MA1522 Notes

compiled by:
m. zaidan
qianbo dong

for:
Academic Year 2023-2024, Semester 2

Contents

1 Linear Systems **4**
1.1 Elementary Row Operations 4

2 Matrices **5**
2.1 Matrix multiplication 5
2.2 Representation of linear system 5
2.3 Inverses 6
2.3.1 Properties of invertible matrices 6
2.3.2 Finding the inverse 7
2.4 LU Decomposition 7
2.5 Partial pivoting 8
2.6 Elementary matrices 9
2.7 Determinant 9

3 Vector Spaces **11**
3.1 Linear span 11
3.2 Solution space 11
3.3 Linear independence 11
3.4 Basis 12
3.5 Transition matrices 12
3.6 Row & column space 14

3.7 Finding a basis/extending basis 14
3.8 Consistency 16
3.9 Rank 16
3.10 Nullspace 17
3.10.1 Dimension theorem 17

4 Orthogonality **18**
4.1 Orthogonality 18
4.2 Projection 21
4.3 Gram-Schmidt Process 23
4.4 Least Squares 24
4.5 QR Decomposition 25
4.6 Orthogonal Matrices 28

5 Eigenvalues, eigenvectors & eigenspaces **30**
5.1 Eigenvalues & eigenvectors 30
5.2 Characteristic equation 30
5.3 Eigenspace 31

6 Diagonalization **33**
6.1 Diagonalizable matrices 33
6.2 Algebraic and geometric multiplicity 33
6.3 Criterion of diagonalizability 33
6.4 Algorithm to find diagonalization 34

7 Orthogonal Diagonalization **35**
7.1 P^{-1} 35
7.2 Orthogonally diagonalizable 35
7.3 Algorithm to find orthogonal diagonalization 36
7.4 Singular Value Decomposition 36

8 Linear Transformations **38**
8.1 Linearity 39
8.2 Change of Bases 40
8.3 Composition 42
8.4 Ranges 42
8.5 Kernel 43

1 Linear Systems

1.1 Elementary Row Operations

Definition

1. **Type I**: Multiply a row by **non-zero** constant
2. **Type II**: Interchange two rows
3. **Type III**: Add a constant multiple of a row to another row

Inversing EROs

1. **Type I**: Divide scalar number
2. **Type II**: Interchange
3. **Type III**: Minus

$$A \xrightarrow{E_1} \cdot \rightarrow \ldots \xrightarrow{E_k} B$$

$$A \xleftarrow{E_1^{-1}} \cdot \leftarrow \ldots \xleftarrow{E_k^{-1}} B$$

$$E_k \ldots\ldots E_1 A = B$$

$$A = E_1^{-1} \ldots E_k^{-1} B$$

2 Matrices

2.1 Matrix multiplication

Definition

AB is the $m \times n$ matrix, such that its $(i, j)-$entry is

$$a_{i1}b_{1j} + ... + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$$

To find the entry

1. Extract i^{th} row of A
2. Extract j^{th} column of b
3. Multiply componentwise
4. Add products

AB: **pre-multiplication** of A to B.
BA: **post-multiplication** of A to B.

2.2 Representation of linear system

$$Ax = b$$

A: coefficient matrix.
x: variable matrix (solution)
b: constant matrix

2.3 Inverses

Definition

Given a square matrix of order n A,
if there exists square matrix B of order n

$$AB = BA = I_n$$

A is **invertible**, B is an **inverse** of A. *If A is not invertible, it is **singular**.*

2.3.1 Properties of invertible matrices

A square matrix A of order n is invertible $\iff$

1. A is row equivalent to I_n
2. A has n pivot columns (in REF)
3. $Ax = 0$ has only the trivial solution $x = 0$
4. Columns of A are linearly independent
5. $x \to Ax$ is one-to-one:
   $Ax = Ay \implies x = y$
6. For each column vector $b \in \mathbb{R}^n$, $Ax = b$ has an unique solution.
7. Columns of A span $\mathbb{R}^n$
8. $x \to Ax$ is surjection:
   $\forall T : \mathbb{R}^n \to \mathbb{R}^m$, all elements in $\mathbb{R}^m$ can be expressed as Ax in $\mathbb{R}^n$
9. There exists square matrix of order n B such that $BA = I_n$
10. There exists square matrix of order n B such that $AB = I_n$
11. A^T is invertible
12. Columns of A are a basis for $\mathbb{R}^n$
13. Column space of A = $\mathbb{R}^n$
14. Dimension of column space of A = n
15. Rank of $A = 0$

16. Null space of $A = \{0\}$
17. 0 is not an eigenvalue of A.
18. $det(A) \neq 0$
19. Row space of A is $\mathbb{R}^n$
20. $c \neq 0, (cA)^{-1} = \frac{1}{c}A^{-1}$

2.3.2 Finding the inverse

Gauss-Jordan Elimination

1. Write matrix in form $A|I$
2. RREF to get $I|A^{-1}$

Adjoint Matrix

$$A^{-1} = \frac{1}{det(A)}adj(A)$$

2.4 LU Decomposition

Definition

If exclusively Type III operations are used in Gaussian elimination for matrix A, there exists

$$A = LU$$

L: lower-triangular matrix L with diagonal entries 1. U: row-echelon form of A

Finding matrices L and U

1. $A = LU$
2. Find U by getting the REF of A.
3. L is the inverse EROs ($A = E_1^{-1}...E_k^{-1}U, L = E_1^{-1}...E_k^{-1}$)

Solving linear systems with LU decomposition

If $A = LU$, to solve $Ax = b$

1. We get $LUx = b$
2. Substituting $y = Ux$, we get $Ly = b$, where Ly is lower triangular
3. Solve for y, by solving $(L \mid b)$
4. Solve for x, by solving $(U \mid y)$

