Matrix Operations I [gmd6-A]

Matrices are at the core of linear algebra. Sets of linear equations need solving in most areas of physics, most prominently in quantum mechanics.

Matrices are lists of elements (real or complex numbers) arranged in m rows and n columns. Each element has two indices, the first stating its row and the second its column. The order of a matrix is $m \times n$.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix},$$

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} d_{11} \\ d_{21} \\ d_{31} \end{pmatrix}, \quad \mathbf{E} = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \\ e_{31} & e_{32} \end{pmatrix},$$

$$\mathbf{F} = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{33} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}, \quad \mathbf{G} = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{33} \\ g_{31} & g_{32} & g_{33} \end{pmatrix}.$$

Matrix $\mathbf{C}$ is a *row vector* and matrix $\mathbf{D}$ a *column vector*. The redundant index of their elements is omitted if they are interpreted as vectors.

Matrices $\mathbf{F}$ and $\mathbf{G}$ are each a *square matrix*. Vectors and square matrices are more important by far in physics than rectangular matrices.

Fundamental properties, operations, and relations:

- *Equality*: Two matrices are equal if they are of the same order (e.g. $\mathbf{A}$ and $\mathbf{B}$ or $\mathbf{F}$ and $\mathbf{G}$) and have identical elements.
- *Addition*: Only matrices of the same order can be added or subtracted. The elements of $\mathbf{A} \pm \mathbf{B}$, for example, are $a_{ij} \pm b_{ij}$.
- *Scaling*: Multiplying a matrix with a real or complex number λ means multiplying each element of the matrix with that number.
- *Transpose*: Transposing a matrix changes its order from $m \times n$ to $n \times m$ with the row and column indices of all elements interchanged.

  $\mathbf{E}$ is the transpose of $\mathbf{B}$ if $b_{ij} = e_{ji}$. A row vector transposes into a column vector and vice versa. The transpose of a transpose is the original matrix.

- *Multiplication*: Matrix multiplication is not commutative. If the first factor has order $m \times p$ then the second factor must have order $p \times n$ and the product matrix has order $m \times n$.

  The product $\mathbf{AD}$ yields a 2×1 column vector and the product $\mathbf{CE}$ yields a 1×2 row vector.

  The product $\mathbf{AE}$ yields a 2×2 square matrix whereas the product $\mathbf{EA}$ yields a 3×3 square matrix.

  The matrix $\mathbf{P} = \mathbf{AE}$, for example, has elements, $p_{ij} = \sum_k a_{ik} e_{kj}$.

  Matrix multiplication is associative if the factors are compatible for multiplication: $(\mathbf{CG})\mathbf{D} = \mathbf{C}(\mathbf{GD})$ and $(\mathbf{AF})\mathbf{E} = \mathbf{A}(\mathbf{FE})$. Powers of square matrices are possible: $\mathbf{F}^2 = \mathbf{FF}$ and $\mathbf{G}^3 = \mathbf{GGG}$.

  Transpose of a product: $(\mathbf{FG})^T = \mathbf{G}^T\mathbf{F}^T$ and $(\mathbf{AFG})^T = \mathbf{G}^T\mathbf{F}^T\mathbf{A}^T$.
- *Symmetric matrix*: $\mathbf{H}^T = \mathbf{H}$. An $n \times n$ symmetric square matrix has $n(n+1)/2$ independent elements: those on the diagonal and on one side of the diagonal.
- *Antiymmetric matrix*: $\mathbf{J}^T = -\mathbf{J}$. An $n \times n$ antisymmetric (or skew-symmetric) square matrix has $n(n-1)/2$ independent elements: those on one side of the diagonal. The diagonal elements are zero.
- *Decomposition*: Any real square matrix $\mathbf{F}$ can be expressed as the sum of a symmetric and an antisymmetric matrix:

$$\mathbf{F} = \mathbf{F}_\mathrm{s} + \mathbf{F}_\mathrm{a}, \quad \mathbf{F}_\mathrm{s} = \frac{1}{2}\big(\mathbf{F} + \mathbf{F}^T\big), \quad \mathbf{F}_\mathrm{a} = \frac{1}{2}\big(\mathbf{F} - \mathbf{F}^T\big).$$

- *Complex conjugate matrix*: If the elements a_{ij} of the matrix $\mathbf{A}$ are complex numbers then the complex conjugate matrix $\bar{\mathbf{A}}$ has complex conjugate elements a_{ij}^*. The *conjugate transpose* $\bar{\mathbf{A}}^T$ has elements a_{ji}^*.
- *Hermitian matrix*: Complex square matrix $\mathbf{F}$ with $\bar{\mathbf{F}}^T = \mathbf{F}$.
- *Unit matrix*: Square matrix $\mathbf{I}$ with elements δ_{ik} (Kronecker delta).
- *Null matrix*: Matrix $\mathbf{0}$ with of any order with all elements zero.
- *Diagonal matrix*: A square matrix $\mathbf{F}$ with zero off-diagonal elements.
- *Trace*: Sum of diagonal elements of a square matrix: $\mathrm{Tr}[\mathbf{F}] = \sum_i f_{ii}$.

