

Linear Algebra

Version: 1.0. December 24, 2024

Ramasamy Kandasamy

1. Matrices and Gaussian Eliminations

1.1. Geometry of linear equations

Matrix representation: $Ax = b$

- *Row picture*: Lines in a (\mathbb{R}^2) or planes (\mathbb{R}^3) .
- *Column picture*: b is linear combination of the columns of A . $\therefore b$ is in the column space of A .

Singular case: No solution or infinite solution.

1.2. Gaussian Elimination

- Apply row operations and row exchange to convert the system of linear equations to triangular form.
- Use back substitution to solve the system.

Row operations preserve the nullspace and row space of A , but this normally alters the eigenvalues.

Gaussian elimination in matrix form

- Each row operation on A can be represented as a matrix multiplication on A . If a triangular form U is obtained by three row operations on A , each of them can be represented by matrices, say G, F and E . Then, $G \times F \times E \times A = U$.
- The product of G, F and E would be a lower triangular matrix (L). The matrix A can be written as $PA = LU$. L can be found by keeping track of the row operations. P is a permutation matrix.
- LU can also be written as LDU by splitting U to DU . D is a diagonal matrix containing the pivot elements.

1.3. Matrix multiplication

5 ways to multiply matrices

Given, $A \times B = C$

- Standard.
- Each column of C is a linear combination of columns of A . $[Ab_1 \ \cdots \ Ab_n]$, where b_i s are columns of B .
- Each row of C is a linear combination of rows of B .

$$\begin{bmatrix} a_1 B \\ \vdots \\ a_n B \end{bmatrix}, \text{ where } a_i \text{s are rows of } A.$$

$$\sum_{i=1}^n [ca_i] \times [rb_i]$$

Where, ca_i and rb_i are i -th column and row of A and B , respectively.

- Multiplying Blocks of matrix. Eg:

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \times \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix}$$

A_i, B_i and C_i are blocks of the matrices A, B and C . The blocks are treated as a single entity and multiplied like a normal matrix. Eg: $C_1 = A_1 B_1 + A_2 B_3$

Last approach would be useful for proving matrix theorems by inductions.

Properties

- Associative: $(AB)C = A(BC)$.
- Distributive: $A(B + C) = AB + AC$; $(B + C)D = BD + CD$.
- **NOT** commutative: Usually $EF \neq FE$.
- $AB = 0 \not\Rightarrow A = 0$ (or) $B = 0$.
- $AC = AD \not\Rightarrow C = D$
- Let A, B, C be $n \times n$ matrices. Then:
- If $\text{rank}(A) = n$ and $AB = AC$, then $B = C$.
- If $\text{rank}(A) = n$, then $AB = 0 \Rightarrow B = 0$. If $AB = 0$ but $A \neq 0$ and $B \neq 0$, then $\text{rank}(A) < n$ and $\text{rank}(B) < n$.

1.4. Inverse and Transpose of a Matrix

If the system of equations is non-singular, then it could be solved by using inverse of A .

- A^{-1} is inverse of $A \Leftrightarrow A^{-1}A = A^{-1}A = I$.
- If A^{-1} exist, then A is invertible or non-singular matrix.

Properties

- A square matrix is singular or not invertible if:
 - Its determinant is 0.
 - $\text{Dimension}(\text{Column Space}) < \text{Matrix dimension}$.
Proof (by contradiction): $\exists x \neq 0, \exists Ax = 0$.
Assuming A^{-1} exists $\Rightarrow A^{-1}Ax = A^{-1}0 \Rightarrow x = 0$.
- If A is invertible, then the one and only solution to $Ax = b$ is $x = A^{-1}b$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(AB)^T = B^T A^T$
- $(A^{-1})^T = (A^T)^{-1}$
- For any rectangular matrix R , RR^T and $R^T R$ is symmetric.
- A symmetric matrix could be decomposed to LDL^T .

Finding inverse by Gauss-Jordan elimination

Start with augmented matrix. Achieve upper triangular matrix on the left part and then identity matrix by row operations:

$$[A|I] \longrightarrow [U|X] \longrightarrow [I|A^{-1}]$$

2. Vector Spaces

2.1. Vector spaces and subspaces

Vector space Set of vectors with vector addition and scalar multiplication, satisfying the following properties.