2.5 Partial pivoting

Definition

If interchanging of rows is needed (Type III operations),

$$A \xleftarrow{E_1^{-1}} \cdot \leftarrow \dots \cdot \xleftarrow{E_i} \cdot \leftarrow \dots \cdot \xleftarrow{E_k^{-1}} U, E_i : R_i \leftrightarrow R_j$$

$$A = E_1^{-1} \dots E_i \dots E_k^{-1} U$$

$$E_i A = (E_i E_1^{-1} \dots) E_i \dots E_k^{-1} U$$

$$P = E_i, L = E_i E_1^{-1} \dots E_i \dots E_k^{-1}$$

For every matrix A, there exists,

- A permutation matrix P: Product of type II elementary matrices
  The matrix of all the Type II EROs
- L: lower-triangular matrix L with diagonal entries 1.
- U: row-echelon form of A

such that $PA = LU$

2.6 Elementary matrices

Definition

A matrix obtained by doing an ERO on an identity matrix.

2.7 Determinant

2 by 2 matrix

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$det(A) = |A| = ad - bc$$

3 by 3 matrix

Let $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$

$$det(A) = |A| = a(ei - fh) - b(di - fg) + c(dh - eg)$$

Elementary matrix of order 3

Let A be a square matrix of order 3, and E be an elementary matrix of order 3.

$$det(EA) = det(E)det(A)$$

If A is invertible, then $R = I, det(R) = 1$,

$$det(A) = det(E_1^{-1})....det(E_k^{-1})$$

Cofactor expansion

Let A_{ij} denote $(i, j)-$cofactor of A.

$$det(A) = a_{i1}A_{i1} + ... + a_{in}A_{in}$$

General strategy:

Let A = $(a_{ij})_{n\times n}$.

- If A has a zero row/zero column, then $det(A) = 0$.
- If A is triangular, $det(A) = a_{ii}...a_{nn}$.
- Otherwise:
  - If $n = 2$, $det(A) = a_{11}a_{22} - a_{12}a_{21}$
  - If row or column has many 0 entries, use cofactor.
  - Use Gaussian elimination:
    $det(A) = (-1)^t det(R)$, t is number of type II operations.

3 Vector Spaces

3.1 Linear span

Definition

Let $S = \{v_1 ..., v_k\}$ be a subset of $\mathbb{R}^n$.
The **linear span** is the set of all linear combinations

$$span(S) = \{c_1 v_1 + ...c_k v_k \mid c_1, ...c_k \in \mathbb{R}\}$$

3.2 Solution space

Definition

The **solution space** of a homogenous linear system of n variables is the **solution set.**

3.3 Linear independence

Definition

Let $S = \{v_1 ..., v_k\}$ be a subset of $\mathbb{R}^n$.
Given the equation $c_1 v_1 + ... + c_k v_k = 0$:

- Non-trivial solution exists
  - S is a linearly dependent set
  - $v_1, ..., v_k$ are linearly dependent.
- Non-trivial solution does not exist
  - S is a linearly independent set
  - $v_1, ..., v_k$ are linearly independent.

- S is linearly dependent $\iff$ some v_i is a linear combination of other vectors in S

- S is linearly independent $\iff$ no vector in S can be written as a linear combination of other vectors

3.4 Basis

Definition

Let S be a subset of vector space V.
S is a basis if it is the minimal subset of V where $span(S) = V$

To check if it is a basis

It is **sufficient** to show two of these conditions to prove S is a basis for V:

1. S is linearly independent
2. $span(S) = V$
3. $dim(V) = |S|$

3.5 Transition matrices

Definition

Coordinate vector, $(v)_S$ is a row vector that represents the scalar coefficient of the basis, S, of a vector space V to represent vector v relative to basis, S.

$$v = c_1v_1 + c_2v_2 + ... + c_kv_k$$

$$(v)_S = (c_1, c_2, ..., c_k)$$

Coordinate vector, $[v]_S$ is the column vector form of $(v)_S$
i.e. $[v]_S = (v)_S^T$
Suppose the column vectors of A forms a basis for vector space V, then $A[v]_S = v$

Definition

Transition matrix, P is a matrix that transitions a coordinator vector, $[v]_S$, relative to basis S to a coordinator vector, $[v]_T$, relative to basis T.
Essentially, Switching the representation of a point in a vector space V from a particular basis S to another particular basis T.
Suppose S and T are basis for vector space V, $P[w]_S = [w]_T$

Properties of Transition Matrix

Transition matrix is transitive.
Suppose S_1, S_2, S_3 are bases for vector space V and P and Q be the transitive matrix from S_1 to S_2 and from S_2 to S_3 respectively. Then,

$$[v]_{S_1} \xrightarrow{P} [v]_{S_2} \xrightarrow{Q} [v]_{S_3}$$

$$[v]_{S_3} = QP[v]_{S_1}$$

P^{-1} is the transition matrix from S_2 tp S_1

$$[v]_{S_1} \xrightarrow{P} [v]_{S_2} \xrightarrow{P^{-1}} [v]_{S_1}$$

Transition matrix from S to T

Used when $span(S) = span(T)$ = vector space, V

Let $S = \{u_1, u_2, u_3\}$ and $T = \{v_1, v_2, v_3\}$
$(T \mid S) = (v_1 \ v_2 \ v_3 \mid u_1 \mid u_2 \mid u_3\,) \xrightarrow[Elimination]{Gauss-Jordan} (I|P)$

3.6 Row & column space

Definition

Row space is the vector space spanned by the rows.
Let $r_1, ...$ be the rows of A.

