PHY 564
Advanced Accelerator Physics
Lecture 3:
Review of Linear Algebra

Gang Wang
Vladimir N. Litvinenko
Yichao Jing

CENTER for ACCELERATOR SCIENCE AND EDUCATION
Department of Physics & Astronomy, Stony Brook University
Collider-Accelerator Department, Brookhaven National Laboratory

Matrix: definition and properties

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

Addition: $A + B = C \Leftrightarrow a_{ij} + b_{ij} = c_{ij}$

Multiplied by a constant: $kA = B \Leftrightarrow ka_{ij} = b_{ij}$

Equality: $A = B \Leftrightarrow a_{ij} = b_{ij}$

Multiplication (inner product): $AB = C \Leftrightarrow \sum_k a_{ik} b_{kj} = c_{ij}$

$$(AB)C = A(BC), \qquad A(B+C) = AB + AC$$

In general $AB \neq BA$

Multiplication demands that A has the same number of columns as B has rows.

Matrix: special cases I

- Diagonal matrix:

$$a_{ij} = 0 \quad \text{for} \;\; i \neq j$$

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & 0 & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

If A and B are both diagonal matrix, they are commutative:

$$AB = BA$$

- Identity matrix:

$$AI = IA = A \qquad \text{for} \quad \forall \text{A}$$

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & 0 & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

$$I_{ij} = \delta_{ij} = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases}$$

Matrix: special cases II

- Block diagonal matrix: A and A_i are square matrix.

$$A = \begin{bmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A_k \end{bmatrix},$$

$$A = \left[\begin{array}{ccc|c|cc} 1 & 3 & 2 & 0 & 0 & 0 \\ 7 & 0 & 2 & 0 & 0 & 0 \\ 1 & 1 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 6 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 3 & 3 \end{array}\right]$$

- Triangular matrix:

Upper diagonal matrix: elements below diagonal are all zero

$$U = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}$$

Lower diagonal matrix: elements below diagonal are all zero

$$L = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

Matrix V: transpose matrix

- A matrix, B, is called the transpose matrix of a matrix A if

$$A_{ij} = B_{ji}$$

The transpose matrix is often denoted as A^T,

i.e. $$A_{ij} = \left(A^T\right)_{ji}$$

- A square matrix A is called an orthognal matrix if
- $$A^T = A^{-1}$$
- A square matrix A is called symmetric matrix if

$A^T = A$ and anti-symmetric if $A^T = -A$

Matrix: trace

- In any square matrix, the sum of the diagonal elements is called the trace.

$$Tr(A) = \sum_i a_{ii}$$

- A useful property: $Tr(AB) = Tr(BA)$

- In general, $Tr(ABC) = Tr(BCA) \neq Tr(BAC)$

- Trace is a linear operator:

$$Tr(A + kB) = Tr(A) + k \cdot Tr(B)$$

Matrix: determinant of a matrix

- For a square matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

- The determinant

$$D = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \det(A)$$

is called the determinant of matrix A and is denoted by det(A).

Determinant I

$$D = \underbrace{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}}_{\text{n columns}} \Big\} \text{n rows} \quad = \sum_{i,j,k} \varepsilon_{ijk\dots} a_{1i} a_{2j} a_{3k} \cdots$$

$\varepsilon_{ijk\cdots}$ is Levi-Civita symbol

$$\varepsilon_{ijk\cdots} = \begin{cases} 1 & \text{if } (i,j,k\dots) \text{ is even permutation of } (1,2,3\dots) \\ -1 & \text{if } (i,j,k\dots) \text{ is odd permutation of } (1,2,3\dots) \\ 0 & \text{if any of the two indices is repeated} \end{cases}$$

$$\varepsilon_{ijk\cdots l\cdots m\cdots} = -\varepsilon_{ijk\cdots m\cdots l\cdots}$$

Determinant II

A determinant of n dimension can be expanded over a column (or a row) into a sum of n determinants of n-1 dimension:

$$D = \begin{vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nj} & \dots & \dots \end{vmatrix} = \sum_i C_{ij} a_{ij}$$

n-1 columns

$$M_{ij} = \begin{vmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix}$$

n-1 rows

$C_{ij} = (-1)^{i+j} M_{ij}$ is called the ijth cofactor of D.

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

Determinant III

- Multiplied by a constant

$$\begin{vmatrix} ka_{11} & a_{12} & a_{13} \\ ka_{21} & a_{22} & a_{23} \\ ka_{31} & a_{32} & a_{33} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

- The value of a determinant is unchanged if a multiple of one column (row) is added to another column (row)

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11}+ka_{12} & a_{12} & a_{13} \\ a_{21}+ka_{22} & a_{22} & a_{23} \\ a_{31}+ka_{32} & a_{32} & a_{33} \end{vmatrix}$$

- A determinant is equal to zero if any two columns (rows) are proportional

$$\begin{vmatrix} a_{12} & ka_{12} & a_{13} \\ a_{22} & ka_{22} & a_{23} \\ a_{32} & ka_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = 0$$

