MA1522 Linear Algebra for Computing

AY 24/25 Sem 1 — github/omgeta

1. Vector Spaces

A vector space V is a nonempty set of vectors with the following properties for all vectors $\vec{u}, \vec{v}, \vec{w} \in V$ and for all scalars c and d:

i. $\vec{u} + \vec{v} \in V$
ii. $c\vec{u} \in V$
iii. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
iv. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
v. $\vec{0} \in V$ such that $\vec{u} + \vec{0} = \vec{u}$
vi. $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
vii. $(c + d)\vec{u} = c\vec{u} + d\vec{u}$
viii. $c(d\vec{u}) = (cd)\vec{u}$
ix. $1\vec{u} = \vec{u}$

A subspace W of a vector space V is a subset with the following properties for all vectors $\vec{u}, \vec{v} \in W$ and for all scalars c:

i. $0 \in W$ (Includes zero vector)
ii. $\vec{u} + \vec{v} \in W$ (Closure over addition)
iii. $c\vec{u} \in W$ (Closure over multiplication)

A linear map from V to W is a function $T : V \to W$ with the following properties for linear maps R, S, T and for all scalars c, d for which the following are defined:

i. $S + T = T + S$ (Commutative)
ii. $(R + S) + T = R + (S + T)$ (Associative)
iii. $T + \mathbf{0} = T$ (Additive Identity)
iv. $T + (-T) = \mathbf{0}$ (Additive Inverse)
v. $c(dT) = (cd)T$ (Associative)
vi. $c(S + T) = cS + cT$ (Distributive)
vii. $(c + d)T = cT + dT$ (Scalar Addition)
viii. $R(ST) = (RS)T$ (Associative)
ix. $R(S + T) = RS + RT$ (Distributive)
x. $(S + T)R = SR + TR$ (Distributive)
xi. $c(ST) = (cS)T = S(cT)$ (Associative)
xii. $TI = IT = T$ (Identity)

2. Vectors

For some vector $\vec{v} \in \mathbb{R}^n$, where $v_1, \ldots, v_n \in \mathbb{R}$:

$$\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

Linear Combinations

For vectors $\vec{v}_1, \ldots, \vec{v}_p \in V$ and scalars $c_1, \ldots, c_p$, the vector $\vec{y}$ given by:

$$\vec{y} = c_1\vec{v}_1 + \ldots + c_p\vec{v}_p$$

is a linear combination of $\vec{v}_1, \ldots, \vec{v}_p$ with weights $c_1, \ldots, c_p$

Linear Span

For a set of vectors $S = \{\vec{v}_1, \ldots, \vec{v}_p\} \subseteq V$, $\text{Span}(S) \subseteq V$ denotes the set of all linear combinations of $\vec{v}_1, \ldots, \vec{v}_p$ and is given by:

$$\text{Span}(S) = \{\sum_{i=1}^{p} c_i\vec{v}_i | \vec{v}_i \in S, c_i \in K\}$$

Linear Dependence

For a set of non-zero vectors $S = \{\vec{v}_1, \ldots, \vec{v}_p\} \subseteq V$, S is linearly dependent if and only if some vector $\vec{v}_i$ is a linear combination of the others. Any set $\{\vec{v}_1, \ldots, \vec{v}_p\} \in \mathbb{R}^n$ is linearly dependent if $p > n$.

A linearly independent set of vectors forms a matrix with a pivot position in every column.

Basis

For a set of non-zero vectors $S = \{\vec{v}_1, \ldots, \vec{v}_p\} \subseteq V$, S is a basis for $W \subseteq V$ if:

i. S is a linearly independent set, and
ii. $\text{Span}(S) = W$

3. Matrices

For some matrix $A \in \mathbb{R}^{m \times n}$, where $a_{11}, \ldots, a_{mn} \in \mathbb{R}$:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

For some matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$:

$$\begin{aligned} AB &= A \begin{bmatrix} \vec{b_1} & \cdots & \vec{b_p} \end{bmatrix} \\ &= \begin{bmatrix} A\vec{b_1} & \cdots & A\vec{b_p} \end{bmatrix} \\ (AB)_{ij} &= \sum_{k=1}^{n} A_{ik}B_{kj} \\ \text{row}_i(AB) &= \text{row}_i(A) \cdot B \end{aligned}$$