$$\text{Row space } = span\{r_1, ...\}$$

Column space is the vector space spanned by the cols.
Let $c_1, ...$ be the column of A.

$$\text{Column space } = span\{c_1, ...\}$$

- row space of $A =$ column space of A^T
- column space of $A =$ row space of A^T

3.7 Finding a basis/extending basis

Using row vectors given a set of vectors

Let $V = span\{v_1, ..., v_k\}$

1. View each v_i as a row vector
2. Form a matrix $\begin{pmatrix} v_1 \\ ... \\ v_k \end{pmatrix}$
3. Perform Gaussian elimination to get REF R
4. Non-zero rows of REF are the basis of the vector space

Using column vectors given a set of vectors

Let $V = span\{v_1, ..., v_k\}$

1. View each v_i as a column vector
2. Form a matrix $\begin{pmatrix} v_1 & ... & v_k \end{pmatrix}$
3. Perform Gaussian elimination to get REF R
4. Columns in V that correspond to the pivot columns in R are a basis
   Note: The pivot columns in R form a basis for the column space of R

Extending a basis

Used when $span(S)$ is smaller than $\mathbb{R}^n$

Let $S = \{v_1, ..., v_k\}$

1. Gaussian elimination on R to reduce it to row-echelon form.
   - Remove vectors that are linearly dependent
   - Note columns that are not pivot
2. Add rows to R such that columns are all pivot

3.8 Consistency

Consistency

Let $b \in \mathbb{R}^m$

$$\begin{aligned} Ax = b \text{ is consistent } &\iff Av = b \text{ for some } v \in \mathbb{R}^n \\ &\iff b \text{ is in the column space of } A \end{aligned}$$

Using dimension,
let R be a row echelon form of A.
Thus, row echelon form of $(A \mid b)$ is of the form $(R \mid b')$

$$\begin{aligned} Ax = b \text{ is consistent } &\iff b' \text{ is non pivot in } (R \mid b') \\ &\iff rank(R) = rank(R \mid b') \\ &\iff rank(A) = rank(A \mid b) \end{aligned}$$

3.9 Rank

Definition

Let R be a row echelon form of A.

$$dim(\text{row space of A}) = \text{ number of nonzero rows of } R$$

$$dim(\text{column space of A}) = \text{ number of pivot columns of } R$$

$$rank(A) = dim(\text{column space of A}) = dim(\text{row space of A})$$

Let A be a $m \times n$ matrix.

- $rank(A) = rank(A^T)$
- $rank(A) = 0 \implies A = 0$
- $rank(A) \leq min(m, n)$
- $rank(A) = min(m, n) \implies A$ is full rank

- A full rank $\iff$ A invertible

3.10 Nullspace

Definition

Let A be a $m \times n$ matrix. The **nullspace** of A is the solution space of $Ax = 0$.

$$\{v \in \mathbb{R}^n \mid Av = 0\}$$

$$nullity(A) = dim(\text{nullspace of } A)$$

Let R be a row-echelon form of A.

$$\begin{aligned} nullity(A) &= nullity(R) \\ &= \text{the number of non-pivot columns of } R \end{aligned}$$

3.10.1 Dimension theorem

Dimension Theorem

Let A be a $m \times n$ matrix.

$$rank(A) + nullity(A) = n$$

Properties of product of matrix multiplication

Let A be a $m \times n$ matrix and B be an $n \times p$ matrix.

- column space of $AB \subseteq$ column space of A
- row space of $AB \subseteq$ row space of A
- $rank(AB) \leq min(rank(A), rank(B))$

4 Orthogonality

4.1 Orthogonality

Definition

Dot Product (inner product) of u and v is defined as:

$$u \cdot v = u_1 v_1 + u_2 v_2 = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

= uv^T when u, v are viewed as row vectors

= u$^T v$ when u, v are viewed as column vectors

When $u \cdot v$ is 0, u and v are perpendicular/orthogonal, denoted by $u \perp v$.
(Special Case) When $u = 0$, $0 \cdot v = 0 \Rightarrow 0$ is orthogonal to every vector $v \in \mathbb{R}^n$
Angle between u and v is given by

$$cos\ \theta = \frac{u \cdot v}{||u||||v||}, where\ u, v \neq 0$$

Length of a vector, u is given as:

$$||u|| = \sqrt{u_1^2 + ... + u_n^2} = \sqrt{u \cdot u}, u \in \mathbb{R}^n$$

When $||u|| = 1$, u is a unit vector.

Properties of dot product

- Commutative Law: $u \cdot v = v \cdot u$
- Distributive law: $(u + v) \cdot w = u \cdot w + v \cdot w = w \cdot (u + v)$
- Scalar: $(\mathsf{c}u) \cdot v = u \cdot (\mathsf{c}v) = \mathsf{c}(u \cdot v), \mathsf{c} \in \mathbb{R}$
- $||\mathsf{c}v|| = |\mathsf{c}|||v||$
- $u \cdot u \geq 0$
- $u \cdot u = 0 \Leftrightarrow u = 0$
- $|u \cdot v| \leq ||u||\ ||v||$ (Cauchy-Schwaz ineqality)
- $||u + v|| \leq ||u|| + ||v||$
- $d(u, w) \leq d(u, v) + d(v, w)$ where $d(s, e)$ represents distance from s to e