- Associativity of vector addition.
- Commutativity of vector addition.
- Identity element of vector addition.
- Inverse elements of vector addition.
- Compatibility of scalar multiplication with field multiplication. i.e., $a(bv) = ab(v)$
- Identity element of scalar multiplication.
- Distributivity of scalar multiplication w.r.t. vector addition. i.e., $a(u + v) = au + av$
- Distributivity of scalar addition w.r.t. vector multiplication. i.e. $(a + b)v = av + bv$

If S and P are two vector spaces, $S \cup P$ need not be a vector space, but $S \cap P$ is a vector space.

Subspace $Q \subseteq P$ is a subspace of P , if it satisfies the following property in addition to all the properties of a vector space.

- $a \in Q$ and $b \in Q \Rightarrow a + b \in Q$.
- $a \in Q \Rightarrow ca \in Q$ for all scalar c .

2.2. Solving $Ax = b$ [Revisited]

Augmented matrix $[A|b] \rightarrow$ Row operation leading to echelon form.

Pivot elements First non-zero element in a row with a column of zeroes below it.

Rank = number of pivot elements.

Free columns = Columns without pivot elements.

Null space: of a matrix A , is the set of all x that satisfies $Ax = 0$.

General solution: $x = x_p + x_n$, where x_n is a vector from the nullspace of A .

The existence and number of solutions depends on the rank of A and that of the augmented matrix $[A|b]$.

2.3. Linear independence, basis and dimension

Linear independence: A set of vectors v_1, v_2, \dots, v_k are said to be linearly independent iff

$$c_1 v_1 + c_2 v_2 + \dots + c_k v_k = 0 \Rightarrow c_1 = c_2 = \dots = c_k = 0.$$

Linear Combination: $LC(v_i) = \{u \mid u = \sum_j c_j v_j\}$

Span of a set of vectors $\{v_1, v_2, \dots, v_k\} \subset V$ is the smallest subspace $U \subseteq V$ containing these vectors.

Theorem: $LC(\{v_i\}) = \text{Span}(\{v_i\})$.

Basis: Basis of a vector space is a set of vectors with the following two properties.

- The vectors in the set are linearly independent.
- They span the vector space.

Theorems

- If $A \subseteq V$ is linearly independent and $B \subseteq V$ is a basis of V , then $|A| \leq |B|$.
- If $B^i \subseteq V$ and $B^j \subseteq V$ are each a basis of V then $|B^i| = |B^j|$.

Dimension of a vector space: Size of the basis of the vector space.

2.4. Four fundamental subspaces

For a matrix A of dimensions $m \times n$ and rank r .

- Column space of A , $C(A)$: Smallest subspace containing the columns of A .
 $C(A) \subseteq \mathbb{R}^m$.
Dimension of $C(A)$ is r .
- Row space of A , $R(A)$: Smallest subspace containing the rows of A .
 $R(A) \subseteq \mathbb{R}^n$.
Dimension of $R(A)$ is r .
- Null space of A , $N(A)$: Set of all x such that $Ax = 0$.
 $N(A) \subseteq \mathbb{R}^n$.
Dimension of $N(A)$ is $n - r$.
- Null space of A^T , $N(A^T)$: Set of all x such that $A^T x = 0$.
 $N(A^T) \subseteq \mathbb{R}^m$.
Dimension of $N(A^T)$ is $m - r$.

Relationship between the subspaces:

- $C(A) \perp N(A^T)$ and $C(A) \cup N(A^T) = \mathbb{R}^m$.
- $R(A) \perp N(A)$ and $R(A) \cup N(A) = \mathbb{R}^n$.
- $C(A)$ and $R(A)$ are isomorphic.
- In otherwords, A is invertible within the domain / co-domain of $C(A)$ and $R(A)$.

Proof outline: If $x \notin R(A)$, find $x_\perp \perp R(A)$ and show that $Ax_\perp = 0$.

2.5. Linear Transformation

Linear Transformation: A transformation T that converts a vector u to $v (= T(x))$, is a linear transformation if it satisfies the following properties.

- $T(0) = 0$.
- $T(u + v) = T(u) + T(v)$
- $T(au) = aT(u)$

Important notes:

- Every matrix multiplication to a vector is a linear transformation.
- Every linear transformation can be represented by a matrix multiplication.
- NOTE: The transformation matrix need not be a square matrix.