Linear equation system

- Existence of non-trivial solution of homogeneous equations

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = 0 \\ a_{21}x + a_{22}y + a_{23}z = 0 \\ a_{31}x + a_{32}y + a_{33}z = 0 \end{cases}$$

$$x \cdot D \equiv x \cdot \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11}x + a_{12}y + a_{13}z & a_{12} & a_{13} \\ a_{21}x + a_{22}y + a_{23}z & a_{22} & a_{23} \\ a_{31}x + a_{32}y + a_{33}z & a_{32} & a_{33} \end{vmatrix} = 0$$

Similarly, $y \cdot D = 0$ and $z \cdot D = 0$

- Thus a set of homogenous linear equations have non-trivial solutions only if the determinant of the coefficients, D, vanishes.

Matrix: properties of the determinant of a matrix

- Some properties of the determinant of matrices
  - $\det\left(A^T\right)=\det\left(A\right)$
  - $\det\left(kA\right)=k^n\det\left(A\right)$
  - $\det\left(AB\right)=\det\left(A\right)\det\left(B\right)$

Proof of det(AB)=det(A)det(B):

(1)

$$\sum_{i,j,k\dots}\varepsilon_{ijk\dots}a_{\beta i}a_{\alpha j}a_{\gamma k}\dots$$
$$=\sum_{j,i,k\dots}\varepsilon_{jik\dots}a_{\beta j}a_{\alpha i}a_{\gamma k}\dots$$
$$=\sum_{i,j,k\dots}\varepsilon_{jik\dots}a_{\alpha i}a_{\beta j}a_{\gamma k}\dots$$
$$=-\sum_{i,j,k\dots}\varepsilon_{ijk\dots}a_{\alpha i}a_{\beta j}a_{\gamma k}\dots$$

(2)

$$\sum_{i,j,k\dots}\varepsilon_{ijk\dots}a_{\alpha i}a_{\beta j}a_{\gamma k}\dots$$
$$=\varepsilon_{\alpha\beta\gamma\dots}\sum_{i,j,k\dots}\varepsilon_{ijk\dots}a_{1i}a_{2j}a_{3k}\dots$$
$$=\varepsilon_{\alpha\beta\gamma\dots}\left|A\right|$$

(3)

$$\left|AB\right|=\sum_{i,j,k\dots}\varepsilon_{ijk\dots}\left(AB\right)_{1i}\left(AB\right)_{2j}\left(AB\right)_{3k}\dots$$
$$=\sum_{i,j,k\dots}\sum_{\alpha,\beta,\gamma}\varepsilon_{ijk\dots}A_{1\alpha}B_{\alpha i}A_{2\beta}B_{\beta j}A_{2\gamma}B_{\gamma j}\dots$$
$$=\sum_{\alpha,\beta,\gamma}A_{1\alpha}A_{2\beta}A_{2\gamma}\dots\left\{\sum_{i,j,k\dots}\varepsilon_{ijk\dots}B_{\alpha i}B_{\beta j}B_{\gamma j}\dots\right\}$$
$$=\left|B\right|\sum_{\alpha,\beta,\gamma}\varepsilon_{\alpha\beta\gamma\dots}A_{1\alpha}A_{2\beta}A_{2\gamma}\dots$$
$$=\left|A\right|\left|B\right|$$

Matrix IV: inversion

- Inversion of a square matrix A is to find a square matrix B such that

$$AB = BA = I$$

B is called the inverse matrix of A and often denoted by A^{-1}, i.e.

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

- One way to find the inverse matrix is by

$$\left(A^{-1}\right)_{ij} = \frac{C_{ji}}{|A|}, \quad \text{where } C_{ji} \text{ is the } ji^{th} \text{ cofactor of } A$$

Matrix VI: similarity transformation and diagonalization

- Two matrix, A and B, are called similar if there exists a invertible matrix P such that

$$B = P^{-1}AP,$$

and the transformation from A to B is called similarity transformation.

- Diagnolization of a matrix, A, is to find a similarity transformation matrix, P, such that

$P^{-1}AP$ is a diagonal matrix:

Matrix VII: diagonalization

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \lambda_n \end{pmatrix} \quad \text{or} \quad AP = P\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

- If we look at the jth column of the second equation, it follows

$$\sum_k a_{ik} p_{kj} = \sum_k p_{ik} \lambda_k \delta_{kj} = \lambda_j p_{ij} \qquad (*)$$

Defining a nx1 matrix (i.e. a column vector) $\left|P^J\right\rangle$ such that $\left|P^j\right\rangle_i = p_{ij}$ (note: j is fixed)

Equation (*) becomes: $A\left|P^j\right\rangle = \lambda_j \left|P^j\right\rangle$

Matrix VIII: eigenvalue and eigenvector

- For a matrix A, a vector matrix X is called an eigenvector of A if

$$A \cdot X = \lambda X$$

where λ is called the eigenvalue associated with the eigenvector X.