Definition

Let $S = \{v_1, ..., v_k\}$ be a subset of R^n
S is **orthogonal** if every pair of distinct vectors in S are orthogonal:
$v_i \cdot v_j = 0\ , \forall i \neq j$
S is **orthonormal** if every distinct vector in S is pairwise orthogonal and a unit vector

$$v_i \cdot v_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

- S orthonormal $\Rightarrow S$ orthogonal
- S orthogonal $\Rightarrow T \subseteq S$ is orthogonal
- S orthonormal $\Rightarrow T \subseteq S$ is orthonormal
- S orthogonal $\Rightarrow S \cup 0$ is orthogonal
- S orthonormal $\Rightarrow 0 \notin S$

Definition

Normalizing is the process of converting an orthogonal set of nonzero vectors in $\mathbb{R}^n$, $S = u_1, .., u_k$ to an orthonormal set of vectors, $T = v_1, .., v_k$.

$$u_i \mapsto v_i = \frac{u_i}{||u_i||},\ u_i \in S,\ v_i \in T$$

Check if a set of vectors is orthogonal/orthonormal

Let $A = \{v_1, v_2, ..., v_k\}$ be a subset of $\mathbb{R}^n$.

$$A^T A = \begin{cases} \text{diagonal matrix} & if A \text{ is orthogonal} \\ I_k & if A \text{ is orthonormal} \end{cases}$$

Linear Independance in orthogonal sets

- An orthogonal set of nonzero vectors is linearly independent.
- An orthonormal set of vectors is linearly independent

Orthogonal sets as basis

Let $V = span(S)$, where S is a set of nonzero vectors in $\mathbb{R}^n$

- S is orthogonal set $\Rightarrow$ S is a orthogonal basis for V
- S is orthonormal set $\Rightarrow$ S is a orthonormal basis for V

Checking if S is a basis for V is made simpler as now you have to only check one of the following:

- $|S| = dim(V)$
- $span(S) = V$

The condition that S is linearly indepedent is covered as S is either a nonzero orthogonal set of vectors or a orthonormal set of vectors.

Let $S = \{u_1, ..., u_k\}$ be a basis for a vector space V.

$$\forall w \in V, w = c_1u_1 + ... + c_ku_k$$

$$(w)_S = (c_1, ..., c_k), \text{the coordinate vector of } w \text{ relative to } S$$
$$c_i = \frac{w \cdot u_i}{u_i \cdot u_i} = \frac{w \cdot u_i}{||u_i||^2}$$

Thus, finding the solution, $[w]_S$ can be found easily.

4.2 Projection

Definition

Projection is the vector that is imposed onto a vector v by another vector u.

The project vector, $p = ||p||\frac{v}{||v||} = ||u||cos\theta\frac{v}{||v||} = ||u||\frac{u \cdot v}{||u||||v||}\frac{v}{||v||} = \frac{u \cdot v}{||v||^2}v$

If v is a unit vector, then $p = (u \cdot v)v$

Projection on orthogonal basis

Suppose $S = v_1, ..., v_k$ be an orthonormal basis for vector space V
The projection of w on V is

$$(w \cdot v_1)v_1 + (w \cdot v_2)v_2 + ... + (w \cdot v_k)v_k$$

Suppose $S = u_1, ..., u_k$ be an orthogonal basis for vector space V
The projection of w on V is

$$(w \cdot \frac{u_1}{||u_1||^2})u_1 + (w \cdot \frac{u_2}{||u_2||^2})u_2 + ... + (w \cdot \frac{u_k}{||u_k||^2})u_k$$

Projection on vector space

Suppse $Ax = b$ where the column space of A is V
Method 1:

1. Find a set S such that $V = span(S) \Rightarrow S =$ column vectors of A
2. Use Gram-Schmidt process to convert S to an orthogonal basis for V
3. Find projection p of b onto V

Method 2:

1. Let A be a matrix whose column space is V
2. Find a least squared solution u to $Ax = b$ using $A^T Ax = A^T b$
3. The projection $p = Au$ where u is the solution found in the previous step

4.3 Gram-Schmidt Process

Definition

Gram-Schmidt Process allows us to find an orthogonal basis from a given basis for a vector space.

Let $u_1, u_2, ..., u_k$ be a basis for a vector space V

$$v_1 = u_1$$
$$v_2 = u_2 - \frac{u_2 \cdot v_1}{||v_1||^2} v_1$$
$$\vdots$$
$$v_k = u_k - \frac{u_2 \cdot v_1}{||v_1||^2} v_1 - ... - \frac{u_k \cdot v_{k-1}}{||v_{k-1}||^2} v_{k-1}$$

$\{v_1, v_2, ..., v_k\}$ is an orthogonal basis for V.

After normalizing, $v_i \mapsto w_i = \frac{v_1}{||v_1||}$
$\{w_1, w_2, ..., w_k\}$ is an orthonormal basis for V.

4.4 Least Squares

Least Squared Solutions

If $Ax = b$ is inconsistent, its least squared solution, u, is $Au = p$ where p is the projection of b onto V

Method 1:
Steps:

1. Set the columns space of A as V
2. Use Gram-Schmidt process to find the orthogonal basis for V
3. Find the projection of b onto V
4. Solve for $Ax = p$, going to RREF

Method 2:
Solve the system $A^T Ax = A^T b$
The method is guaranteed to get a solution regardless of consistency of system.

Remark: From Method 2: $x = (A^T A)^{-1}(A^T b)$.
This can be used to find the projection of w on a vector space V given the basis of V, which would be A.
Remark: Given two least-squared solutions v_1, v_2, if A has an eigenvalue 0,

$$\begin{aligned} A^T Av_1 =& A^T b \\ A^T Av_2 =& A^T b \\ A^T A(v_1 - v_2) =& A^T(b - b) \\ A^T A(v_1 - v_2) =& 0 \end{aligned}$$

$v_1 - v_2$ is thus in the nullspace of $A^T A$, making it an eigenvector associated to 0.
Remark: Least-squared solutions are **non-unique**, unless the matrix A is linearly independent.