- If $Ax / T(x)$ is known for all the basis of the vector space, then it is known for all the vectors in vector space.

Examples of linear transformation:

Scaling, Rotation, Reflection, Projection, Differentiation and Integration.

2.6. Change of Basis and Similarity Transform

$x \rightarrow$ In standard basis.

$x_B \rightarrow$ w.r.t. the basis B .

$x_{B^1} \rightarrow$ w.r.t. the basis B^1 .

$x_{B^2} \rightarrow$ w.r.t. the basis B^2 .

$x_{B^1} = B^{1-1}x$, or, $x = B^1 x_{B^1}$

$x_{B^2} = B^{2-1}B^1 x_{B^1}$

Similarity Transform:

Give a transformation matrix w.r.t one basis get the equivalent transformation another basis.

Eg: Standard to B :

Say, $y = Ax$, we want A' such that $y_B = A'x_B$.

$\Rightarrow By_b = ABx_b, \forall x \in V. \Rightarrow y_b = B^{-1}ABx_b \Rightarrow A' = B^{-1}AB$.

3. Orthogonality

3.1. Orthogonal vectors and subspaces

Orthogonal vectors: x and y are orthogonal iff $x^T y = 0$.

Length of a vector: $\|x\| = x^T x$.

Orthogonal subspaces: Two subspaces V and W are said to be orthogonal if $\forall v \in V$ and $\forall w \in W$ $v \perp w$.

Orthogonal complement: Given a subspace V in \mathbb{R}^n , the space of all vectors orthogonal to V is called orthogonal complement of V . It is denoted by V^\perp .

Theorems:

- If a set of vectors v_1, v_2, \dots, v_k of non-zero length are mutually perpendicular, then those vectors are linearly independent.
- Row space is orthogonal complement of null space.
- Column space is orthogonal complement of left null space.

A deeper meaning of matrix multiplication

The matrix multiplication Ax transforms the row space component of x (x_r) to column-space of A and the nullspace component of x (x_n) to 0.

Here $x = x_r + x_n$.

The real action is between row space and column space.

3.2. Cosines and Projections onto lines

Cosine The cosine of the angle between any two non-zero vectors is defined as follows:

$$\cos\theta = \frac{a^T b}{\|a\| \|b\|}$$

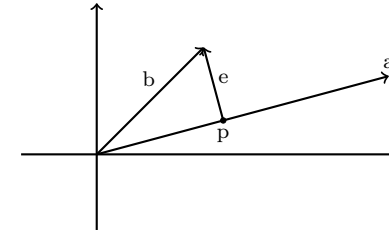
Law of cosines

$$\|b - a\|^2 = \|a\|^2 + \|b\|^2 - 2\|a\| \|b\| \cos\theta$$

Schwarz inequality

$$\|a^T b\| \leq \|a\| \|b\|, \because \cos\theta \leq 1$$

Projection onto a line (1D subspace)



p is projection of b on a . $\therefore p = xa$ for some x .

$e = b - p$ and $e \perp a$,

$$\therefore a^T (b - xa) = 0$$

$$xa^T a = a^T b$$

$$\Rightarrow x = \frac{a^T b}{a^T a}$$

$$p = a \frac{a^T b}{a^T a}$$

This can be written as,

$p = \frac{aa^T}{a^T a} b = Pb$, where $P = \frac{aa^T}{a^T a}$ is the projection matrix of a .

Properties of P

- $C(P)$ = line through a .
- $\text{Rank}(P) = 1$.
- P is symmetric: $P = P^T$
- $P^2 = P$

Projection onto a higher dimension subspace

Finding projection of b on A

Let p be the projection of b on subspace spanned by columns of A , then there exists \hat{x} such that $A\hat{x} = p$

$$b - p \perp A \Rightarrow A^T (b - A\hat{x}) = 0$$

$$\therefore \hat{x} = (A^T A)^{-1} A^T b$$

$$\therefore p = A(A^T A)^{-1} A^T b$$

Here the projection matrix $P = A(A^T A)^{-1} A^T$

If A is invertible then $P = I$

$A^T A$ has the same nullspace of A
i.e., $A^T A$ is invertible if columns of A are independent.