Additional properties:

i. $AB \neq BA$ in general (Not commutative)
ii. $\exists A, B \neq 0,\ AB = 0$ (Zero divisor)
iii. $A0 = 0A = 0$ (Zero matrix)
iv. A has zero row $\implies$ AB has zero row
v. B has zero column $\implies$ AB has zero column

Transpose

For some matrix $A \in \mathbb{R}^{m \times n}$, the transpose $A^T \in \mathbb{R}^{n \times m}$ is given by:

$$A^T_{ij} = A_j i$$
$$\text{row}_i(A^T) = \text{column}_i(A)$$

Properties:

i. $(A^T)^T = A$
ii. $(cA)^T = cA^T$
iii. $(A + B)^T = A^T + B^T$
iv. $(AB)^T = B^T A^T$
v. $A^T A = 0 \iff A = 0$

Systems of Linear Equations

Systems of linear equations of the form:

$$a_{11}x_1 + \dots + a_{1n}x_n = b_1$$
$$\vdots$$
$$a_{m1}x_1 + \dots + a_{mn}x_n = b_n$$

can be expressed in the matrix-vector form, $A\vec{x} = \vec{b}$, where $A \in \mathbb{R}^{m\times n}$ is the coefficient matrix and $\vec{x} \in \mathbb{R}^n$ is the solution vector:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

By finding the RREF of the augmented matrix, we can find the solution set for the original system.

Row Equivalence

Two matrices $A, B \in \mathbb{R}^{m\times n}$ are row equivalent if one can be changed to the other by a sequence of elementary row operations, that is:

$$B = E_k E_{k-1} \dots E_1 A \iff A \sim B$$

Elementary row operations:

i. Add a multiple of a row to another row $(R_n + cR_m)$
ii. Scale a row by a nonzero constant (cR_n)
iii. Interchange two rows $(R_n \leftrightarrow R_m)$

Row Echelon Forms

A matrix is in row echelon form (REF) if:

i. All nonzero rows are above all zero rows
ii. Each pivot is to the right of the pivot of the row above it
iii. All entries below a pivot are zeros

A matrix is in reduced row echelon form (RREF) if:

i. All pivots are 1
ii. All other entries in the pivot column are 0

Inverse

For some square matrix $A \in \mathbb{R}^{n\times n}$, A is invertible/nonsingular if there exists the inverse $A^{-1} \in \mathbb{R}^{n\times n}$ such that:

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

For matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Using row reduction, we can solve for A^{-1} by:

$$[A \mid I] \xrightarrow{RREF} [I \mid A^{-1}]$$

Properties:

i. $(A^{-1})^{-1} = A$
ii. $(cA)^{-1} = c^{-1}A^{-1}$
iii. $(A^T)^{-1} = (A^{-1})^T$
iv. $(AB)^{-1} = B^{-1}A^{-1}$
v. $AB = AC \implies B = C$
vi. $BA = CA \implies B = C$

Invertible Matrix Theorem

Let A be a $n \times n$ matrix. The following statements are equivalent:

i. A is invertible
ii. A^T is invertible
iii. A has a left inverse, C, such that $CA = I$
iv. A has a right inverse, D, such that $AD = I$
v. RREF of A is I
vi. Columns of A form a basis for $\mathbb{R}^n$
vii. $A\vec{x} = \vec{0}$ has only the trivial solution
viii. $A\vec{x} = \vec{b}$ has a unique solution, $\forall \vec{b} \in \mathbb{R}^n$
ix. $\mathrm{Nul}(A) = \{\mathbf{0}\} \iff \mathrm{nullity}(A) = 0$
x. $\mathrm{Col}(A) = \mathbb{R}^n \iff \mathrm{rank}(A) = n$
xi. $\det(A) \neq 0$
xii. 0 is not an eigenvalue

Determinant

For some square matrix $A \in \mathbb{R}^{n\times n}$, the determinant $\det(A)$ or $|A|$ is given by:

$$\det(A) = \sum_{j=1 \text{ or } i=1}^{n} a_{ij}A_{ij}$$

where A_{ij} is the (i, j) cofactor of A given by:

$$A_{ij} = (-1)^{i+j}\det(M_{ij})$$

where $M_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ is the (i, j) matrix minor of A obtained by deleting the ith row and jth column of A.

For matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $\det(A) = ad - bc$

Properties:

i. $\det(A^T) = \det(A)$
ii. $\det(AB) = \det(A)\det(B)$
iii. $\det(A^{-1}) = \frac{1}{\det(A)}$
iv. $\det(cA) = c^n \det(A)$
v. If A is triangular, $\det(A)$ is the product of the entries on the main diagonal of A

For the elementary matrix $E \in \mathbb{R}^{m\times n}$, $\det(A)$ is given by the type of elementary row operation:

i. $R_n + cR_m \implies \det(E) = 1$
ii. $cR_n \implies \det(E) = c$
iii. $R_n \leftrightarrow R_m \implies \det(E) = -1$

Cramer's Rule

Let an invertible matrix A and any $b \in \mathbb{R}^n$, the unique solution of $A\vec{x} = \vec{b}$ has it's entries given by:

$$x_i = \frac{\det(A_i(\vec{b}))}{\det(A)}$$

where $A_i(\vec{b})$ is the matrix obtained by replacing $\mathrm{column}_i(A)$ with $\vec{b}$.

Adjoint

For some square matrix $A \in \mathbb{R}^{n\times n}$, the adjoint $\text{adj}(A)$ is given by:

$$\text{adj}(A) = (A_{ij})^T = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n_1} \\ A_{12} & A_{22} & \cdots & A_{n_2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

Properties:

i. $A \cdot \text{adj}(A) = \det(A) \cdot I$
ii. A is singular $\iff$ $\text{adj}(A)$ is singular
iii. $\text{adj}(AB) = \text{adj}(B)\,\text{adj}(A)$
iv. $\det(\text{adj}(A)) = \det(A)^{-1}$
v. $\text{adj}(cA) = c^{n-1}\,\text{adj}(A)$
vi. $\text{adj}(A^{-1}) = \text{adj}(A)^{-1}$
vii. $\text{adj}(\text{adj}(A)) = (\det(A))^{n-2}A$

Change of Basis

For bases $\mathcal{B} = \{\vec{b_1}, \ldots, \vec{b_n}\}$ and $\mathcal{C} = \{\vec{c_1}, \ldots, \vec{c_n}\}$ of a vector space $\mathbb{R}^n$, there is a unique change of basis matrix $P_{\mathcal{C}\leftarrow\mathcal{B}} \in \mathbb{R}^{n\times n}$ such that:

$$\begin{aligned} \left[\vec{x}\right]_{\mathcal{C}} &= P_{\mathcal{C}\leftarrow\mathcal{B}} \left[\vec{x}\right]_{\mathcal{B}} \\ (P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} \left[\vec{x}\right]_{\mathcal{C}} &= \left[\vec{x}\right]_{\mathcal{B}} \\ \implies (P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} &= P_{\mathcal{B}\leftarrow\mathcal{C}} \end{aligned}$$

where $\left[\vec{x}\right]_{\mathcal{B}}$ and $\left[\vec{x}\right]_{\mathcal{C}}$ are the vector $\vec{x}$ represented in the coordinate system used by the bases $\mathcal{B}$ and mcC respectively.

Using row reduction, we can solve for $P_{C\leftarrow B}$ by:

$$\begin{bmatrix} \vec{c_1} & \cdots & \vec{c_n} & | & \vec{b_1} & \cdots & \vec{b_n} \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} I & | & P_{C\leftarrow B} \end{bmatrix}$$

When converting from a basis $\mathcal{B}$ to the standard basis $\mathcal{E} = \{\vec{e_1}, \ldots, \vec{e_n}\}$, change of basis matrix $P_{\mathcal{B}}$ is given by $P_{\mathcal{B}} = \begin{bmatrix} \vec{b_1} \cdots \vec{b_n} \end{bmatrix}$ such that:

$$\begin{aligned} \vec{x} &= P_{\mathcal{B}} \left[\vec{x}\right]_{\mathcal{B}} \\ (P_{\mathcal{B}})^{-1}\vec{x} &= \left[\vec{x}\right]_{\mathcal{B}} \end{aligned}$$