4.5 QR Decomposition

QR Decomposition

Let A be a $m \times n$ matrx whose column are linearly independent. Then there exist:

- an $m \times n$ matrix Q whose column form a orthonormal set
- an invertiable upper triangluar matrix R of order n

Steps:

1. Use Gram-Schmidt process to obTain an orthonormal basis $\{w_1, w_2, ..., w_n\}$ for the column space of A
2. $Q = \begin{pmatrix} w_1 & w_2 & ... & w_n \end{pmatrix}$
3. $R = \begin{pmatrix} w_1 \cdot u_1 & w_1 \cdot u_2 & \dots & w_1 \cdot u_n \\ 0 & w_2 \cdot u_2 & \dots & w_2 \cdot u_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n \cdot u_n \end{pmatrix}$

Application:

Suppose A has linearly independent columns with QR decomposition QR

$$Ax = b \Leftrightarrow QRx = b \Leftrightarrow Rx = Q^T b$$

where u is the least squares solution to $Rx = Q^T b$ which is also the least squares solution to $Ax = b$

Full QR Decomposition

Used when the basis may not be independent.

Suppose $V = span(S), S = \{u_1, u_2, ..., u_k\}$ where S may not be independent.
$\Rightarrow$ Gram-Schmidt process generates an orthonormal basis by ignoring the zero vector, 0.

$$v_1 = u_1$$

$$v_2 = u_2 - \frac{u_2 \cdot v_1}{||v_1||^2} v_1 = 0, \text{ Vector is ignored}$$

$$v_3 = u_3 - \frac{u_3 \cdot v_1}{||v_1||^2} v_1$$

$$\vdots$$

After normalizing, $v_i \mapsto w_i = \frac{v_1}{||v_1||}$
$\{w_1, w_3, ...\}$ is an orthonormal basis.
Since v_2 is a zero vector, we need to extend the orthonormal basis by choosing a arbituary vector (preferably elementary vectors such as $e_1 = (1, 0, 0)$) and process it through the above Gram-Schmidt process to obtain a hopefully non-zero orthogonal basis and normalise it to obtain the extended orthonormal basis.

Let A be an $m \times n$ matrix with $m \geq n$. Then there exist

- An orthogonal matrix Q of order m
- An invertible upper triangular matrix R' of order n

such that $A = QR$, where $R = \begin{pmatrix} R' \\ 0 \end{pmatrix}$ is an $m \times n$ matrix

Let $\{u_1, u_2, ..., u_m\}$ be the columns of A

1. Use Gram-Schmidt process to find an orthogonal basis.
2. Extend it to an orthonormal basis $\{w_1, ..., w_m\}$ for $\mathbb{R}^m$ if necessary.
3. $Q = \begin{pmatrix} w_1 & \dots & w_m \end{pmatrix}$
4. $R' = \begin{pmatrix} w_1 \cdot u_1 & w_1 \cdot u_2 & \dots & w_1 \cdot u_n \\ 0 & w_2 \cdot u_2 & \dots & w_2 \cdot u_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n \cdot u_n \end{pmatrix}$ a upper triangular matrix.

4.6 Orthogonal Matrices

Definition

Orthogonal Matrix is a square matrix where $A^TA = I$ or equivalently $A^{-1} = A^T$

Properties:
Let A and B be orthogonal matricies of the same order

$$(AB)^T(AB) = B^TA^TAB = I$$

AB is a orthogonal matrix
Let S = $\{v_1, ..., v_k\} \subseteq \mathbb{R}^n$ and $A = \begin{pmatrix} v_1 & \dots & v_k \end{pmatrix}$

$$A^TA = I_k \Leftrightarrow S \text{ is an orthonormal set}$$

If $k = n$, then A is a square matrix and $|S| = dim(\mathbb{R}^n) \Rightarrow A$ is orthogonal $\Leftrightarrow A^TA = I_n \Leftrightarrow S$ is an orthonormal basis for $\mathbb{R}^n$

Let A now be a $m \times n$ matrix
$A^TA = I_n \Leftrightarrow$ columns of A form an orthonormal set in $\mathbb{R}^m$
$AA^T = I_n \Leftrightarrow$ rows of A form an orthonormal set in $\mathbb{R}^n$

Specialised to a orthogonal matrix, A, of order n:
$\Leftrightarrow$ columns of A form an orthonormal set in $\mathbb{R}^n$
$\Leftrightarrow$ rows of A form an orthonormal set in $\mathbb{R}^n$

Premultiplication of Orthogonal matrix with another matrix

Let $A = \begin{pmatrix} v_1 & \dots & v_k \end{pmatrix}$ where v_1 to v_k represent the column vectors of A,
$PA = \begin{pmatrix} Pv_1 & \dots & Pv_k \end{pmatrix}$
Then, $(PA)^T(PA) = A^T P^T PA = I \Rightarrow \{Pu_1, ..., Pu_k\}$ is an orthonormal set

If A consist of orthonormal set of column vectors, then PA is also a orthonormal set of vectors
Essentially, the pre-multiplication of an orthogonal matrix converts an orthonormal set to another orthonormal set.