Least squares problem Solve: $Ax = b$

- Normal equations: $A^T A \hat{x} = A^T b$
- Best estimate: $\hat{x} = (A^T A)^{-1} A^T b$
- Projection: $p = A \hat{x} = A(A^T A)^{-1} A^T b$
- $b = p + e$
- p is in $C(A)$
- e is in $N(A^T)$
- If b is in $C(A)$,
 $p = A(A^T A)^{-1} A^T b = A(A^T A)^{-1} A^T A x = A x = b$
- If b is in $N(A^T)$, $p = A(A^T A)^{-1} A^T b = A(A^T A)^{-1} 0 = 0$
- If A is invertible, $p = A(A^T A)^{-1} A^T b = b$
- (I think) $\|e\|^2 = \|b - p\|^2$ is the squared error.

3.3. Orthogonal basis and Gram-Schmidt

Orthogonal basis: A set of bases, q_1, q_2, \dots, q_k , is said to be orthogonal bases, if for all $i \neq j$ $q_i^T q_j = 0$.

Orthonormal basis: Orthonormal bases, q_1, q_2, \dots, q_k , satisfies the following property:

$$q_i^T q_j = \begin{cases} 1 & \text{if } i = j. \\ 0 & \text{otherwise} \end{cases}$$

Gram-Schmidt process

- Eg: A case with two bases a and b :
 $a' = a$
 $b' = b - \frac{aa^T}{a^T a} b$, here, $b' = e$ and $b = p + e$
- Eg: Case with three bases a, b and c .
 $a' = a$
 $b' = b - \frac{aa^T}{a^T a} b$
 $c' = c - \frac{aa^T}{a^T a} c - \frac{bb^T}{b^T b} c$
From c subtract projection of c on a and projection of c on b .

4. Determinant & Trace

4.1. Properties of determinants

Three most basic properties of determinants

1. $\det I = 1$.
2. Determinant changes sign when two rows are exchanged.
3. Determinant depend linearly on the first row.
Eg:

$$\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$$

$$\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

Derivative properties of determinants

4. $|cA| = c^n |A|$.
5. If two rows of A are equal, then $\det A = 0$
Follows from rule 2.
6. Adding a multiple of one row to another does not alter the value of the matrix.
Follows from rule 3 and rule 4.
7. If A has a row of 0s, then $\det A = 0$.
Follows from rule 4 and rule 5.
8. If A is triangular, $\det A = \prod d_i$.
Proof: Convert A to an equivalent diagonal determinant by row operation. Now apply rule 3 and finally rule 1.
9. If A is singular, then $\det A = 0$. If A is invertible, then $\det A \neq 0$.
Proof:
If A is singular, elimination leads to zero row, $\therefore \det A = 0$
If A is invertible, elimination leads to non-zero pivots.
 $\det A = \prod \text{pivots} \neq 0$.
10. $\det(AB) = \det(A) \times \det(B)$
Proof:
For a diagonal matrix D , $\det(DB) = \det(D)\det(B)$
 A can be converted to a diagonal matrix D by row operations.
Using the exact same row operation as above AB can be converted to DB
For another proof see book.
11. $\det(A^T) = \det(A)$.
Proof:
 $PA = LDU$ and $\det(PA) = \det(LDU)$
 $\det(A^T P^T) = \det(U^T D^T L^T) = \det(LDU) = \det(PA)$
 $\det(P^T) = \det(P) \therefore \det(A^T) = \det(A)$.

4.2. Formulae for determinant

Based on LDU decomposition

If A is invertible $PA = LDU$.
 $\det(A) = \pm \det(L)\det(D)\det(U) = \pm(\text{product of pivots})$

Using linearity of determinants

Eg: 2×2 case.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix} = ad - bc$$

Expansion in cofactors

The determinant of A is any row i times its cofactors

That is, $\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \dots + a_{in}C_{in}$

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

Proof:

Split the $n \times n$ matrix to sum of $n \times n$ matrices by linearity property of determinants. Each of the constituent matrices can be converted to LDU form.

