Notes on Calculus

DSBA Mathematics Refresher 2024

Abstract

1 Linear Independence Property

Given a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V, the set is said to be **linearly independent** if the only solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

is

$$c_1 = c_2 = \cdots = c_n = 0.$$

In other words, none of the vectors in the set can be written as a linear combination of the others.

Example: Consider the vectors $\mathbf{v}_1 = \begin{pmatrix} 3 \\ 2 \\ 4 \\ -1 \\ 2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -2 \\ 5 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ and $\mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 5 \\ 3 \\ -4 \end{pmatrix}$ in $\mathbb{R}^5$. These vectors are linearly independent because the only solution to

$$c_1v_1 + c_2v_2 + c_3v_3 = \mathbf{0}$$

is $c_1 = c_2 = c_3 = 0$.

2 Spanning Property

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V is said to **span** V if every vector in V can be expressed as a linear combination of these vectors. Mathematically, the set spans V if for every vector $\mathbf{v} \in V$, there exist scalars $c_1, c_2, \ldots, c_n$ such that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n.$$

Example: In $\mathbb{R}^2$, the vectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ and $\mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ span $\mathbb{R}^2$ because any vector $\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}$ can be written as

$$\mathbf{v} = x\mathbf{v}_1 + (y - x)\mathbf{v}_2 + (y - x)\mathbf{v}_2.$$

3 Basis

A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ in a vector space V is called a **basis** for V if:

- The set is linearly independent.
- The set spans the vector space V.

The number of vectors in a basis is called the **dimension** of the vector space.

Example: The vectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ form a basis for $\mathbb{R}^2$. The dimension of $\mathbb{R}^2$ is 2.

4 (Reduced) Row Echelon Form

A matrix is in **row echelon form** if it satisfies the following conditions:

1. All non-zero rows are above any rows of all zeros.
2. The leading entry of each non-zero row after the first occurs to the right of the leading entry of the previous row.
3. The leading entry in any non-zero row is 1 (this condition is for the **reduced** row echelon form).
4. The leading 1 is the only non-zero entry in its column (for **reduced** row echelon form).

Example: The matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$$

is in row echelon form.

5 Inverting a Matrix

The **inverse** of a square matrix A is denoted by A^{-1} and is defined as the matrix that satisfies

$$AA^{-1} = A^{-1}A = I,$$

where I is the identity matrix.

Steps to find the inverse of a matrix:

1. Form the augmented matrix $[A|I]$.
2. Use row operations to convert A into the identity matrix.
3. The matrix that results from the identity matrix on the left side is A^{-1} on the right side.

Example: For the matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

the augmented matrix is

$$[A|I] = \begin{pmatrix} 1 & 2 & | & 1 & 0 \\ 3 & 4 & | & 0 & 1 \end{pmatrix}$$

after reducing to row echelon form:

$$[I|A^{-1}] = \begin{pmatrix} 1 & 0 & | & -2 & 1 \\ 0 & 1 & | & 1.5 & -0.5 \end{pmatrix}$$

hence, the inverse is

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ 1.5 & -0.5 \end{pmatrix}$$

6 Determinant

The **determinant** of a square matrix A, denoted by $\det(A)$, is a scalar value that can be computed from the elements of the matrix. The determinant has important properties and applications, including:

- A matrix is invertible if and only if its determinant is non-zero.
- The determinant of a product of matrices is the product of their determinants: $\det(AB) = \det(A)\det(B)$.

For a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the determinant is calculated as:

$$\det(A) = ad - bc.$$

For a 3×3 matrix $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$, the determinant is calculated as:

$$\det(A) = a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}.$$

Expanding the 2×2 determinants (minors):

$$\det(A) = a(ei - fh) - b(di - fg) + c(dh - eg).$$

For higher dimension matrices, ask a computer (it's outside the scope of this course).
Example: For the matrix
$$B = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix},$$
$\det B = 0$, hence, the matrix is not invertible (you can try to invert it, you will encounter a problem).

While for the matrix
$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$
$\det A = -2 \neq 0$, hence, the matrix is invertible (and we found the inverse in the previous section).

