MA 106 : Linear Algebra
Course Syllabus and Evaluation Scheme

Department of Mathematics

Indian Institute of Technology Bombay

Syllabus

1. **Matrix operations and linear equations.** We shall study matrix operations. Use this knowledge to study the problem of solving systems of linear equations in any number of unknowns. The Gauss elimination method will be introduced which is one of the most efficient methods to solve a system of linear equations.

2. **Axiomatic approach to determinants of matrices.** We shall introduce determinants of matrices using the axiomatic approach. Determinants can be used to solve systems of linear equations, check invertibility of matrices and to find volumes of polytopes.

3. **Vector spaces and inner product spaces.** We shall introduce the fundamental notions of vector spaces and inner product spaces. Using the notion of dimension of a vector space, we will study rank of matrices. This will lead to a basic result called the rank-nullity theorem.

Syllabus

4. **Linear transformations.** Vector spaces are studied using linear transformations. Many operations such as matrix action on vectors, differentiation and integration of functions, rotations and reflections turn out to be examples of linear transformations. The concept of linear transformation helps us to study properties of matrices in a co-ordinate free manner.

5. **Eigenvalues and eigenvectors of matrices.** Matrices are simplified using eigenvalues and eigenvectors. These notions are important in all areas of pure and applied mathematics. We shall learn algorithms to diagonalise matrices of special types such as symmetric, Hermitian, orthogonal, unitary and more generally normal matrices. We will use results about diagonalisation of matrices for identification of curves and surfaces represented by quadratic equations.

Textbooks and References

(1) Serge Lang, Introduction to Linear Algebra, 2nd Ed. Springer, India.

(2) Gilbert Strang, Linear Algebra and its applications, Indian Edition, 2020

(3) M.K. Srinivasan and J.K. Verma, Notes on Linear Algebra, 2014

Plan for lectures and tutorials

Tut. D1, D2	Tut. D3, D4	Topic	hrs of lectures
Wed, 8 March	Wed, 8 March	Matrix operations	1
Wed, 15 March	Wed, 8 March	Linear equations	2
Tue, 21 March	Mon, 20 March	Determinant of matrices	2
Wed, 29 March	Wed, 29 March	Vector spaces	3
Wed, 5 April	Wed, 5 April	Linear Transformations	2
Wed, 12 April	Wed, 12 April	Inner Product Spaces	2
Tue, 18 April	Thu ,13 April	Eigenvalues	4

Evaluation scheme		
Mode	**Time**	**Marks**
2 quizzes	TBA	20
Final exam	19-26 April	30
Total		50

Notation for numbers and vectors

1. We shall use the following notation throughout this course.

$\mathbb{P} = \{1, 2, 3, \ldots\}$	$\mathbb{N} = \{0, 1, 2, 3, \ldots\}$
$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$	$\mathbb{Q}$ = the set of all rational numbers
$\mathbb{R}$ = the set of all real numbers	$\mathbb{C}$ = the set of all complex numbers
The n-dimensional Euclidean space	$\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_j \in \mathbb{R}, j = 1, \ldots, n\}$
The zero vector	$\mathbf{0} = (0, \ldots, 0)$

2. We use the following operations among vectors in $\mathbb{R}^n$.
For $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, and $\alpha \in \mathbb{R}$, we define

$$\begin{aligned} \text{(sum)} \quad \boldsymbol{x} + \boldsymbol{y} &= (x_1 + y_1, \ldots, x_n + y_n) \in \mathbb{R}^n, \\ \text{(scalar multiple)} \quad \alpha \boldsymbol{x} &= (\alpha x_1, \ldots, \alpha x_n) \in \mathbb{R}^n, \\ \text{(dot product)} \quad \boldsymbol{x} \cdot \boldsymbol{y} &= x_1 y_1 + \cdots + x_n y_n \in \mathbb{R}. \end{aligned}$$

Notation for matrices

1. Let $m, n \in \mathbb{N}$. An $m \times n$ **matrix A** with real entries is a rectangular array of real numbers arranged in m rows and n columns, written as:

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \cdots & \boxed{a_{jk}} & \cdots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} & \cdots & a_{mn} \end{bmatrix} = [a_{jk}],$$

2. where $a_{jk} \in \mathbb{R}$ is called the $(j, k)^{th}$ **entry** of **A** for $j = 1, \ldots, m$ and $k = 1, \ldots, n$.
3. Let $\boxed{\mathbb{R}^{m \times n}}$ denote the set of all $m \times n$ matrices with real entries.
4. If $\mathbf{A} := [a_{jk}]$ and $\mathbf{B} := [b_{jk}]$ are in $\mathbb{R}^{m \times n}$, then we define
5. $\mathbf{A} = \mathbf{B} \iff a_{jk} = b_{jk}$ for all $j = 1, \ldots, m$ and $k = 1, \ldots, n$.