Block matrices

- $\det \begin{bmatrix} A & 0 \\ C & D \end{bmatrix} = \det(A)\det(D) = \det \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$
- If A is invertible
 $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B)$
- If blocks are square matrices of the same size. And $AC = CA$.
 $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(AD - CB)$
- If blocks are square matrices of same size and if $A = D$ and $B = C$. Here A and B need not commute.
 $\det \begin{bmatrix} A & B \\ B & A \end{bmatrix} = \det(A - B)\det(A + B)$

4.3. Application of determinants

Definition 4.1 (Adjugate Matrix):

Transpose of cofactor matrix: $\text{adj}(A) = C^T$. **NOTE:**
 $|\text{adj}(A)| = |A|^{n-1}$.

Computation of A^{-1}

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

C is cofactor matrix.

Proof:

$$A^{-1} = \frac{C^T}{\det(A)} \implies \det(A)I = AC^T$$

Now the diagonal of AC^T will be $\det(A)$. To complete the proof we just need to show that the off-diagonal elements are all 0.

Each off-diagonal element is represented as:

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \dots + a_{in}C_{jn}$$

where $i \neq j$

This is equal to the determinant of modified A where the j 'th row is replaced by a copy of i 'th row times ± 1 . This determinant = 0

Cramer's rule; Solution to $Ax = b$

$$x_j = \frac{\det(B_j)}{\det A}$$

Where, B_j is A with it's j 'th column replaced by b .

Proof: $x = A^{-1}b = \frac{1}{\det(A)} C^T b$

Volume of a box

$\det(A)$ represents the volume of a box (higher dimensional parallelogram) whose edges are represented by the row vectors of A (or the column vectors of A).

4.4. Trace of a Matrix

Basic Properties

- **Linearity:**
 - $tr(A + B) = tr(A) + tr(B)$.
 - $tr(cA) = c \times tr(A)$
- **Cyclic properties:**
 - $tr(AB) = tr(BA)$.
 - $tr(ABC) = tr(BCA) = tr(CAB)$.
- If $B = P^{-1}AP$, then $tr(A) = tr(B)$.

5. Eigen values and eigen vectors

5.1. Introduction

Eigenvalues & Eigenvectors: The scalars λ and the vectors x that satisfy $Ax = \lambda x$ are the eigenvalues and eigenvectors of the matrix A .
 $Ax = \lambda x \implies (A - \lambda I)x = 0 \implies |A - \lambda I| = 0$
 Setting λ as unknown gives rise to a n dimensional polynomial equation in λ called the (**Characteristic equation**).

Solving this gives eigen values.

Algebraic multiplicity(M_λ): is the number of times λ occurs as a root in the characteristic equation.

Geometric multiplicity(m_λ): Dimension of the eigenspace corresponding to the eigenvalue λ .

Key Results:

- $\sum \lambda_i = tr(A)$
- $\prod \lambda_i = det(A)$
- For a triangular matrix the Eigen values are along the diagonal.
- If $Ax = \lambda x$, then $A^k x = \lambda^k x$.
- Eigen vectors corresponding to each Eigen value form a vector space.
- $m_\lambda \leq M_\lambda$

TODO: Proof outlines.

5.2. Diagonalization

If a $n \times n$ matrix A has n independent eigen vectors then, it can be written as follows:

$$A = S\Lambda S^{-1}$$

where, each column of S is one of the eigenvectors of A (e_i)

Caley-Hamilton's theorem.

Theorem 5.1:

A satisfies it's own characteristic equation.

Proof.

Substitute $A = S\Lambda S^{-1}$ in $(A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_n I)$. \square

TODO: Proof for case where A is not diagonalizable.

5.3. Difference equations and powers A^k

Difference equations

$$u_{k+1} = Au_k$$

In some sense difference equations are analogous to differential equations.

Examples

Fibonacci numbers

$$F_{k+2} = F_{k+1} + F_k$$

Let, $u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$ then, difference representation of Fibonacci sequence would be:

$$u_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = Au_k$$

The solution to this difference equation, $u_{k+1} = Au_k$ is

$$u_k = A^k u_0$$

If A can be diagonalized then $u_k = S\Lambda^k S^{-1}u_0$.

Markov matrices

Markov matrix is similar to the above examples along with the following two properties:

1. Each column of Markov matrix adds to 1.
2. The numbers outside and inside can never become negative.

$u_{k+1} = Au_k$ and the solution is $u_k = A^k u_0$

If A can be diagonalized, then $u_k = S\Lambda^k S^{-1}u_0$. *Steady state*(u_∞):
 $Au_\infty = u_\infty$.

