Hare
Krishna

Matrices

CHEAT SHEET

RAGHAVENDRA BAZARI

Contents

Matrices

Definition

Rectangular table of elements arranged in rows & columns

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \updownarrow \text{ number of rows } (n)$$

$$\xleftrightarrow{\qquad\qquad}$$

number of columns (m)

Each element of a matrix is often denoted by a variable with two subscripts. For example, a_{21} represents the element at the second row and first column of the matrix.

Dimension of a matrix A are the n rows (horizontal) and the m columns (vertical), represented by $n \times m$.
If $n == m$ then matrix is called *Square Matrix* with dimension n.
If $n! = m$ then matrix is called *Rectangular Matrix*.

Types of Matrices

	Definition/Prerequisite	Example/Properties
Null or Zero Matrix (O_n)	All entries of the matrix are zero. Represented by O_n where n is dimension of square matrix. (It can be rectangle as well but most generally used with Square matrices)	$O_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$
Identity or Unit Matrix (I_n)	A Diagonal matrix with each $d_i = 1$. Represented by I_n where n is dimension of square matrix.	$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Diagonal Matrix (D_n)	A Square matrix with all elements is zero except diagonal. Diagonal elements can be either zero or non-zero. Represented by D_n where n is dimension of square matrix.	$D_3 = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}$
Positive (Semi) Definite Matrix	Any of below condition is fulfilled. $\vec{v}^T A \vec{v} \geq 0 \; for \; all \; \vec{v}$ $\lambda_i \geq 0$ All the principal minors of A are ≥ 0. i.e., All the upper left (or lower right) submatrices A_k have nonnegative determinants.	
Idempotent matrix	$A^n = A$	*Eigen values of idempotent matrices are either 0 or 1*
Diagonalizable	If such decomposition of A exists: $A = V \, \Lambda V^{-1}$ $where \; V \; is \; eigen \; matrix = [\overrightarrow{v_0}, \overrightarrow{v_1}, \dots, \overrightarrow{v_n}]$ Λ is diagonal matrix whose diagonals are eigen values	
Symmetric matrix	$A^T = A$	
Skew Symmetric Matrix	$A^T = -A$	
Orthogonal Matrix	$A^{-1} = A^T$	

Operations

	Result (R)	Prerequisite	$\dim(R)$	$r_{ij} =$	Properties
Rank of Matrix, $\rho(A)$	ρ(A) : The number of linearly independent rows or columns in the matrix. ρ(A) is used to denote the rank of matrix A. (i.e., number of non-zero rows after doing any number of row operations $r_j \to r_j + \alpha r_k$ for any j and k).	NA	NA	For example: $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ $\rho(A) = 1$	1. A matrix whose rank is equal to its dimensions is called a full rank matrix. 2. When the rank of a matrix is smaller than its dimensions, the matrix is called rank-deficient, singular, or multicollinear. 3. Only full rank matrices have an inverse. 4. Only O (Zero) matrix can have zero rank.
Determinant	$\|A\|$	$\# \, of \, cols(A) == \# \, of \, rows(A)$	NA	$if \; A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ $then \; \|A\| = a_{11}a_{22} - a_{12}\,a_{21}$ $if \; A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ $then \; \|A\| = a_{11}(a_{22}a_{33} - a_{23}a_{31}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$	$\|cA\|^n = c^n\|A\|^n for \; any \; constant \; c$ $\|A^T\| = \|A\|$ $\|A^{-1}\| = \frac{1}{\|A\|}$ $\|AB\| = \|A\|\|B\|$ A matrix with determinant value of zero is called *Singular Matrix.* Consequently, a matrix with non-zero determinant value is called *non-singular matrix* or *invertible matrix.*

Trace	$Tr(A) = sum\ of\ diagonal\ elements\ of\ matrix\ A$	A is Square matrix	NA	$\sum_{i=0}^{n-1} a_{ii}$	Tr(A) = Sum of eigen values of a matrix
Addition or Subtract	$A + B$ or $A - B$	$\dim(A) == \dim(B)$	$\dim(A)$ or $\dim(B)$	$a_{ij} + b_{ij}$	$A + O_n = O_n + A = A$ $A + B = B + A$
Multiplication	$A \times B\ or\ AB$	$\#\ of\ cols(A) == \#\ of\ rows(B)$	$(\#\ of\ rows(A)$, $\#\ of\ cols(B))$	$\sum_{k=0}^{m} a_{ik} \times b_{kj}$	$A \times I_n = I_n \times A = A$ $A \times O_n = O_n \times A = O_n$ $A \times (B + C) = AB + AC$ or $(B + C)A = BA + CA$ For a Square matrix A, its self-multiplication n times is represented by A^n.
Equality	$A == B$	$\dim(A) == \dim(B)$ $a_{ij} == b_{ij}$ for all i & j	NA	NA	NA
Transpose	A^T	NA	$(\#\ of\ cols(A),$ $\#\ of\ rows(A))$	a_{ji}	$(A^T)^T = A$ $(A \pm B)^T = A^T \pm B^T$ $(cA)^T = cA^T\ for\ any\ constant\ c$ $(AB)^T = B^T A^T$

