Primer on Linear Algebra

Linear Algebra

- *Basic Concepts on Vectors and Matrices*

Reading:

- *Many primers (check internet)*

Matrices & Vectors

- Vector $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \cdot \\ \vdots \\ x_n \end{bmatrix} = (x_1 \ldots x_n)^T = [x_i]$
- Dot product $\boldsymbol{x}^T \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i$, outer $\boldsymbol{x}\boldsymbol{y}^T = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & \ddots & \vdots \\ x_n y_1 & \cdots & x_n y_n \end{pmatrix}$
- Matrix $\boldsymbol{A} = \begin{bmatrix} a_{11} & \cdots & a_{1l} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nl} \end{bmatrix} = (\boldsymbol{a_1} \ldots \boldsymbol{a}_l) = \begin{bmatrix} \tilde{\boldsymbol{a}}_1{}^T \\ \cdot \\ \vdots \\ \tilde{\boldsymbol{a}}_n{}^T \end{bmatrix} = [a_{ij}]$

Matrix Multiplication

- Matrix multiplication $\boldsymbol{A} = [a_{\boldsymbol{ij}}] \in R^{nxl}$ $\mathbf{B} = [b_{\boldsymbol{ij}}] \in R^{lxm}$

$$\boldsymbol{AB} = \left[\sum_{i=1}^{M} a_{ik} b_{kj}\right] = \begin{bmatrix} \tilde{\boldsymbol{a}}_1{}^T \\ \vdots \\ \tilde{\boldsymbol{a}}_n{}^T \end{bmatrix} [\boldsymbol{b}_1 \quad \cdots \quad \boldsymbol{b}_m] = [\tilde{\boldsymbol{a}}_i{}^T \boldsymbol{b}_j]$$

$$= [\boldsymbol{Ab}_1 \quad \cdots \quad \boldsymbol{Ab}_m] = \begin{bmatrix} \tilde{\boldsymbol{a}}_1{}^T \boldsymbol{B} \\ \vdots \\ \tilde{\boldsymbol{a}}_n{}^T \boldsymbol{B} \end{bmatrix}$$

Matrix Multiplication

- Matrix/vector multiplication $\boldsymbol{Ab} = \begin{bmatrix} \tilde{\boldsymbol{a}}_1{}^T \\ \vdots \\ \tilde{\boldsymbol{a}}_n{}^T \end{bmatrix} \boldsymbol{b} = \begin{bmatrix} \tilde{\boldsymbol{a}}_1{}^T \boldsymbol{b} \\ \vdots \\ \tilde{\boldsymbol{a}}_n{}^T \boldsymbol{b} \end{bmatrix}$ $=[\tilde{\boldsymbol{a}}_j{}^T \boldsymbol{b}]$

- Identity matrix $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ 3x3 Identity matrix

Matrix Transpose

- Matrix transpose

$$\boldsymbol{A}^T = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1l} & \cdots & a_{nl} \end{bmatrix} = (\tilde{\boldsymbol{a}}_1 \ldots \tilde{\boldsymbol{a}}_n) = \begin{bmatrix} {\boldsymbol{a}_1}^T \\ . \\ \vdots \\ {\boldsymbol{a}_l}^T \end{bmatrix}$$

- Property $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$

- Symmetric Matrix

$$\boldsymbol{A} = \boldsymbol{A}^T$$

Example

$$\begin{bmatrix} 4 & 5 & -3 \\ 5 & 7 & 2 \\ -3 & 2 & 10 \end{bmatrix}$$

Trace

$$tr(\boldsymbol{A}) = \sum_{i}^{n} a_{ii}$$

- properties $$tr(\boldsymbol{A}^T) = tr(\boldsymbol{A})$$

$$tr(\boldsymbol{A} + \boldsymbol{B}) = tr(\boldsymbol{A}) + tr(\boldsymbol{B})$$

$$tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$$

- in general $$tr(\boldsymbol{AB}) \neq tr(\boldsymbol{B})tr(\boldsymbol{A})$$

Determinants

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$|\boldsymbol{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