Properties of a Markov matrix A :

1. $\lambda_1 = 1$ is an eigenvalue of A .
2. Its eigenvector x_1 is nonnegative and $Ax_1 = x_1$.
3. The other eigenvalues satisfy $|\lambda_i| \leq 1$.
4. If A or any power of a A has all positive entries, these other $|\lambda_i|$ are below 1.
5. The solution $A^k u_0$ approaches a multiple of x_1 which is the steady state u_∞ .

Stability of $u_{k+1} = Au_k$

The difference equation $u_{k+1} = Au_k$ is:

- *stable* if all eigenvalues satisfy $|\lambda_i| < 1$.
- *neutrally stable* if some $|\lambda_i| = 1$ and all other $|\lambda_i| < 1$.
- *unstable* if at least one $|\lambda_i| > 1$.

5.4. Differential equations and e^{At}

System of differential equations

Matrices are useful to solve a system of differential equations.
 Eg:

$$\begin{aligned} \frac{du_1}{dt} &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\ \frac{du_2}{dt} &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\ &\vdots \\ \frac{du_n}{dt} &= a_{n1}u_1 + a_{n2}u_2 + \cdots + a_{nn}u_n \end{aligned}$$

This system in matrix notation is $\frac{du}{dt} = Au$, where

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

Solution:

$$u(t) = e^{At}u(0)$$

Is the above solution valid if A is not diagonalizable.

Proof: If A is diagonalizable then let $v = S^{-1}u$, where S is the eigenvector matrix of A . Using eigenvector matrix the system of linear equations is decoupled.

$$\begin{aligned} \therefore S \frac{dv}{dt} &= ASv \\ \implies \frac{dv}{dt} &= S^{-1}ASv = \Lambda v \\ v(t) &= e^{\Lambda t}v(0) \\ u(t) &= Se^{\Lambda t}S^{-1}u(0) = e^{At}u(0) \end{aligned}$$

where e^{At} is the exponential of the matrix A .

Matrix exponential

Matrix exponential is defined as follows:

$$e^{At} = I + At + \frac{(At)^2}{2} + \cdots + \frac{(At)^n}{n!} + \cdots$$

This series **always converges** and has the following properties:

- $(e^{As})(e^{At}) = e^{A(s+t)}$
- $(e^{At})(e^{-At}) = I$
- $\frac{d}{dt}(e^{At}) = Ae^{At}$
- If A can be diagonalized, $A = SAS^{-1}$, then $du/dt = Au$ has the solution $u(t) = e^{At}u(0)$.
- e^{At} is never singular.

Proof 1: If λ is an eigenvalue of A , then $e^{\lambda t}$ is the corresponding eigenvalue of e^{At} , which is never 0.

Proof 2: $\det(e^{At}) = e^{\lambda_1 t} e^{\lambda_2 t} \cdots e^{\lambda_n t} = e^{\text{trace}(At)} \neq 0$.

Stability of differential equations

Behaviour of $u(t)$ as $t \rightarrow \infty$: The differential equation $du/dt = Au$ is:

- **stable:** and $e^{At} \rightarrow 0$ whenever, all $\text{Re}(\lambda_i) < 0$.
- **neutrally stable:** when all $\text{Re}(\lambda_i) < 0$ and $\text{Re}(\lambda_1) = 0$.
- **unstable:** and e^{At} is unbounded if any eigenvalue has $\text{Re}(\lambda_i) > 0$.

These results are true even if A is not diagonalizable.

5.5. Complex matrices

Length of a vector: $\|x\| = |x_1| + |x_2| + \cdots + |x_n|$

Hermitian (and symmetric) matrices

Conjugate transpose: $\overline{A}^T = A^H$ (Read as A Hermitian) has entries $(A^H)_{ij} = \overline{A_{ji}}$

Hermitian matrices: A is Hermitian iff $A^H = A$.

Theorem: If $A = A^H$ then $\forall x, x^H Ax$ is real.

Proof: $(x^H Ax)^H = x^H A^H x = x^H Ax$.

Theorem: If $A = A^H$ then all the eigenvalues are real.

Proof: $x^H Ax = \lambda x^H x$. LHS is real and $x^H x$ is real, therefore λ is real.