Determinants:

Determinants are sums of products of elements of square matrices constructed as shown here for matrices of orders $n = 2, 3$.

$$\mathrm{Det}[\mathbf{A}] = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\mathrm{Det}[\mathbf{B}] = \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$
$$= b_{11}b_{22}b_{33} + b_{12}b_{23}b_{31} + b_{13}b_{21}b_{32} - b_{13}b_{22}b_{31} - b_{12}b_{21}b_{33} - b_{11}b_{23}b_{32}.$$

- The number of terms is $n!$, which is even for $n \geq 2$. Each term is the product of n elements from different rows and columns.
- Each term can be positive, zero, negative, or complex, depending on the value of the elements.
- Half the terms come with a plus sign and the other half with a minus sign.
- If the first index is ordered, $\{1, 2, \ldots, n\}$, and the second index an even (odd) permutation of $\{1, 2, \ldots, n\}$, then the term comes with a plus (minus) sign.
- Even (odd) permutations require an even (odd) number of transpositions of nearest-neighbor indices in the sequence.

Laplace expansion of a determinant:

- The *minor* associated with element h_{ik} of a square matrix $\mathbf{H}$ is the determinant constructed from the elements which omit those in row i and column j.
- The minor associated with element h_{ik} multiplied with the sign factor $(-1)^{i+j}$ is named the *cofactor* H_{ij}.[1] The cofactor matrix is $\mathbf{H}_{ij}$.
- The determinant of $\mathbf{H}$ can be reconstructed from the cofactors and elements of one row or one column:

$$\mathrm{Det}[\mathbf{H}] = \sum_j h_{ij}H_{ij} = \sum_i h_{ij}H_{ij}.$$

[1] Watch out for ambiguities in the definition of minors. Some software assign the indices of minors differently.

Theorems about determinants (stated without proof):

- A square matrix and its transpose have the same determinant:

$$\mathrm{Det}[\mathbf{A}^T] = \mathrm{Det}[\mathbf{A}].$$

- Nonzero determinants require that each row and each column of the matrix have at least one nonzero element.
- Switching any two rows or any two columns of the matrix changes the sign of its determinant.
- Multiplying all elements of one row or one column by the same number implies multiplying the determinant by that number.
- Nonzero determinants require that no two rows or columns of the matrix are proportional.
- Adding the elements of one row (column) to those of a different row (column) leaves the determinant invariant.
- Even though matrix multiplication is not commutative ($\mathbf{FG} \neq \mathbf{GF}$, in general), both products have the same determinant, which is equal to the product of determinants:

$$\mathrm{Det}[\mathbf{FG}] = \mathrm{Det}[\mathbf{GF}] = \mathrm{Det}[\mathbf{F}]\mathrm{Det}[\mathbf{G}].$$

- The determinant of a matrix vanishes if its rows or columns interpreted as vector components are linearly dependent.

Inverse matrix:

A square matrix $\mathbf{H}$ with nonvanishing determinant has an *inverse matrix*, i.e. a matrix which satisfies the relation $\mathbf{H}^{-1}\mathbf{H} = \mathbf{H}\mathbf{H}^{-1} = \mathbf{I}$:

$$\mathbf{H}^{-1} = \frac{(\mathbf{H}_{ij})^T}{\mathrm{Det}[\mathbf{H}]},$$

where $(\mathbf{H}_{ij})^T$ is the transpose of the cofactor matrix introduced earlier.

The inverse of an inverse matrix is the original matrix: $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$. The inverse of a product of matrices is related to the inverse matrices as follows:

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

Determinant of inverse matrix: $\mathrm{Det}[\mathbf{A}^{-1}] = \dfrac{1}{\mathrm{Det}[\mathbf{A}]}$.

Orthogonality and unitarity:

A real square matrix $\mathbf{A}$ is *orthogonal* if its transpose is equal to its inverse:

$$\mathbf{A}^T = \mathbf{A}^{-1} \quad \Rightarrow \ \mathbf{A}^T\mathbf{A} = \mathbf{I}.$$

Determinant of orthogonal matrix: $\mathrm{Det}[\mathbf{A}] = \pm 1$.

A complex matrix $\mathbf{B}$ is *unitary* if its complex conjugate transpose (Hermitian adjoint) is equal to its inverse:

$$\mathbf{B}^\dagger \doteq \bar{\mathbf{B}}^T = \mathbf{B}^{-1} \quad \Rightarrow \ \mathbf{B}^\dagger\mathbf{B} = \mathbf{I}.$$

Determinant of unitary matrix: $\mathrm{Det}[\mathbf{B}] = e^{\imath\phi}, \quad \phi \in \mathbb{R}$.