Special types of matrices

1. **Definitions.** An $n \times n$ matrix is called a **square matrix** of size n.
2. A square matrix $\mathbf{A} = [a_{jk}]$ is called **symmetric** if $a_{jk} = a_{kj}$ for all j, k.
3. A square matrix $\mathbf{A} = [a_{jk}]$ is called **skew-symmetric** if $a_{jk} = -a_{kj}$ for all j, k.
4. A square matrix $\mathbf{A} = [a_{jk}]$ is called a **diagonal matrix** if $a_{jk} = 0$ for all $j \neq k$.
5. A diagonal matrix $\mathbf{A} = [a_{jk}]$ is called a **scalar matrix** if all diagonal entries of $\mathbf{A}$ are equal.
6. **Definitions.** The **identity matrix I** is a scalar matrix in which all diagonal elements are equal to 1, and the **zero matrix O** is the matrix with all entries equal to 0.
7. A square matrix $\mathbf{A} = [a_{jk}]$ is called **upper triangular** if $a_{jk} = 0$ for all $j > k$, and **lower triangular** if $a_{jk} = 0$ for all $j < k$.
8. **Remarks.** (1) A matrix $\mathbf{A}$ is upper triangular as well as lower triangular if and only if $\mathbf{A}$ is a diagonal matrix.
9. (2) Every diagonal entry of a skew-symmetric matrix is 0 since $a_{jj} = -a_{jj}$ $\implies a_{jj} = 0$ for $j = 1, \ldots, n$.

Examples of matrices

1. **Example.** $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ is symmetric and $\begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & 5 \\ -3 & -5 & 0 \end{bmatrix}$ is skew-symmetric.
2. **Example.** The matrix $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ is diagonal, while $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is scalar.
3. **Example.** The matrix $\begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 4 \end{bmatrix}$ is upper triangular,while the matrix $\begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & 4 \end{bmatrix}$ is lower triangular.

Row and column vector of matrices

1. **Definitions.** A **row vector a** of length n is a matrix with only one row consisting of n scalars. It is written as

$$\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_k & \cdots & a_n \end{bmatrix}.$$

2. A **column vector b** of length n is a matrix with only one column. it is written as

$$\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_k \\ \vdots \\ b_n \end{bmatrix}$$

3. **Definitions.** Suppose $\mathbf{A} := [a_{jk}]$ and $\mathbf{B} := [b_{jk}]$ are $m \times n$ matrices. Then the $m \times n$ matrix $\mathbf{A} + \mathbf{B} := [a_{jk} + b_{jk}]$ is called the **sum** of **A** and **B**.
4. If $\alpha \in \mathbb{R}$, then $\alpha\mathbf{A} := [\alpha a_{jk}]$ is called the **scalar multiple** of **A** by α.

Operations on Matrices

1. **Proposition.** The addition and scalar multiplication satisfy the following properties for all matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and scalars α and β.
   1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
   2. $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$, which we write as $\mathbf{A} + \mathbf{B} + \mathbf{C}$
   3. $\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$
   4. $(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$
   5. $\alpha(\beta\mathbf{A}) = (\alpha\beta)\mathbf{A}$, which we write as $\alpha\beta\mathbf{A}$
2. **Notation.** We write $(-1)\mathbf{A}$ as $-\mathbf{A}$, and $\mathbf{A} + (-\mathbf{B})$ as $\mathbf{A} - \mathbf{B}$.
3. (2) The **transpose** of an $m \times n$ matrix $\mathbf{A} := [a_{jk}]$ is defined to be the $n \times m$ matrix $\mathbf{A}^{\mathsf{T}} := [a_{kj}]$.
4. **Remarks.** The row vectors of $\mathbf{A}$ are the column vectors of $\mathbf{A}^{\mathsf{T}}$.
5. It is easy to prove that
$$(\mathbf{A}^{\mathsf{T}})^{\mathsf{T}} = \mathbf{A}, \quad (\mathbf{A} + \mathbf{B})^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} + \mathbf{B}^{\mathsf{T}} \text{ and } (\alpha\mathbf{A})^{\mathsf{T}} = \alpha\mathbf{A}^{\mathsf{T}}.$$
6. Note that A square matrix $\mathbf{A}$ is symmetric $\Longleftrightarrow \mathbf{A}^{\mathsf{T}} = \mathbf{A}$.
7. A square matrix $\mathbf{A}$ is skew-symmetric $\Longleftrightarrow \mathbf{A}^{\mathsf{T}} = -\mathbf{A}$.
8. The transpose of a row vector is a column vector, and vice versa.