Theorem: Eigenvectors of Hermitian matrices that come from two different eigenvalues are orthogonal.

Proof:

Let, $Ax = \lambda x$ and $Ay = \mu y$

$$\lambda x^H y = (Ax)^H y = x^H Ay = x^H \mu y \implies x^H y = 0, \because \lambda \neq \mu$$

Orthonormal matrices: $QQ^T = I$.

Theorem (Spectral theorem): A real symmetric matrix can be factored into $A = Q\Lambda Q^T$. Its orthonormal eigenvectors are in the columns of Q .

Unitary matrices

$$U^H U = I$$

Multiplication by U has no effect on inner products, angles and lengths.

Property 1:

- $(Ux)^T(Uy) = x^T U^T U y = x^T y$
- $\|Ux\|^2 = x^H U^H U x = x^H x = \|x\|^2$

Property 2:

For every eigenvalue(λ) of U , $|\lambda| = 1$

Property 3:

Eigen vectors corresponding different eigenvalues are orthogonal.

Let, $Ux = \lambda x$ and $Uy = \mu y$.

$$x^H y = (Ux)^H (Uy) = (\lambda x)^H (\mu y) = \bar{\lambda} \mu x^H y$$

By property-2 $\bar{\lambda} \lambda = 1, \therefore \bar{\lambda} \mu \neq 1$

$$\implies x^H y = 0.$$

If A is Hermitian then $K = iA$ is skew-Hermitian.

Theorem: Eigenvalues of K are purely imaginary.

Fourier matrix

THE FOLLOWING IS INCOMPLETE

$$F_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)^2} \end{bmatrix}$$

5.6. Similarity transformation

Definition: A and B are similar if $\exists M \ni B = M^{-1}AM$.

Going from one to another is known as **similarity transformations**.

In the case where $M = S$, $M^{-1}AM$ becomes the diagonal matrix Λ . This is the best case scenario. But other M 's are also useful. Usually M is chosen such that, $M^{-1}AM$ is easier to work with than A .

Theorem: Suppose $B = M^{-1}AM$, then A and B have the same eigenvalues.

Proof: $Ax = \lambda x \implies B(M^{-1}x) = \lambda(M^{-1}x)$

$M^{-1}x$ is the eigenvector of B corresponding to λ .

Change of Basis = Similarity transformation

Every linear transformation is represented by a matrix. Similar matrices represent the same transformation T with respect to different basis.

$$\begin{array}{ccc} [T]_{V \text{ to } V} & = [I]_{V \text{ to } V} & [T]_{V \text{ to } V} \\ B & = M^{-1} & A \end{array} \quad \begin{array}{ccc} [I]_{V \text{ to } V} & & [I]_{V \text{ to } V} \\ & & M \end{array}$$

Triangular forms with unitary M

Shcur's lemma:

There is a unitary matrix $M = U$ such that $U^{-1}AU = T$ is triangular. The eigenvalues of A appear along the diagonal of this similar matrix T .

Diagonalizing Hermitian matrices

Theorem:

If $A = A^H$ This triangular form T , is a diagonal matrix.

Proof: $(U^{-1}AU)^H = U^H A^H (U^{-1})^H = U^{-1}AU \implies T = T^H$.

Spectral theorem:

Every real symmetric matrix can be diagonalized by an orthogonal matrix Q . Every Hermitian matrix can be diagonalized by a unitary matrix U :

$$\begin{array}{ccc} \text{(real)} & Q^{-1}AQ = \Lambda & \text{or} & A = Q\Lambda Q^T \\ \text{(complex)} & U^{-1}AU = \Lambda & \text{or} & A = U\Lambda U^H \end{array}$$

Normal matrices

The matrix N is normal if it commutes with N^H : $NN^H = N^H N$.

Theorem:

A triangular matrix T that is normal must be diagonal

Proof:

Tip: Use induction and multiplication by blocks.

For details see: <https://math.stackexchange.com/a/2538528/633346>

Theorem:

$T = U^{-1}AU$ is diagonal if and only if A is a normal matrix.