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

$$|\boldsymbol{A}| = \sum_{j=1}^{n} (-1)^{j+k} a_{jk} |\boldsymbol{A}_{jk}|$$

Determinants

- Properties

$$|\boldsymbol{AB}| = |\boldsymbol{A}||\boldsymbol{B}| \qquad |\boldsymbol{A}+\boldsymbol{B}| \neq |\boldsymbol{A}| + |\boldsymbol{B}|$$

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & 0 & . & 0 \\ 0 & a_{22} & . & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & . & a_{nn} \end{bmatrix} \quad \text{then} \quad |\boldsymbol{A}| = \prod_{i=1}^{n} a_{ii}$$

Matrix Inverse

- Matrix inverse $\boldsymbol{A}^{-1}$ $\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\,\boldsymbol{A}^{-1} = \boldsymbol{I}$

  $\boldsymbol{A}^{-1}$ exists iff $|\boldsymbol{A}| \neq 0$

- Solving systems $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \;\Rightarrow\; \boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$

- Properties $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$ $\boldsymbol{A}^{-1^T} = \boldsymbol{A}^{T^{-1}}$

$$|\boldsymbol{A}^{-1}| = \frac{1}{|\boldsymbol{A}|}$$

Matrix Inverse

- Inverse is only for square matrices. What happens for non-square (pseudo-inverse)?

- Solving the system $\boldsymbol{A} \in R^{nxl}$

$$\boldsymbol{Ax} = \boldsymbol{b} \underset{\boldsymbol{A}^T}{\Rightarrow} \boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$$

$$\xrightarrow[(\boldsymbol{A}^T\boldsymbol{A})^{-1}]{} \boldsymbol{x} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$$

- Pseudo-inverse $\boldsymbol{A}^+ = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T \quad \boldsymbol{A}^+\boldsymbol{A} = \boldsymbol{I}$

Rank of matrix (Definition 1)

Equal to the dimension of the largest square sub-matrix of a matrix that has a non-zero determinant.

$$\begin{bmatrix} 4 & 5 & 2 & 14 \\ 3 & 9 & 6 & 21 \\ 8 & 10 & 7 & 28 \\ 1 & 2 & 9 & 5 \end{bmatrix}$$ has rank 3

$$|\boldsymbol{A}| = 0 \quad \text{but} \quad \left|\begin{bmatrix} 4 & 5 & 2 \\ 3 & 9 & 6 \\ 8 & 10 & 7 \end{bmatrix}\right| = 63$$

Rank of matrix (Definition 2)

- A set of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \dots, \boldsymbol{a}_l$ are linearly independent then $c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + c_3\boldsymbol{a}_3 + \cdots + c_l\boldsymbol{a}_l = 0$

  iff $c_1, c_2, c_3, \dots, c_n = 0$

<u>Alternative definition of rank</u>: the maximum number of linearly independent columns (or rows) of $\boldsymbol{A}$

Example:
$$1\begin{bmatrix}4\\3\\8\\1\end{bmatrix} + 2\begin{bmatrix}5\\9\\10\\2\end{bmatrix} + 0\begin{bmatrix}2\\6\\7\\9\end{bmatrix} - 1\begin{bmatrix}14\\21\\28\\5\end{bmatrix} = 0$$

Therefore, rank is not 4 !

Properties

- If $\boldsymbol{A}$ in $nxn, rank(\boldsymbol{A}) = n$ iff $\boldsymbol{A}$ is nonsingular (i.e., invertible)
- If $\boldsymbol{A}$ in $nxn, rank(\boldsymbol{A}) = n$ iff $|\boldsymbol{A}| \neq 0$ (full rank).
- If $\boldsymbol{A}$ in $nxn, rank(\boldsymbol{A}) < n$ iff $\boldsymbol{A}$ is singular.