- The eigenvalues are found by solving the following polynomial equation

$$(A - \lambda I) \cdot X = 0 \Rightarrow \det(A - \lambda I) = 0$$

Defective Matrix

- Not all square matrix can be diagonalized:

$$A = \begin{pmatrix} 2 & -3 \\ 3 & -4 \end{pmatrix}; \ \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & -3 \\ 3 & -4-\lambda \end{vmatrix} = 0 \Rightarrow (\lambda+1)^2 = 0 \Rightarrow \lambda = -1$$

$$\begin{pmatrix} 2 & -3 \\ 3 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow \begin{cases} 3x_1 - 3x_2 = 0 \\ 3x_1 - 3x_2 = 0 \end{cases} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix};$$

We end up with only one eigenvector.

- A square matrix that does not have a complete set of eigenvectors is not diagonalizable and is called a defective matrix.
- If a matrix, A, is defective (and hence is not similar to a diagonal matrix), then what is the simplest matrix that A is similar to?

Jordan form matrix

- Definition: a Jordan block with value λ is a square, upper triangular matrix whose entries are all λ on the diagonal, all 1 on the entries immediately above the diagonal, and zero elsewhere:

$$J(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix}$$

1D: $[\lambda]$

2D: $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$

3D: $\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$

- Definition: a Jordan form matrix is a block diagonal matrix whose blocks are all Jordan blocks

- Theorem: Let A be a nxn matrix. Then there is a Jordan form matrix that is similar to A.

$$\left[\begin{array}{cc|cc|c|c} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right]$$

Symplectic Matrix

- As long as the system has a Hamiltonian, the Jacobian matrix, $M_{\alpha\beta} = \frac{\partial X_\alpha}{\partial (X_0)_\beta}$, which describe the motion of the particles, satisfies

$$M^T S M = S \qquad (**)$$

$$S = \begin{pmatrix} S_{1D} & 0 & \dots & 0 \\ 0 & S_{1D} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & S_{1D} \end{pmatrix}; \qquad S_{1D} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

- A matrix satisfying condition (**) is called a symplectic matrix.
  - inverse: $M^T S M = S \Rightarrow S M^T S M = S^2 = -I \Rightarrow \left(-S M^T S\right) M = I \Rightarrow M^{-1} = -S M^T S$
  - if M and N are both symplectic, then their product, MN, is also symplectic $(MN)^T S (MN) = N^T M^T S M N = N^T S N = S$
  - if M is symplectic, M^T is also symplectic

$$\left(M^T S M\right)^{-1} = -S \Rightarrow M^{-1} S \left(M^T\right)^{-1} = S \Rightarrow S = M S M^T \Rightarrow \left(M^T\right)^T S M^T = S$$

Symplectic Matrix II

- If λ is eigen value then $1/\lambda$ is also an eigenvalue and the multiplicity of λ and $1/\lambda$ is the same.
  - It implies that the eigenvalues are coming in pairs $\{\lambda, 1/\lambda\}$.
- As a consequence of above property, the determinant of a symplectic matrix is 1.

Symplectic Matrix

- If a motion of a particle in n-D space can be described by a Hamiltonian, H, it follows

$$\dot{x}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial x_i} \qquad \text{for} \quad i = 1,2...n$$

We can write above equations into a matrix form:

$$\dot{X} = S\frac{\partial H}{\partial X} \Leftrightarrow \left(\dot{X}\right)_\alpha = S_{\alpha\beta}\left(\frac{\partial H}{\partial X}\right)_\beta \quad \text{for} \quad \alpha,\beta,\gamma = 1,\ 2,\ ...,\ 2n$$

$$X = \begin{pmatrix} x_1 \\ p_1 \\ \cdots \\ \cdots \\ x_n \\ p_n \end{pmatrix}; \quad \dot{X} = \begin{pmatrix} \dot{x}_1 \\ \dot{p}_1 \\ \cdots \\ \cdots \\ \dot{x}_n \\ \dot{p}_n \end{pmatrix}; \quad S = \begin{pmatrix} S_{1D} & 0 & ... & 0 \\ 0 & S_{1D} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & 0 & S_{1D} \end{pmatrix}; \ S_{1D} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \frac{\partial H}{\partial X} = \begin{pmatrix} \frac{\partial H}{\partial x_1} \\ \frac{\partial H}{\partial p_1} \\ ... \\ ... \\ \frac{\partial H}{\partial x_n} \\ \frac{\partial H}{\partial p_n} \end{pmatrix}$$

Symplectic Matrix

- Let M to be the Jacobian matrix of a map (for linear motion, this is the transfer matrix)

$$M_{\alpha\beta} = \frac{\partial X_\alpha}{\partial (X_0)_\beta}$$

where X is the coordinate vector at some final location s and X_0 is the cooridnates vector at the initial location s=0

- It can be shown that $\frac{d}{ds}\left(M^T S M\right) = 0$, which means $M^T S M$ does not change with s. When the particle is still at its initial location, M is a unit matrix, and hence $M^T S M = I^T S I \Rightarrow M^T S M = S$