Proof:

If A is normal,

$$\begin{aligned} TT^H &= U^{-1}AUU^H A^H U = U^{-1}AA^H U = U^{-1}A^H AU = \\ U^H A^H U U^{-1}AU &= T^H T \implies T \text{ is diagonal.} \end{aligned}$$

Jordan form

Theorem: If A has s independent eigenvectors, it is similar to a matrix with s blocks.

$$\text{Jordan form } J = M^{-1}AM = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix}$$

Each Jordan block is a triangular matrix that has only a single eigenvalue (λ_i) along the diagonal corresponding to only one eigenvector. For each missing eigenvector there will be a 1 just above the diagonal.

$$\text{Jordan block } J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

6. Positive definite matrices

Definition:

A $n \times n$ symmetric matrix A is positive-definite if $\forall x \neq 0, x \in \mathbb{R}^n$ $x^T A x > 0$.

Test for positive definiteness Each of the following test is a necessary and sufficient condition for a matrix to be positive definite.

1. $x^T A x > 0$ for all nonzero vectors x .
Definition.

2. All eigenvalues of A satisfy $\lambda_i > 0$.

Proof: $x^T A x = \lambda x^T x = \lambda \|x\|^2$

Converse: Every $\forall y > 0, y \in \mathbb{R}^n$, $y = a_1 x_1 + \dots + a_n x_n$

$$A y = a_1 \lambda_1 + \dots + a_n \lambda_n$$

$$y^T A y = a_1^2 \lambda_1 + \dots + a_n^2 \lambda_n > 0$$

3. All the upper left submatrices A_k have positive determinants.

$$\text{Clue: } x^T A x = \begin{bmatrix} x_k^T & 0 \end{bmatrix} \begin{bmatrix} A_k & * \\ * & * \end{bmatrix} \begin{bmatrix} x_k \\ 0 \end{bmatrix} = x_k^T A_k x_k$$

4. All the pivots (without row exchanges) satisfy $d_k > 0$.

Proof: Incomplete

Theorem: Another test for positive definiteness

The symmetric matrix A is positive definite if and only if:

$\exists R$ with independent columns $\ni A = R^T R$.

Positive semidefinite Matrices

The tests for semidefinite matrices will relax to allow zeros.

Definition:

$\forall x \ x^T A x \geq 0$

Test for positive semi definiteness

- $\forall x \neq 0 \ x^T A x \geq 0$.
- All eigenvalues satisfy $\lambda_i \geq 0$.
- No principal submatrices of A have negative determinants.
- No pivots are negative.
- $\exists R$, possibly with dependent columns such that $A = R^T R$.

6.1. Singular value decomposition

Any $m \times n$ matrix A can be decomposed in to:

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

Σ : Diagonal matrix, $m \times n$. This diagonal matrix has eigenvalues from $A^T A$, not from A .

The positive entries $\sigma_1, \sigma_2, \dots, \sigma_s$, form the first r diagonal elements of Σ . These elements are called **singular values**. The remainder of entries in Σ is 0.

U : Orthogonal matrix, $m \times m$.

The columns of U are the eigenvectors of $A^T A$. V : Orthogonal

matrix, $n \times n$.

The columns of V are the eigenvectors of $A A^T$.

Remark 1

For positive definite matrices, Σ is Λ and $U \Sigma V^T$ is identical to $Q \Lambda Q^T$.

Remark 2

U and V give orthonormal bases for all four fundamental subspaces.

first	r	columns of U :	column space of A
last	$m - r$	columns of U :	left nullspace of A
first	r	columns of V :	row space of A
last	$n - r$	columns of V :	nullspace of A

Remark 3

$A v_j = \sigma_j u_j$ or $AV = U \Sigma$

Remark 4

Eigenvectors of $A A^T$ and $A^T A$ goes into columns of U and V .

$$A A^T = (U \Sigma V^T)(V \Sigma^T U^T) = U \Sigma \Sigma^T U^T, \text{ and,}$$

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

7. Matrix Factorizations

LU factorizability

$$A = LU$$

Requirements:

An invertible matrix A is LU factorizable, if the determinant of all its **principal minors** are > 0 . Or if A has rank k , A is LU factorizable, if all its first k principal minors are > 0 .

Procedure:

Do Gaussian elimination without row exchanges.

LDU:

Pivots in U are divided out into D to have 1s in the diagonal of U .

PLU factorization

$$PA = LU$$

Always exists.

Procedure:

Do Gaussian elimination.

If A is invertible, then L and U are invertible. P is always invertible and $P^{-1} = P^T$.