Transition matrix of orthogonal bases

Let $A = \{u_1, u_2, u_3\}$ and $B = \{v_1, v_2, v_3\}$ be orthonormal bases for V. Let $w \in V$

$$w = A[w]_S = B[w]_T$$

$$[w]_T = B^T B[w]_T = B^T A[w]_S \Rightarrow P = B^T A \text{ is transition matrix from S to T.}$$

$$[w]_S = A^T A[w]_S = A^T B[w]_T \Rightarrow Q = A^T B \text{ is transition matrix from T to S.}$$

$P^{-1} = Q$ Both P and Q are orthogonal matrix.
$\therefore$ The transition matrix between orthonormal bases is an orthogonal matrix

5 Eigenvalues, eigenvectors & eigenspaces

5.1 Eigenvalues & eigenvectors

Definition

Let A be a square matrix of order n.
Suppose that for some $\lambda \in \mathbb{R}$ and nonzero $v \in \mathbb{R}^n$:

$$Av = \lambda v$$

Then,

- λ is an **eigenvalue** of A.
- v is an **eigenvector** of A associated to the eigenvalue λ.

5.2 Characteristic equation

$$\begin{aligned} \lambda \in \mathbb{R} \text{ is an eigenvalue of A} &\iff Av = \lambda v \text{ for some } 0 \neq v \in \mathbb{R}^n \\ &\iff \lambda I v - Av = 0 \text{ for some } 0 \neq v \in \mathbb{R}^n \\ &\iff (\lambda I - A)v = 0 \text{ for some } 0 \neq v \in \mathbb{R}^n \\ &\iff \lambda I - A \text{ is singular} \\ &\iff det(\lambda I - A) = 0 \end{aligned}$$

Additionally,

$$0 \text{ is not an eigenvalue of } A \iff A \text{ is an invertible matrix}$$

$\boxed{\text{Remark:}}$ Note that $A^T A$ has the same nullspace as A.
Thus, if A has an eigenvalue of 0, the eigenvector v associated to eigenvalue 0 is in the nullspace of both $A^T A$ and A.

Main Theorem for Invertible Matrices

A is an invertible matrix, if and only if any (and hence, all) of the following hold:

- columns of A form a basis for $\mathbb{R}^n$
- $rank(A) = n$
- $nullity(A) = 0$
- 0 is not an eigenvalue of A.

Definition

The **characteristic polynomial** of A:

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

The **characteristic equation** of A:

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

Eigenvalues are the **roots** of the characteristic equation.

Special case for triangular matrices

Eigenvalues of a triangular matrix are its diagonal entries.

5.3 Eigenspace

Let A be a square matrix of order n, and let λ be an eigenvalue of A.
For any nonzero vector $v \in \mathbb{R}^n$,

$$\begin{aligned} v \text{ is an eigenvector of } A \text{ associated to } \lambda &\iff Av = \lambda v \\ &\iff (\lambda I - A)v = 0 \\ &\iff v \in \text{ nullspace of } (\lambda I - A) \end{aligned}$$

Definition

Let λ be an eigenvalue of a square matrix A.

The **eigenspace** of A associated to λ, denoted $E_{A,\lambda}$ or E_λ, is the nullspace of $(\lambda I - A)$.
Note: $(\lambda I - A)$ *is singular* $\therefore dim(E_\lambda) \geq 1$

Eigenvectors

Let λ be an eigenvalue of a square matrix A.

The eigenvectors of A associated to the eigenvalue λ are precisely all nonzero vectors in the eigenspace E_λ.

6 Diagonalization

6.1 Diagonalizable matrices

Definition

A square matrix A is **diagonalizable** if
there is an invertible matrix P where $P^{-1}AP$ is a diagonal matrix.

6.2 Algebraic and geometric multiplicity

Definition

Algebraic multiplicity of λ, denoted $a(\lambda)$ is its multiplicity as a root of the characteristic polynomial of A.
e.g: If the characteristic equation is $(\lambda - 1)^3 = 0$, $a(1) = 3$

Geometric multiplicity of λ, denoted $g(\lambda)$ is the dimension of E_λ.

Algebraic & geometric multiplicity

Let A be a square matrix, λ be an eigenvalue of A. Then,

$$1 \leq g(\lambda) \leq a(\lambda)$$

6.3 Criterion of diagonalizability

Diagonalization Theorem

Let A be a square matrix of order n.

A is diagonalizable $\iff$ A has n linearly independent eigenvectors.
$\iff$ sum of geometric multiplicities of eigenvalues of $A = n$
$\iff$ sum of algebraic multiplicities of eigenvalues of $A = n$
and for each eigenvalue $\lambda, g(\lambda) = a(\lambda)$.

If a square matrix A has n distinct eigenvalues, then A is diagonalizable.

Remark: The inverse is not true. If A is diagonalizable, A might not have n distinct eigenvalues.

Remark: The diagonal matrix D is not unique unless A only has one eigenvalue.

Remark: For this course, diagonalizable = diagonalizable over $\mathbb{R}$

6.4 Algorithm to find diagonalization

Algorithm

Let A be a square matrix of order n.

1. $det(\lambda I - A)$ cannot be completely factorized in $\mathbb{R}$
   - A is not diagonalizable.
2. $det(\lambda I - A)$ can be completely factorized in $\mathbb{R}$
   - Let $\lambda_1, ...\lambda_k$ be distinct eigenvalues of A.
   - Find basis S_i for eigenspace E_{λ_i}.

   (a) $g(\lambda_i) < a(\lambda_i)$ for some i
      - A is not diagonalizable.

   (b) $g(\lambda_i) = a(\lambda_i)$ for all i
      - A is diagonalizable.

7 Orthogonal Diagonalization

7.1 P^{-1}

Given $P^{-1}AP = D$, the inverse of P, P^{-1} is required for computations such as:

- Gauss-Jordan elimination: $(P|I) \to (I|P^{-1})$
- Adjoint matrix: $P^{-1} = \frac{1}{det(P)} adj(P)$

$\boxed{\text{Remark:}}$ P is orthogonal $\implies P^{-1} = P^T$

7.2 Orthogonally diagonalizable

Definition

A square matrix A is **orthogonally diagonalizable** if there exists an orthogonal matrix P such that P^TAP is a diagonal matrix.

