient matrix of a linear system and $\det \mathbf{A} \neq 0$, then the solution of the system is given by

$$\begin{aligned} x_1 &= \frac{\det \mathbf{A}_1}{\det \mathbf{A}} \\ x_2 &= \frac{\det \mathbf{A}_2}{\det \mathbf{A}} \\ &\vdots \\ x_n &= \frac{\det \mathbf{A}_n}{\det \mathbf{A}} \end{aligned}$$

where $\mathbf{A}_n$ is the matrix obtained by replacing column n of $\mathbf{A}$ with the constants of the system.

2.8 The Eigenvalue Problem

- If $\mathbf{A}$ is an $n \times n$ matrix, a number λ is said to be an **eigenvalue** of $\mathbf{A}$ if there exists a nonzero solution vector $\mathbf{K}$ of the linear system $\mathbf{AK} = \lambda\mathbf{K}$. The solution vector $\mathbf{K}$ is said to be an **eigenvector** corresponding to the eigenvalue λ.

- Rearranging the equation above we find

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{K} = \mathbf{0}$$

which only has nontrivial solutions if $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

- Calculating $\det(\mathbf{A} - \lambda\mathbf{I})$ results in an n-th degree polynomial in λ called the **characteristic equation** of $\mathbf{A}$, the solutions to which are its eigenvalues.

- The eigenvector associated with a particular eigenvalue can be found by applying Gauss-Jordan elimination to the augmented matrix $(\mathbf{A} - \lambda\mathbf{I}|\mathbf{0})$.

- A nonzero constant multiple of an eigenvector is another eigenvector.

- If λ is a complex eigenvalue of a matrix, then its conjugate $\lambda*$ is also an eigenvalue. If $\mathbf{K}$ is an eigenvector corresponding to λ then its conjugate $\mathbf{K}*$ is an eigenvector corresponding to $\lambda*$.

- $\lambda = 0$ is an eigenvalue of a matrix iff the matrix isn't invertible

- The determinant of a matrix is the product of its eigenvalues

- If λ is an eigenvalue of a matrix $\mathbf{A}$ with eigenvector $\mathbf{K}$, then $1/\lambda$ is an eigenvalue of $\mathbf{A}^{-1}$ with the same eigenvector.

- The eigenvalues of a triangular matrix are the entries along the main diagonal.

2.9 Powers of Matrices

- Any $n \times n$ matrix $\mathbf{A}$ satisfies its own characteristic equation, i.e. λ can be replaced with $\mathbf{A}$ in the characteristic equation.

- This gives us an expression for $\mathbf{A}^n$ as a linear combination

$$\mathbf{A}^n = c_0\mathbf{I} + c_1\mathbf{A} + c_2\mathbf{A}^2 + \cdots + c_{n-1}\mathbf{A}^{n-1}.$$

  If we multiply this expression by $\mathbf{A}$ we get an expression for $\mathbf{A}^{n+1}$ and we can replace the $\mathbf{A}^n$ term with the original expression. This can be repeated an arbitrary number of times to find expressions for any power of $\mathbf{A}$.

- The constants of the linear combination can be determined by substituting the matrix's eigenvalues into the characteristic equation, resulting in a linear system where the constants are the variables. Solving the system determines the constants.

- If $\mathbf{A}$ is a nonsingular matrix, the fact that it satisfies its own characteristic equation can be used to determine its inverse. This can be achieved by replacing λ with $\mathbf{A}$ in its characteristic equation, solving for $\mathbf{I}$, and multiplying both sides by $\mathbf{A}^{-1}$. This results in an expression for $\mathbf{A}^{-1}$ as a linear combination of powers of $\mathbf{A}$.

2.10 Orthogonal Matrices

- If $\mathbf{A}$ is a symmetric matrix with real entries, then the eigenvalues of $\mathbf{A}$ are real.

- If $\mathbf{A}$ is a symmetric matrix, then the eigenvectors corresponding to different eigenvalues are orthogonal.

- An $n \times n$ nonsingular matrix $\mathbf{A}$ is **orthogonal** if $\mathbf{A}^{-1} = \mathbf{A}^T$.

- An $n \times n$ matrix $\mathbf{A}$ is orthogonal iff its columns form an orthonormal set.

- If an $n \times n$ matrix $\mathbf{A}$ has n distinct eigenvalues, an orthogonal matrix can be formed by normalizing its eigenvectors and using them as column vectors in a new matrix.