If the rows of an $n \times n$ orthogonal matrix $\mathbf{A}$ are thought of as a set of row vectors $\mathbf{a}_i^{(r)}$, then these same vectors are present as column vectors in $\mathbf{A}^T$. Likewise, the set of column vectors $\mathbf{a}_i^{(c)}$ are present as row vectors in $\mathbf{A}^T$.

The matrix products $\mathbf{A}\mathbf{A}^T = \mathbf{I}$ and $\mathbf{A}^T\mathbf{A} = \mathbf{I}$ are then equivalent to the following scalar products, which state the mutual orthonormality of all row vectors and of all column vectors:

$$\mathbf{a}_i^{(r)} \cdot \mathbf{a}_j^{(r)} = \delta_{ij}, \quad \mathbf{a}_i^{(c)} \cdot \mathbf{a}_j^{(c)} = \delta_{ij}.$$

The corresponding statement for unitary matrices involves a complex conjugation in the definition of the scalar product:

$$(\mathbf{b}_i^{(r)})^* \cdot \mathbf{b}_j^{(r)} = \delta_{ij}, \quad (\mathbf{b}_i^{(c)})^* \cdot \mathbf{b}_j^{(c)} = \delta_{ij}.$$

Systems of linear equations:

The system of linear equations,

$$\sum_{j=1}^{n} a_{ij}x_j = r_i \quad : \ i = 1, \ldots, n,$$

can be expressed as a matrix equation,

$$\mathbf{A}\mathbf{X} = \mathbf{R},$$

where $\mathbf{A}$ is the (square) matrix of (given) coefficients a_{ij}. The column vectors $\mathbf{X}$ and $\mathbf{R}$ represent the (unknown) variables x_i and the (given) nonhomogeneities r_i. A unique solution exists if $\mathrm{Det}[\mathbf{A}] \neq 0$:

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{R}.$$

In situations with $\mathrm{Det}[\mathbf{A}] = 0$ no solution or infinitely many solutions may exist, depending on the specifics of $\mathbf{A}$ and $\mathbf{R}$.

Eigenvalues and eigenvectors:

Eigenvalue equations are a frequent occurrence in physics. The prototypical forms are

$$\mathbf{XA} = \lambda\mathbf{X}, \quad \mathbf{AY} = \lambda\mathbf{Y},$$

where $\mathbf{A}$ is an $n \times n$ square matrix, $\mathbf{X}$ is a left (row) eigenvector, and $\mathbf{Y}$ is a right (column) eigenvector.

The left-eigenvector equation is equivalent to the right-eigenvector problem of the transposed matrix:

$$\mathbf{XA} = \lambda\mathbf{X} \quad \Rightarrow \; \mathbf{A}^T\mathbf{X}^T = \lambda\mathbf{X}^T.$$

The eigenvectors $\mathbf{X}_i$, $\mathbf{Y}_i$ and eigenvalues λ_i are, in general, complex. From the products,

$$(\mathbf{X}_i\mathbf{A})\mathbf{Y}_j = \mathbf{X}_i(\mathbf{AY}_j) \quad \Rightarrow \; \lambda_i\mathbf{X}_i\mathbf{Y}_j = \lambda_j\mathbf{X}_i\mathbf{Y}_j,$$

follows that the left and right eigenvectors belonging to different eigenvalues must be orthogonal to each other.

The n eigenvalues λ_i are the roots of the characteristic polynomial,

$$\mathrm{Det}[\mathbf{A} - \lambda\mathbf{I}] = 0.$$

The eigenvalues of orthogonal or unitary matrices have unit norm. The norm of a real (complex) number is its magnitude (modulus).

For symmetric or Hermitian matrices, the left and right eigenvalue problems are equivalent and all eigenvalues are real. Left and right eigenvectors are (pairwise) identical, forming an orthogonal set.

The eigenvectors of an $n \times n$ Hermitian matrix $\mathbf{H}$ form a unitary matrix $\mathbf{U}$. We can write

$$\mathbf{Hu}_i = \lambda_i\mathbf{u}_i, \quad \mathbf{U} = \big(\mathbf{u}_1 \cdots \mathbf{u}_n\big) \quad \Rightarrow \; \mathbf{HU} = \mathbf{U\Lambda},$$

where $\mathbf{\Lambda}$ is a diagonal matrix with elements $\lambda_1, \ldots, \lambda_n$ along the diagonal. Multiplying the last equation with $\mathbf{U}^{-1} = \mathbf{U}^\dagger$ from the left demonstrates the unitary transformation which diagonalizes a Hermitian matrix:

$$\mathbf{U}^{-1}\mathbf{HU} = \mathbf{\Lambda}.$$

A corresponding relation pertains to the diagonalization of a real, symmetric matrix $\mathbf{S}$ by an orthogonal matrix $\mathbf{O}$ with the attribute $\mathbf{O}^{-1} = \mathbf{O}^T$.