Orthogonal/Orthonormal matrices

$$\boldsymbol{A} = \begin{bmatrix} \widetilde{\boldsymbol{a}}_1{}^T \\ \vdots \\ \widetilde{\boldsymbol{a}}_n{}^T \end{bmatrix} \qquad \widetilde{\boldsymbol{a}}_j{}^T \widetilde{\boldsymbol{a}}_k = 0 \text{ for every } j \neq k \qquad \boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}} = \mathbf{L}$$

$$\boldsymbol{A} = [\boldsymbol{a}_1 \ldots \boldsymbol{a}_n] \qquad \boldsymbol{a}_{\boldsymbol{j}}{}^T \boldsymbol{a}_{\boldsymbol{k}} = 0 \text{ for every } j \neq k \qquad \boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}} = \mathbf{L}$$

$$\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}} = \mathbf{I} \qquad \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A} = \mathbf{I}$$

Eigenvalues & Eigenvectors

- Vector ***x*** is an eigenvector of matrix ***A*** and λ is an eigenvalue of ***A*** if:

$$\boldsymbol{Ax} = \lambda \boldsymbol{x} \qquad \boldsymbol{x} \neq \boldsymbol{0}$$

- **Interpretation:** the linear transformation implied by ***A*** cannot change the direction of the eigenvectors ***x***, only their magnitude.

Eigenvalues & Eigenvectors

- To find the eigenvalues λ of a matrix *A*, find the roots of the *characteristic polynomial*:

$$|\boldsymbol{A} - \lambda \boldsymbol{I}| = 0$$

Example: $\boldsymbol{A} = \begin{bmatrix} 5 & -2 \\ 6 & -2 \end{bmatrix}$

$$\left| \begin{bmatrix} 5-\lambda & -2 \\ 6 & -2-\lambda \end{bmatrix} \right| = 0 \Rightarrow \lambda^2 - 3\lambda + 2 = 0 \Rightarrow \lambda_1 = 1, \lambda_2 = 2 \Rightarrow$$

Eigenvalues & Eigenvectors

Find the eigenvectors by solving the two homogenous systems:

$$(\boldsymbol{A} - \lambda_1 \boldsymbol{I})\boldsymbol{v}_{\mathbf{1}} = 0 \Rightarrow \boldsymbol{v}_{\mathbf{1}} = k \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$$

$$(\boldsymbol{A} - \lambda_2 \boldsymbol{I})\boldsymbol{v}_{\mathbf{1}} = 0 \Rightarrow \boldsymbol{v}_{\mathbf{2}} = k \begin{bmatrix} 1/3 \\ 1 \end{bmatrix}$$

Eigenvalues & Eigenvectors

- Eigenvalues and eigenvectors are only defined for square matrices (i.e., $n = l$)
- Eigenvectors are not unique: If $\boldsymbol{v}$ is an eigenvector, so is $k\boldsymbol{v}$ (solution of a homogenous system)
- Suppose $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of $\boldsymbol{A}$, then:

$$\sum_{i=1}^{n} \lambda_i = tr(\boldsymbol{A}) \quad \prod_{i=1}^{n} \lambda_i = |\boldsymbol{A}| \quad \lambda_i = 0 \Rightarrow \boldsymbol{A} \text{ is singular}$$

Diagonalisation

- Given matrix $\boldsymbol{A}$, find $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is diagonal (i.e., $\boldsymbol{P}$ diagonalises $\boldsymbol{A}$)
- Take $\boldsymbol{P} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n]$, where $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ are the eigenvectors of $\boldsymbol{A}$

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

Diagonalisation

Are all $n \; x \; n$ matrices diagonalizable with $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$?

- Only if $\boldsymbol{P}^{-1}$ exists (i.e., $\boldsymbol{A}$ must have n linearly independent eigenvectors, that is, $rank(\boldsymbol{A}) = n$)

- If $\boldsymbol{A}$ has n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then the corresponding eigen-vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ form a basis:

  (1) linearly independent

  (2) span R^n

Diagonalisation

Symmetric Matrices

- The eigenvalues of symmetric matrices are all real.

- The eigenvectors corresponding to distinct eigenvalues are orthogonal. $\boldsymbol{P}^{-1} = \boldsymbol{P}^{T}$

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{T} = \sum_{i=1}^{n} \lambda_{i} \boldsymbol{v}_{i} (\boldsymbol{v}_{i})^{T}$$

Properties in the white board