7 Eigenvalues

Let A be an $n \times n$ matrix. A scalar λ is called an **eigenvalue** of A if there exists a non-zero vector $\mathbf{v} \in \mathbb{R}^n$ (or $\mathbb{C}^n$) such that
$$A\mathbf{v} = \lambda \mathbf{v}.$$
The vector $\mathbf{v}$ is referred to as the corresponding **eigenvector** of A associated with λ.

Finding Eigenvalues To find the eigenvalues of a matrix A, we solve the characteristic equation:
$$\det(A - \lambda I) = 0,$$
where I is the identity matrix of the same size as A, and $\det(\cdot)$ denotes the determinant. The solutions λ are the eigenvalues of A.
Example: Consider the matrix
$$A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}.$$
The characteristic equation is given by:
$$\det(A - \lambda I) = \det\left(\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$
$$= \det\left(\begin{pmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{pmatrix}\right) = (4-\lambda)(3-\lambda) - 2 = \lambda^2 - 7\lambda + 10 = 0.$$
The eigenvalues are the roots of this quadratic equation, $\lambda_1 = 5$ and $\lambda_2 = 2$.

8 Eigenvectors

Given an eigenvalue λ of a matrix A, the corresponding **eigenvector $\mathbf{v}$** is any non-zero vector that satisfies the equation:

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

To find the eigenvectors associated with λ, we solve the system:

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

This is a system of linear equations.
Example: For the matrix A from the previous example and $\lambda_1 = 5$, we solve:

$$(A - 5I)\mathbf{v} = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix} \mathbf{v} = \mathbf{0}.$$

One solution to this system is $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ [1].

Similarly, for $\lambda_2 = 2$, we solve:

$$(A - 2I)\mathbf{v} = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \mathbf{v} = \mathbf{0},$$

which gives $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ [2].

9 Diagonalization

A square matrix A is said to be **diagonalizable** if there exists an invertible matrix P and a diagonal matrix D such that:

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

When this is possible, the diagonal elements of D are the eigenvalues of A, and the columns of P are the corresponding eigenvectors.

Procedure for Diagonalization

1. Find the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A.
2. For each eigenvalue λ_i, find the corresponding eigenvector $\mathbf{v}_i$.
3. Form the matrix P using the eigenvectors as columns: $P = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n \end{pmatrix}$.

[1] Any vector of the form $\mathbf{v} = \begin{pmatrix} \nu \\ \nu \end{pmatrix}, \nu \in \mathbb{R}^*$ would work, it is common practice to set $\nu = 1$.

[2] Again, any vector of the form $\mathbf{v} = \begin{pmatrix} -\nu \\ 2\nu \end{pmatrix}, \nu \in \mathbb{R}^*$ would work, it is common practice to set $\nu = 1$.

4. The matrix D is a diagonal matrix with eigenvalues on the diagonal:

$$D = \begin{pmatrix} \lambda_1 & & 0 & 0 & 0 \\ & \lambda_2 & & \ddots & 0 \\ 0 & & \ddots & & 0 \\ 0 & \ddots & & \ddots & \\ 0 & 0 & 0 & & \lambda_n \end{pmatrix}.$$

5. Verify that $A = PDP^{-1}$.

Example: Consider the matrix $A = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$. From our earlier work, the eigenvalues are $\lambda_1 = 5$ and $\lambda_2 = 2$, with corresponding eigenvectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

Thus, we can form:

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} 5 & 0 \\ 0 & 2 \end{pmatrix}.$$

Finally, verify that:

$$P^{-1} = \frac{1}{3} \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}, \quad A = PDP^{-1} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}.$$

10 Norms

In vector spaces, a **norm** is a function that assigns a non-negative length or size to vectors. Norms are widely used to measure the magnitude of vectors. The most common norms are the L_1, L_2, and L_p norms.

L_1 Norm (Manhattan Norm) The L_1 norm of a vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ is defined as the sum of the absolute values of its components:

$$\|\mathbf{v}\|_1 = |v_1| + |v_2| + \cdots + |v_n| = \sum_{i=1}^{n} |v_i|.$$

This norm is also known as the **Manhattan norm** or **taxicab norm**, as it represents the distance between two points in a grid-based path, like the blocks in a city.