All orthogonally diagonalizable matrices are symmetric

Let A be a square matrix.

$$A \text{ is orthogonally diagonalizable} \implies A \text{ symmetric}$$

$P^TAP = D = D^T = (P^TAP)^T = P^TA^TP$

$\boxed{\text{Remark:}}$ Given that A is a symmetric matrix with all real entries,

- eigenvalues are all real
- for each eigenvalue λ of $A, g(\lambda) = a(\lambda) = dim(E_\lambda)$
- eigenvectors associated to distinct eigenvalues are orthogonal.

7.3 Algorithm to find orthogonal diagonalization

Algorithm

Let A be a symmetric matrix of order n.

1. Find all eigenvalues of A by solving $det(\lambda I - A) = 0$
2. For each eigenvalue λ_i of A,
   (a) Solve $(\lambda_i I = A)x = 0$ to find basis S_i for eigenspace E_{λ_i}
   (b) Use Gram-Schmidt to convert S_i to orthonormal basis T_i.
3. $T_i \cup .. \cup T_k = \{w_1, ..., w_n\}$ is an orthonormal basis for $\mathbb{R}^n$.
   - $P = \{w_1, ..., w_n\}$ orthogonally diagonalizes A.

7.4 Singular Value Decomposition

Non-negativity of eigenvalues of $A^T A$

If λ was an eigenvalue of $A^T A$, and v was an associated eigenvector:

$$A^T A v = \lambda v \implies v^T A^T A v = v^T \lambda v \implies ||Av||^2 = \lambda ||v||^2 \geq 0$$

Thus,
let A be a $m \times n$ matrix.

The eigenvalues of $A^T A$ are nonnegative.

Definition

Let A be a $m \times n$ matrix, and $\lambda_1, ...$ be its eigenvalues.

$$\sigma_1 = \sqrt{\lambda_1} ......$$

These values are the **singular values** of A.

Algorithm: Singular Value Decomposition

Let A be an $m \times n$ matrix.

1. Find eigenvalues of $A^T A$ $\lambda_1 > ... > \lambda_r + 1$
2. Find corresponding orthonormal set of eigenvectors of $A^T A$ $\{v_1....v_n\}$.
3. Let $\sigma_i = \sqrt{\lambda_i}$, and $u_i = \frac{1}{\sigma_i} A v_i$ for $i = 1, ..., r$
4. Extend $\{u_i, ..., u_r\}$ to orthonormal basis $\{u_i, ..., u_m\}$ for $\mathbb{R}^m$

Let $U = \{u_1...u_m\}, V = \{v_1....v_n\}$. U, V are orthogonal matrices.

Let $\sum = \begin{pmatrix} \sigma_1 & ... & 0 & ... & 0 \\ ... & ... & ... & ... & 0 \\ 0 & ... & \sigma_r & ... & 0 \\ ... & ... & ... & ... & 0 \\ 0 & ... & 0 & ... & 0 \end{pmatrix}$, a **rectangular diagonal matrix**.

Then, $A = U \sum V^T$ is the **singular value decomposition** of A.

8 Linear Transformations

Definition

Linear tranformation can be seen as a mapping f, from domain of $\mathbb{R}^n$ to a codomain of $\mathbb{R}^m$

A mapping T: $\mathbb{R}^n \mapsto \mathbb{R}$ can be defined by:

$$f(x_1, x_2, ..., x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n$$

is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}$

A linear transformation can be viewed in the matrix form:

$$T(x) = Ax, x \in \mathbb{R}^n$$

A is called the **standard matrix** for T

Remark: T is called a linear operator on $\mathbb{R}^n$ if dimension of domain = dimension of codomain

Linear operator

Identity operator is the mapping T: $\mathbb{R}^n$ to $\mathbb{R}^n$ where $I(x) = x$ for $x \in \mathbb{R}^n$

$$I(x) = x = I_nx \Rightarrow I_n \text{ is the standard matrix for } I(x)$$

Zero transformation is the mapping T: $\mathbb{R}^n$ to $\mathbb{R}^m$ where $)(x) = 0$ for $x \in \mathbb{R}^n$

$$O(x) = O = 0_{m\times n}0 \Rightarrow 0_{m\times n} \text{ is the standard matrix for } O(x)$$

Remark: T: $\mathbb{R}^n$ to $\mathbb{R}^n$ has a unique standard transformation

To show that a function T: $\mathbb{R}^n$ to $\mathbb{R}^m$ is a linear transformation, it **suffices** to find an $m \times n$ matrix A such that $T(x) = Ax \quad \forall x \in \mathbb{R}^n$

8.1 Linearity

Linearity

Suppose T: $\mathbb{R}^n$ to $\mathbb{R}^m$ is a linear transformation and A is the standard matrix of T, i.e $T(x) = Ax$.