Linear combinations of row and column vectors

1. Let $m, n \in \mathbb{N}$. Let $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ and $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^{1\times n}$, then $\alpha_1 \mathbf{a}_1 + \cdots + \alpha_m \mathbf{a}_m$ is called a **linear combination** of $\mathbf{a}_1, \ldots, \mathbf{a}_m$.
2. Similarly, if $\mathbf{b}_1, \ldots, \mathbf{b}_m \in \mathbb{R}^{n\times 1}$, then $\alpha_1 \mathbf{b}_1 + \cdots + \alpha_m \mathbf{b}_m \in \mathbb{R}^{n\times 1}$ is called a **linear combination** of $\mathbf{b}_1, \ldots, \mathbf{b}_m$.
3. For $k = 1, \ldots, n$, the vector $\mathbf{e}_k := \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}^{\mathsf{T}}$, where the kth entry is 1 and all other entries are 0.
4. If $\mathbf{b} = \begin{bmatrix} b_1 & \cdots & b_k & \cdots & b_n \end{bmatrix}^{\mathsf{T}}$ is a column vector then $\mathbf{b} = b_1 \mathbf{e}_1 + \cdots + b_k \mathbf{e}_k + \cdots + b_n \mathbf{e}_n$, is a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$.
5. The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are called **basic column vectors** in $\mathbb{R}^{n\times 1}$.
6. Let $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m\times n}$. Then $\mathbf{A}_j := \begin{bmatrix} a_{j1} & \cdots & a_{jn} \end{bmatrix} \in \mathbb{R}^{1\times n}$ is called the jth **row vector of A** for $j = 1, \ldots, m$, and we write $\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_m \end{bmatrix}$.
7. $A^k := \begin{bmatrix} a_{1k} & \cdots & a_{mk} \end{bmatrix}^{\mathsf{T}}$ is called the kth **column vector of A** for $k = 1, \ldots, n$, and we write $\mathbf{A} = \begin{bmatrix} \mathbf{A}^1 & \cdots & \mathbf{A}^n \end{bmatrix}$.

Lecture 2: Matrix Multiplication

1. Before we discuss product of matrices we define the product of a row vector $\mathbf{a} \in \mathbb{R}^{\mathbf{1}\times\mathbf{n}}$ by a column vector $\mathbf{b} \in \mathbb{R}^{\mathbf{n}\times\mathbf{1}}$.
2. **Definitions.** Suppose $\mathbf{a} := \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$ and $\mathbf{b} := \begin{bmatrix} b_1 & \cdots & b_n \end{bmatrix}^{\mathsf{T}}$.
3. The **product** of the row vector $\mathbf{a}$ with the column vector $\mathbf{b}$ is defined as

$$\mathbf{ab} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} := a_1 b_1 + \cdots + a_n b_n \in \mathbb{R}.$$

4. Let $\mathbf{A} \in \mathbb{R}^{m\times n}$. Then $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix}$, where $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^{1\times n}$.

Matrix Multiplication

1. Recalling that $\mathbf{b} \in \mathbb{R}^{n\times 1}$, we define

$$\mathbf{A}\mathbf{b} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} \mathbf{b} := \begin{bmatrix} \mathbf{a}_1\mathbf{b} \\ \vdots \\ \mathbf{a}_m\mathbf{b} \end{bmatrix} \in \mathbb{R}^{m\times 1}.$$

2. Suppose $\mathbf{B} \in \mathbb{R}^{n\times p}$. Then $\mathbf{B} = [\mathbf{b}_1 \cdots \mathbf{b}_p]$, where $\mathbf{b}_1, \ldots, \mathbf{b}_p \in \mathbb{R}^{n\times 1}$.
3. Note that $\mathbf{A}\mathbf{b}_1, \ldots, \mathbf{A}\mathbf{b}_p \in \mathbb{R}^{m\times 1}$, we define
4. $$\mathbf{A}\mathbf{B} = \mathbf{A}[\mathbf{b}_1 \cdots \mathbf{b}_p] := [\mathbf{A}\mathbf{b}_1 \cdots \mathbf{A}\mathbf{b}_p] \in \mathbb{R}^{m\times p}.$$
5. Therefore $\mathbf{A}\mathbf{B} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{bmatrix} [\mathbf{b}_1 \cdots \mathbf{b}_p] = \begin{bmatrix} \mathbf{a}_1\mathbf{b}_1 & \cdots & \mathbf{a}_1\mathbf{b}_p \\ \vdots & & \vdots \\ \mathbf{a}_m\mathbf{b}_1 & \cdots & \mathbf{a}_m\mathbf{b}_p \end{bmatrix} \in \mathbb{R}^{m\times p}.$

Matrix Multiplication

1. Suppose $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m\times n}$, $\mathbf{B} := [b_{jk}] \in \mathbb{R}^{n\times p}$, then $\mathbf{AB} \in \mathbb{R}^{m\times p}$, and for $j = 1, \ldots, m$; $k = 1 \ldots, p$,

$$\mathbf{AB} = [c_{jk}], \quad \text{where } c_{jk} := \mathbf{a}_j\mathbf{b}_k = \sum_{\ell=1}^{n} a_{j\ell}b_{\ell k}.$$