Example: For the vector $\mathbf{v} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$ in $\mathbb{R}^3$, the L_1 norm is:

$$\|\mathbf{v}\|_1 = |3| + |-4| + |1| = 3 + 4 + 1 = 8.$$

L_2 Norm (Euclidean Norm) The L_2 norm of a vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ is the square root of the sum of the squares of its components:

$$\|\mathbf{v}\|_2 = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} = \sqrt{\sum_{i=1}^{n} v_i^2}.$$

This norm is commonly known as the **Euclidean norm** because it represents the usual notion of distance in Euclidean space.

Example: For the vector $\mathbf{v} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$ in $\mathbb{R}^3$, the L_2 norm is:

$$\|\mathbf{v}\|_2 = \sqrt{3^2 + (-4)^2 + 1^2} = \sqrt{9 + 16 + 1} = \sqrt{26}.$$

L_p Norm The L_p norm generalizes the L_1 and L_2 norms and is defined for any real number $p \geq 1$. The L_p norm of a vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ is given by:

$$\|\mathbf{v}\|_p = (|v_1|^p + |v_2|^p + \cdots + |v_n|^p)^{\frac{1}{p}} = \left(\sum_{i=1}^{n} |v_i|^p \right)^{\frac{1}{p}}.$$

Different values of p result in different norms. The L_2 norm is a special case where $p = 2$, and the L_1 norm is a special case where $p = 1$.

Example: For the vector $\mathbf{v} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$ in $\mathbb{R}^3$, the L_p norm with $p = 3$ is:

$$\|\mathbf{v}\|_3 = \left(|3|^3 + |-4|^3 + |1|^3\right)^{\frac{1}{3}} = (27 + 64 + 1)^{\frac{1}{3}} = \sqrt[3]{92}.$$

Properties of Norms To be a norm, a function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ must verify:

1. **Non-negativity:** $\|\mathbf{v}\| \geq 0$ for all vectors $\mathbf{v}$, and $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.
2. **Scalar Multiplication:** For any scalar α and vector $\mathbf{v}$, $\|\alpha\mathbf{v}\| = |\alpha|\|\mathbf{v}\|$.
3. **Triangle Inequality:** For any vectors $\mathbf{v}$ and $\mathbf{w}$, $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$.

11 The Gram-Schmidt Algorithm

We assume the following:

- The vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ are linearly independent in an inner product space V.
- The space V is equipped with an inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ which is usually the standard dot product in $\mathbb{R}^n$.

Given a set of linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, the goal is to find an orthogonal set of vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ such that:

- $\text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \text{span}(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ for all k,
- $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ for all $i \neq j$ (orthogonality),
- Optionally, we can normalize the vectors so that $\|\mathbf{u}_i\| = 1$, forming an orthonormal set.

Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be a set of linearly independent vectors. The Gram-Schmidt process proceeds as follows:

Step 1: Construct the first orthogonal vector
We begin by setting the first vector in the orthogonal set equal to the first vector in the original set:

$$\mathbf{u}_1 = \mathbf{v}_1$$

At this point, $\mathbf{u}_1$ is trivially orthogonal because it's the only vector in the set.

Step 2: Orthogonalize the remaining vectors
For each subsequent vector $\mathbf{v}_k$ (where $k = 2, 3, \ldots, n$), we subtract off the projections onto the previously calculated orthogonal vectors $\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}$.

The orthogonal vector $\mathbf{u}_k$ is computed as:

$$\mathbf{u}_k = \mathbf{v}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{v}_k, \mathbf{u}_j \rangle}{\langle \mathbf{u}_j, \mathbf{u}_j \rangle} \mathbf{u}_j$$

This formula ensures that $\mathbf{u}_k$ is orthogonal to each of the previous vectors $\mathbf{u}_1, \ldots, \mathbf{u}_{k-1}$.

Step 3: Normalize (Optional)
If we want an orthonormal set of vectors, we normalize each $\mathbf{u}_k$ by dividing it by its magnitude:

$$\mathbf{e}_k = \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}$$

The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ now form an orthonormal basis for the same subspace.