4. Subspaces

Null Space

For any matrix $A \in \mathbb{R}^{m\times n}$, the null space $\text{Nul}(A) \in \mathbb{R}^n$ is the solution space to the homogenous equation $A\vec{x} = \vec{0}$ given by:

$$\text{Nul}(A) = \{\vec{x} \in \mathbb{R}^n | A\vec{x} = \vec{0}\}$$

Column Space

For any matrix $A = \begin{bmatrix} \vec{a_1} & \cdots & \vec{a_n} \end{bmatrix} \in \mathbb{R}^{m\times n}$, the column space $\text{Col}(A) \in \mathbb{R}^m$ is the set of all linear combinations of the columns of A given by:

$$\begin{aligned} \text{Col}(A) &= \text{Span}(\{\vec{a_1}, \ldots, \vec{a_n}\}) \\ &= \{\vec{b} \in \mathbb{R}^m : A\vec{x} = \vec{b} \text{ for some } \vec{x} \in \mathbb{R}^n\} \end{aligned}$$

Row Space

For any matrix $A \in \mathbb{R}^{m\times n}$, the row space $\text{Row}(A) \in \mathbb{R}^n$ is the set of all linear combinations of the rows of A given by:

$$\text{Row}(A) = \text{Col}(A^T)$$

Dimension

If a vector space V is spanned by a finite set, V is finite-dimensional and the dimension $\dim(V)$ is the number of vectors in any basis for V.

$$\begin{aligned} \text{rank}(A) &= \dim(\text{Col}(A)) = \dim(\text{Row}(A)) \\ &= \text{number of pivot columns} \\ &= \text{number of pivot rows} \\ \text{nullity}(A) &= \dim(\text{Nul}(A)) \\ &= \text{number of free variables} \end{aligned}$$

Rank-Nullity Theorem

Rank and nullity of any matrix $A \in \mathbb{R}^{m\times n}$ satisfy the equation:

$$\text{rank}(A) + \text{nullity}(A) = \text{number of columns in } A$$

5. Eigenvectors

For some square matrix $A \in \mathbb{R}^{n\times n}$, the nonzero eigenvector $\vec{x}$ and its corresponding eigenvalue λ satisfy the equation:

$$A\vec{x} = \lambda\vec{x}$$

For some eigenvalue λ, the corresponding eigenvectors are found as the nontrivial solutions to $(A - \lambda I) = 0$ as elements of the corresponding eigenspace of A.

A scalar λ is an eigenvalue if and only if λ satisfies the characteristic equation:

$$\det(A - \lambda I) = 0$$

6. Dot Product

Dot product of two vectors, $\vec{u}, \vec{v} \in \mathbb{R}^n$ is given by:

$$\begin{aligned} \vec{u} \cdot \vec{v} = \vec{u}^T\vec{v} &= \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\ &= u_1v_1 + \cdots + u_nv_n \\ &= \|\vec{u}\|\|\vec{v}\| \cos\theta \end{aligned}$$

where θ is the angle between $\vec{u}$ and $\vec{v}$.

Length of $\vec{v}$ is given by:

$$\begin{aligned} \|v\| &= \sqrt{\vec{v}\cdot\vec{v}} = \sqrt{v_1^2 + \cdots + v_n^2} \\ \|v\|^2 &= \vec{v}\cdot\vec{v} \end{aligned}$$

Distance between $\vec{u}$ and $\vec{v}$ is given by:

$$\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$$

Properties:

i. $\vec{u}\cdot\vec{v} = \vec{v}\cdot\vec{u}$
ii. $(\vec{u} + \vec{v})\cdot\vec{w} = \vec{u}\cdot\vec{w} + \vec{v}\cdot\vec{w}$
iii. $(c\vec{u})\cdot\vec{v} = c(\vec{u}\cdot\vec{v}) = \vec{u}\cdot(c\vec{v})$
iv. $\vec{u}\cdot\vec{u} \geq 0$, and $\vec{u}\cdot\vec{u} = 0 \iff \vec{u} = \vec{0}$

7. Orthogonality

Let $\vec{u}, \vec{v} \in \mathbb{R}^n$ be orthogonal vectors. The following statements are equivalent:

i. $\vec{u} \cdot \vec{v} = 0$
ii. $\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$
iii. $\text{dist}(\vec{u}, \vec{v}) = \text{dist}(\vec{u}, -\vec{v})$

A set of vectors is orthogonal if all vectors are mutually orthogonal.