2. Note that the (j,k)th entry of $\mathbf{AB}$ is a product of the jth row vector of $\mathbf{A}$ with the kth column vector of $\mathbf{B}$ as shown below:

$$\begin{bmatrix} \\ a_{j1} & \cdots & a_{j\ell} & \cdots & a_{jn} \\ \\ \end{bmatrix} \begin{bmatrix} b_{1k} \\ \vdots \\ b_{\ell k} \\ \vdots \\ b_{nk} \end{bmatrix}.$$

3. Clearly, the product $\mathbf{AB}$ is defined only when the number of columns of $\mathbf{A}$ is equal to the number of rows of $\mathbf{B}$.
4. Note that $\mathbf{AI} = \mathbf{A}, \mathbf{IA} = \mathbf{A},\ \mathbf{AO} = \mathbf{O}$ and $\mathbf{OA} = \mathbf{O}$, whenever these products are defined.

Matrix Multiplication : Examples

1. Let $\mathbf{A} := \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \end{bmatrix}_{2\times 3}$, $\mathbf{B} := \begin{bmatrix} 1 & 6 & 0 & 2 \\ 2 & -1 & 1 & -2 \\ 2 & 0 & -1 & 1 \end{bmatrix}_{3\times 4}$.
2. Then $\mathbf{AB} = \begin{bmatrix} 2 & 11 & 2 & 1 \\ 8 & -3 & 2 & -5 \end{bmatrix}_{2\times 4}$.
3. Both products $\mathbf{AB}$ and $\mathbf{BA}$ are defined $\iff$ $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$.
4. In general, $\mathbf{AB} \neq \mathbf{BA}$. For example, if $\mathbf{A} := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\mathbf{B} := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then
5. $\mathbf{AB} := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, while $\mathbf{BA} := \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
6. Note that $\mathbf{BA} = \mathbf{O}$, while $\mathbf{AB} = \mathbf{B} \neq \mathbf{O}$. Since $\mathbf{A} \neq \mathbf{I}$, we see that the so-called cancellation law does not hold.

Matrix multiplication

1. Suppose that $\mathbf{A} := [a_{jk}] \in \mathbb{R}^{m\times n}$, and $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^{n\times 1}$. Then for all $k = 1, 2, \ldots, n$.

$$\mathbf{A}\,\mathbf{e}_k = \begin{bmatrix} a_{1k} \\ \vdots \\ a_{jk} \\ \vdots \\ a_{mk} \end{bmatrix}$$

2. This follows from our definition of matrix multiplication.
3. Also, the kth column of $\mathbf{AB}$ ia a linear combination of the columns of $\mathbf{A}$ with coefficients from the kth column of $\mathbf{B}$.
4. If $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$, then $\mathbf{A} = \mathbf{B} \iff \mathbf{A}\mathbf{e}_k = \mathbf{B}\mathbf{e}_k$ for each $k = 1, \ldots, n$.
5. Consider matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\alpha \in \mathbb{R}$. Then it is easy to see that

$$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}, \text{ and } \mathbf{C}(\mathbf{A}+\mathbf{B}) = \mathbf{CA} + \mathbf{CB}$$

and $(\alpha\mathbf{A})\mathbf{B} = \alpha\mathbf{AB} = \mathbf{A}(\alpha\mathbf{B})$, if sums and products are well-defined.

Properties of Matrix Multiplication

1. We prove that the matrix multiplication satisfies the **associative law.**
2. Proposition. If $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$ and $C \in \mathbb{R}^{p \times q}$, then

$$A(BC) = (AB)C$$

3. Proof. Write $A = (a_{jk}), B = (b_{jk}), C = (c_{jk})$. We show that $(A(BC))_{jk} = ((AB)C)_{jk}$ for all j, k. This is clear from the equations below.

$$(A(BC)_{jk} = \sum_{r=1}^{n} a_{jr}(BC)_{rk} = \sum_{r=1}^{n} a_{jr} \sum_{s=1}^{p} b_{rs}c_{sk} = \sum_{s=1}^{p} \left[\sum_{r=1}^{n} a_{jr}b_{rs} \right] c_{sk} = ((AB)C)_{jk}.$$

4. **Proposition.** If $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times p}$, then $(AB)^T = B^T A^T$.
5. Proof. For any matrix C, we write its i^{th} row as C_i and its j^{th} column as C^j.
6. We show that $((AB)^T)_{ij} = (B^T A^T)_{ij}$ for all i, j. This follows from the following:

$$((AB)^T)_{ij} = (AB)_{ji} = A_j B^i = (B^i)^T (A_j)^T = (B^T)_i (A^T)^j = (B^T A^T)_{ij}.$$