Inverse	A^{-1}	A is Square matrix & $\lvert A\rvert$ is not zero	$\dim(A)$	$A^{-1} = \frac{1}{\lvert A\rvert} adj(A)$ Special case: $if\ A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ $then$ $adj(A) = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$ More info on how to compute inverse can be found here	$(A^{-1})^{-1} = A$ $A\,A^{-1} = I_n$ $(A^T)^{-1} = (A^{-1})^T$ $(cA)^{-1} = \frac{1}{c} A^{-1}\ for\ any\ constant\ c$ $(AB)^{-1} = B^{-1}A^{-1}$ $(A^{-1})^n = A^{-n}$
Differentiation	$\frac{\partial A}{\partial B}$	A and B both are column matrices i.e. $\dim(A) = n \times 1$ $\dim(B) = m \times 1$	$n \times m$	$\frac{\partial a_{i0}}{\partial b_{j0}}$	If $n = 1$ then $\frac{\partial A}{\partial B} = \frac{\partial (A^T)}{\partial B}$ Transpose: $\frac{\partial (A^T)}{\partial B} = \left(\frac{\partial A}{\partial B}\right)^T$ Constant Multiplier: $\frac{\partial (CA)}{\partial B} = C\frac{\partial A}{\partial B}$ with dimension $k \times m$, for any constant matrix C with dimension, $k \times n$ Product Rule: $\frac{\partial (DA)}{\partial B} = D\frac{\partial A}{\partial B} + A^T\frac{\partial D^T}{\partial B}$ for any non-constant matrix D with dimension, $1 \times n$

Eigen Value vs Eigen Vector	Eigen Vector: $\vec{v}$ is a vector which even after applying linear transformation(A), at most gets scaled by λ (called eigen value) i.e., $A\vec{v} = \lambda\vec{v}$ or $(A - \lambda I)\vec{v} = 0$ $\vec{v}$ is only eigen vector for A, other transformation might change its direction/magnitude. Hence given a matrix A $(n \times n)$ we can have n eigen vectors, whose values will only affect by at-most scaling them by λ.	A is Square matrix	NA	Solve $\|A - \lambda I\| = 0$ to get n eigenvalues λ_i. Then for each λ_i solve $A\overrightarrow{v_i} = \lambda_1\overrightarrow{v_i}$	1. $tr(A) = \sum \lambda_i$ 2. $\|A\| = \prod \lambda_i$ 3. Rank(A) = number of non-zero eigen values of A. 4. If $\text{Eigenvalue(A)} = \lambda_i$, then • $\text{Eigenvalue}(\text{A}^\text{k}) = \lambda_i^k$ & $\text{Eigenvector}(\text{A}^k) = \text{Eigenvector(A)}$ • $\text{Eigenvalue}(\text{A}^{-1}) = \lambda_i^{-1}$ & $\text{Eigenvector}(\text{A}^{-1}) = \text{Eigenvector(A)}$ $provided\ no\ eigen\ value\ is\ zero$ • $\text{Eigenvalue}(\text{A} + \text{cI}) = \lambda_i + c$ • Eigen Values of upper/lower triangular matrix is its diagonal elements. 5. If a matrix A is symmetric matrix then All eigen vectors are orthogonal, i.e., $\overrightarrow{v_0}\,.\overrightarrow{v_1} = 0$. Eigen matrix, $[\overrightarrow{v_0}, \overrightarrow{v_1}, \dots, \overrightarrow{v_n}]$ of this matrix project the correlation in n orthogonal dimension of most variance: it is eigenvectors and eigenvalues who are behind all the magic explained above, because the eigenvectors of the Covariance matrix are the directions of the axes where there is the most variance (most information) and that we call Principal Components. And eigenvalues are simply the coefficients attached to eigenvectors, which give the amount of variance carried in each Principal Component. In general, to compute principal vectors, first they are normalized (scaled such that length becomes 1) and corresponding eigen values are adjusted accordingly. Sorted in order of eigen values.

Eigen decomposition of Matrix	$A = V\,\Lambda V^{-1}$ $where\ V\ is\ eigen\ matrix = [\overrightarrow{v_0}, \overrightarrow{v_1}, \dots, \overrightarrow{v_n}]$ Λ is diagonal matrix whose diagonals are eigen values	A is Square matrix	NA	NA	NA
Cholesky Decomposition	$A = LL^T$ Where L is lower triangular matrix & L^T is its transpose.	A is Square Symmetric Positive semi definite	NA	NA	All correlation matrices are Cholesky decomposable.

LU Decomposition	$A = LU$ Where L is lower triangular matrix & U is upper triangular matrix.	NA	NA	NA	NA
QR	In linear algebra, a QR decomposition, also known as a QR factorization or QU factorization, is a decomposition of a matrix A into a product $A = QR$ of an orthogonal matrix Q and an upper triangular matrix R.	NA	NA	NA	NA

SVD	Singular value decomposition takes a rectangular matrix (defined as A, where A is a n x p matrix) $A = USV$ Where, $U^T U = I_n$ $V^T V = I_p$ i.e., U and V are orthogonal & S is rectangular diagonal matrix with non-negative real numbers on the diagonal (called singular values)	NA	NA	NA	NA