A set of vectors is orthonormal if it is orthogonal and every vector is a unit vector.

Orthogonal Basis

An orthogonal set of nonzero vectors is linearly independent and a basis for the subspace it spans.

Any orthonormal set is automatically an orthonormal basis for the subspace it spans.

For any orthogonal basis $S = \{\vec{u_1}, \ldots, \vec{u_n}\}$ of subspace V:

$$\vec{v} = \frac{\vec{v} \cdot \vec{u_1}}{\|u_1\|^2}\vec{u_1} + \cdots + \frac{\vec{v} \cdot \vec{u_n}}{\|u_n\|^2}\vec{u_n}$$
$$[\vec{v}]_S = \begin{bmatrix} \vec{v} \cdot \vec{u_1} / \|u_1\|^2 \\ \vdots \\ \vec{v} \cdot \vec{u_n} / \|u_n\|^2 \end{bmatrix}$$

Orthogonal Matrix

For some square matrix $A \in \mathbb{R}^{n \times n}$, A is orthogonal if it has orthonormal columns (and equivalently rows) or:

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

Properties:

i. $\|U\vec{x}\| = \|\vec{x}\|$
ii. $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$
iii. $(U\vec{x}) \cdot (U\vec{y}) = 0 \iff \vec{x} \cdot \vec{y} = 0$

Projection

For any two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$, the projection of $\vec{y}$ onto $\vec{x}$, $\text{proj}_{\vec{x}}\, \vec{y}$, is given by:

$$\begin{aligned} \text{proj}_{\vec{x}}\, \vec{y} &= (\vec{y} \cdot \hat{x})\hat{x} \\ &= \frac{\vec{y} \cdot \vec{x}}{\|\vec{x}^2\|}\vec{x} \end{aligned}$$

For any vector $\vec{y} \in \mathbb{R}^n$ and subspace $W \subseteq \mathbb{R}^n$ with orthogonal basis $\{\vec{v_1}, \ldots, \vec{v_p}\}$, the projection of $\vec{y}$ onto W, $\text{proj}_w\, \vec{y}$ or $\hat{y}$,is given by:

$$\text{proj}_w\, \vec{y} = \hat{y} = \frac{\vec{y} \cdot \vec{v_1}}{\|\vec{v_1}\|^2}\vec{v_1} + \cdots + \frac{\vec{y} \cdot \vec{v_p}}{\|\vec{v_p}\|^2}\vec{v_p}$$

Gram-Schmidt Process

Let $X = \{x_1, \ldots, x_p\}$ be a basis for a nonzero subspace $W \subseteq \mathbb{R}^n$.

$$\begin{aligned} v_1 &= x_1 \\ v_2 &= x_2 - \frac{\vec{x_2} \cdot \vec{v_1}}{\|v_1\|^2}v_1 \\ &\vdots \\ v_p &= x_p - \frac{\vec{x_p} \cdot \vec{v_1}}{\|v_1\|^2}v_1 - \frac{\vec{x_p} \cdot \vec{v_2}}{\|v_2\|^2}v_2 - \cdots - \frac{\vec{x_p} \cdot \vec{v_{p-1}}}{\|v_{p-1}\|^2}v_{p-1} \end{aligned}$$

Then $X' = \{v_1, \ldots, v_n\}$ is an orthogonal basis for W

Least Squares Approximation

For any matrix equation $A\vec{x} = \vec{b}, A \in \mathbb{R}^{m \times n}\vec{b} \in \mathbb{R}^m$, the vector $\hat{x} \in \mathbb{R}^n$ is the least-squares solution such that:

$$\|\vec{b} - A\hat{x}\| \leq \|\vec{b} - A\vec{x}\|, \forall \vec{x} \in \mathbb{R}^n$$

with the least squares solution $\hat{x}$ given by the solution set of:

$$A^T A\vec{x} = A^T\vec{b}$$

8. Factorisations

LU Factorisation

Suppose matrix $A \in \mathbb{R}^{m \times n}$ can be reduced to row echelon form U without row interchanges, then A can be factorised as:

$$\begin{aligned} A &= LU \\ &= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ * & * & \cdots & 1 \end{bmatrix} \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & * \end{bmatrix} \end{aligned}$$

where upper triangular matrix $U \in \mathbb{R}^{m \times n}$ is given by:

$$\begin{aligned} U &= REF(A), \text{ without row interchanges} \\ &= E_k \ldots E_1 A \end{aligned}$$

and unit lower triangular matrix $L \in \mathbb{R}^{m \times m}$ is constructed from unit pivot columns of A (and additional columns of I if there are insufficient pivot columns) such that:

$$E_k \ldots E_1 L = I$$

QR Factorisation

Suppose matrix $A \in \mathbb{R}^{m \times n}$ has linearly independent columns (full column rank), then A can be factorised as:

$$A = QR$$

where orthogonal matrix $Q \in \mathbb{R}^{m \times n}$ is constructed from orthonormal basis vectors for $\text{Col}(A)$ given by:

$$\text{Orthonormal Basis } A' = \text{Gram-Schmidt on } A$$
$$Q = \begin{bmatrix} \vec{a'_1} \cdots \vec{a'_n} \end{bmatrix}$$

and invertible upper triangular matrix $R = P_{A \to Q} \in \mathbb{R}^{n \times n}$ with positive diagonals given by:

$$R = Q^T A$$

ensuring R has positive diagonals by multiplying columns of Q by -1 as needed.

Diagonalisation

Suppose matrix $A \in \mathbb{R}^{n \times n}$ has n linearly independent eigenvectors $\{\vec{v_1}, \ldots, \vec{v_n}\}$, then A can be diagonalised as:

$$A = PDP^{-1}$$

where invertible change of basis matrix $P \in \mathbb{R}^{n \times n}$ is constructed from the linearly independent eigenvectors of A such that:

$$P = \begin{bmatrix} \vec{v_1} \cdots \vec{v_n} \end{bmatrix}$$

and diagonal matrix $D \in \mathbb{R}^{n \times n}$ is constructed from the corresponding eigenvalues of the eigenvectors chosen for the columns of P in the same order:

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Orthogonal Diagonalisation

Suppose matrix $A \in \mathbb{R}^{n \times n}$ is symmetric, then A is orthogonally diagonalisable as:

$$A = PDP^T$$

where invertible change of basis matrix $P \in \mathbb{R}^{n \times n}$ is also orthogonal.

A is symmetric $\iff$ A is orthogonally diagonalisable.

Spectral Theorem

A symmetric matrix $A \in \mathbb{R}^{n \times n}$ has the following properties:

i. A has n real eigenvalues, counting multiplicities
ii. Dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation
iii. Eigenspaces are mutually orthogonal, such that eigenvectors corresponding to different eigenvalues are orthogonal
iv. A is orthogonally diagonalisable

Singular Value Decomposition (SVD)

Any matrix $A \in \mathbb{R}^{m \times n}$ with rank r can be decomposed as:

$$A = U \Sigma V^T$$

where "diagonal" matrix $\Sigma \in \mathbb{R}^{m \times n}$ is constructed with the decreasing r singular values $\sigma = \sqrt{\lambda}$ of A, for eigenvalues σ of the symmetric matrix $A^T A$, such that:

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

$$D = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}$$

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$$

and orthogonal matrix $V \in \mathbb{R}^{n \times n}$ is constructed with the corresponding right singular vectors of A, given by the corresponding unit eigenvectors $\{\vec{v_1}, \ldots, \vec{v_n}\}$ of $A^T A$, such that:

$$V = \begin{bmatrix} \vec{v_1} \cdots \vec{v_n} \end{bmatrix}$$

and orthogonal matrix $U \in \mathbb{R}^{m \times m}$ is constructed with the corresponding left singular vectors of A, such that:

$$\vec{u_i} = \frac{1}{\sigma_i} A \vec{v_i}$$

$$U = [\vec{u_1} \cdots \vec{u_m}]$$

During construction of U and V, if there are insufficient singular vectors to form an orthogonal matrix, additional orthonormal vectors can be formed from the Gram-Schmidt process.