- $T(0) = A0 = 0$
- $T(cv) = A(cv) = c(Av) = cT(v)$
- $T(u+v) = A(u+v) = Au + Av = T(u) + T(v)$
- For any $v_1, ..., v_k \in \mathbb{R}^n$ and $c_1, ..., c_k \in \mathbb{R}$,

$$T(c_1v_1 + ... + c_kv_k) = A(c_1v_1 + ... + c_kv_k)$$

$$= A(c_1v_1) + ... + A(c_kv_k)$$

$$= c_1(Av_1) + ... + c_k(Av_k)$$

$$= c_1T(v_1) + ... + c_kT(v_k)$$

which is the linear transformation of images

$T(c_1v_1 + ... + c_kv_k)$ is completely determined by $T(v_1), ..., T(v_k)$

Determine mapping T is NOT a linear transformation

To show that a mapping T is not a linear transformation

- Show that $T(0) \neq 0$
- Find $v \in \mathbb{R}^n$ and $c \in \mathbb{R}$ such that $T(cv) \neq cT(v)$
- Find $v \in \mathbb{R}^n$ and $c \in \mathbb{R}$ such that $T(u+v) \neq T(u) + T(v)$

8.2 Change of Bases

Change of basis

Suppose $S = \{v_1, .., v_n\}$ be a basis for $\mathbb{R}^n$
$\forall v \in \mathbb{R}^n, (v)_S = (c_1, ..., c_n)$
$v = c_1 v_1 + ... + c_n v_n = \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix} [v]_S = P[v]_S$
Suppose T: $\mathbb{R}^n$ to $\mathbb{R}^m$ is a linear transformation

$$\begin{aligned} T(v) &= c_1 T(v_1) + ... + c_n T(v_n) \\ &= \begin{pmatrix} T(v_1) & \dots & T(v_n) \end{pmatrix} [v]_S \\ &= B[v]_S \\ &= AP[v]_S \end{aligned}$$

$$B = AP \Leftrightarrow A = BP^{-1}$$

To find the standard matrix of T:
$(P|I) \xrightarrow[Elimination]{Gauss-Jordan} (I|P^{-1})$
The standard matrix for T would be BP^{-1}

Suppose T: $\mathbb{R}^n$ to $\mathbb{R}^m$ and $S_1 = \{v_1, .., v_n\}$ and $S_2 = \{u_1, .., u_n\}$ are bases for $\mathbb{R}^n$

$$T(w) = B[w]_{S_1} = \begin{pmatrix} T(v_1) & \dots & T(v_n) \end{pmatrix} [w]_{S_1}$$

$$= C[w]_{S_2} = \begin{pmatrix} T(u_1) & \dots & T(u_n) \end{pmatrix} [w]_{S_2}$$

Let P be the transition matrix from S_1 tp S_2

$$P[w]_{S_1} = [w]_{S_2} \Rightarrow CP[w]_{S_1} = C[w]_{S_2} = T(w) = B[w]_{S_1}$$

$$B = CP$$

If $S_2 = E$ is the standard basis,
$P = \begin{pmatrix} u_1 & \dots & u_n \end{pmatrix}$ and P^{-1} is the transition matrix from E to S
$C = A$ which is the standard matrix of A
$B = AP \Leftrightarrow A = BP^{-1}$
Suppose T: $\mathbb{R}^n \mapsto \mathbb{R}^n$ and $S = \{v_1, .., v_n\}$ represented by P as $\begin{pmatrix} v_1 & \dots & v_n \end{pmatrix}$. P is invertible.

$$v = P[v]_S$$

$$T(v) = P[T(v)]_S \text{ and } Av = AP[v]_S$$

$$P[T(v)]_S = A[v]_S \Rightarrow [T(v)]_S = P^{-1}AP[v]_S$$

T can be represented by $[v]_S \mapsto B[v]_S$ where $B = P^{-1}AP$. A and B are similar.
A square matrix is diagonalisable $\Leftrightarrow$ it is similar to a diagonal matrix.

8.3 Composition

Definition

Let S: $\mathbb{R}^n \mapsto \mathbb{R}^m$ and T: $\mathbb{R}^m \mapsto \mathbb{R}^k$ be linear transformations

Let $T \circ S$: $\mathbb{R}^n \mapsto \mathbb{R}^k$,

Composition of T with S: $(T \circ S)(u) = T(S(u)) \ \forall u \in \mathbb{R}^n$

$\boxed{\text{Remark:}}$ $T \circ S \neq S \circ T$

Let A and B be the standard matrix for the above mapping S and T respectively.

The standard matrix of $T \circ S = BA$

8.4 Ranges

Definition

Let T: $\mathbb{R}^n \mapsto \mathbb{R}^m$

The range of T is the set of all images of T: $R(T) = \{T(v) | v \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$

$R(T) = span(T(v)) = span(T(v_1), ..., T(v_n))$ where $v_1, ..., v_n$ is any basis for $\mathbb{R}^n$

Since T has a standard matrix $A = \begin{pmatrix} T(e_1) & \dots & T(e_n) \end{pmatrix}$

- $R(T) =$(column space of A)
- rank(T) = dim($R(T)$) = rank(A)

8.5 Kernel

Definition

Let T: $\mathbb{R}^n \mapsto \mathbb{R}^m$ with standard matrix A

Kernel of T is the set of all vectors in $\mathbb{R}^n$ whose image is $0 \in \mathbb{R}^m$

$$\begin{aligned} Ker(T) =& \{v \in \mathbb{R}^n \mid T(v) = 0\} \\ =& \{v \in \mathbb{R}^n \mid Av = 0\} \\ =& \text{nullspace of } A \end{aligned}$$

$\boxed{\text{Remark:}}$ Since $T(0) = 0$, $(0 \in \mathbb{R}^n) \in Ker(T)$

The nullity of T, $nullity(T) = dim(Ker(T)) = nullity(A)$

Dimension Theorem for Linear Transformations

Let T: $\mathbb{R}^n \mapsto \mathbb{R}^m$ with standard matrix A

$$R(T) = \ column\ space\ of\ A$$

$$rank(T) = rank(A)$$

$$Ker(T) = \ nullspace\ of\ A$$

$$nullity(T) = nullity(A)$$

$\therefore rank(T) + nullity(T) = n$