Linear Algebra 1 Notes (2023/2024)

Griffin Reimerink

Contents

1 Systems of linear equations

An $m \times n$ **system of linear equations** is a collection of m equations in n unknowns.
By a **solution** of an $m \times n$ system, we mean n numbers $x_1, x_2 ... x_n$ that satisfy all m equations.
A system with no solutions is **inconsistent**, a system with at least 1 solution is **consistent**.

The set of all solutions is called the **solution set**.
A system can only have 0, 1 or infinitely many solutions.
Two systems involving the same number of unknowns are said to be **equivalent** if they have the same solution set.

The coefficients of a system of linear equations can be put into a $m \times n$ matrix.
Left part of a system is the **coefficient matrix**, right part is the **right hand side vector**, the combination of those two is the **augmented matrix**.

If the right hand side vector only contains zeroes, the system is **homogeneous**.
A homogeneous system always has a trivial solution where every variable is 0.
An $m \times n$ homogeneous system has a nontrivial solution if $n > m$.

1.1 Elementary row operations

Elementary row operations:

1. Interchange two rows
2. Multiply a row by a nonzero scalar
3. Replace a row by its sum with a scalar multiple of another row.

Elementary rows on the augmented matrix would not change the solution set and hence create equivalent systems.
A **zero row** is a row that contains only zero entries.
A **nonzero row** is a row that contains at least one nonzero entry.
The **leading nonzero entry** of a nonzero row is the first (from left to right) nonzero entry.
The **leading zero entries** of a nonzero row are the zero entries before the leading nonzero entry.

A matrix is said to be in **row echelon form** if it satisfies the following:

1. The rows below a zero row (if any) are all zero rows.
2. The leading nonzero entry of each nonzero row is 1.
3. The number of leading zero entries on nonzero rows increases from top to bottom.

Every matrix can be put into row echelon form by using the three elementary row operations.
Example:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 3 & 6 & 10 & 8 \\ 0 & 2 & 4 & 6 \end{array}\right] ② \leftarrow ② - 3 * ①$$
$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -4 \\ 0 & 2 & 4 & 6 \end{array}\right] ② \leftrightarrow ③$$
$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 1 & -4 \end{array}\right] ② \leftarrow ② * \frac{1}{2}$$
$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & -4 \end{array}\right]$$

The unknowns corresponding to the leading nonzero entries in each row of the row echelon form will be called **lead variables**, all other unknowns are called **free variables** and can have any value.
The system in row echelon form can be solved with back substitution.
A given system of linear equations is **consistent** iff there is no leading 1 in the last column of the augmented matrix in row echelon form.
A matrix is said to be in **reduced row echelon form** (does not require back substitution to solve) if:

1. It is in row echelon form
2. The leading 1 in each nonzero row is the only nonzero entry in its column.

How to solve a system of linear equations:

1. Given linear system of equations
2. Augmented matrix
3. Row echelon form (with elementary row operations)
4. Consistent/inconsistent
5. Reduced row echelon form

2 Matrices

2.1 Matrix arithmetic

Notation: matrix A, scalar α, vector $\boldsymbol{a}$ or $\overrightarrow{a}$

In general $AB \neq BA$, even if AB is defined BA can be undefined.

- $A + B = B + A$
- $(A + B) + C = A + (B + C)$
- $(AB)C = A(BC)$
- $A(B + C) = AB + AC$
- $(A + B)C = AC + BC$
- $(\alpha\beta)A = \alpha(\beta A)$

If A is an $m \times n$ matrix:
The **transpose** of A is the $n \times m$ matrix A^T defined by $[A^T]_{ij} = [A]_{ji}$.

- $(A^T)^T = A$
- $(\alpha A)^T = \alpha A^T$
- $(A + B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$

Matrices A and B can be divided by solving $Ax = B$

2.2 Matrix types

A matrix whose entries are all zero is called a **zero matrix**.
We denote the $m \times n$ zero matrix by $0_{m,n}$.
The $n \times n$ **identity matrix** is the matrix $I_n = (\delta_{ij})$ where

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise.} \end{cases}$$

A matrix A is said to be

- **square** if it has the same number of rows and columns.
- **symmetric** if it is square and $A^T = A$.
- **skew-symmetric** if it is square and $A^T = -A$.
- **upper triangular** if $[A]_{ij} = 0$ for $i > j$.
- **lower triangular** if $[A]_{ij} = 0$ for $i < j$.
- **triangular** if it is either upper triangular or lower triangular.
- **diagonal** if it is both upper and lower triangular.
- **strictly upper (lower) triangular** if it is upper (lower) triangular and every diagonal entry is zero.
- **nonsingular** or **invertible** if it is square and there exists an $n \times n$ matrix B (the **inverse** of A) such that $AB = BA = I_n$
- **singular** if it does not have an inverse.
- **involutory** if $A^2 = I$
- **idempotent** if $A^2 = A$
- **elementary** if it can be obtained from an identity matrix by performing a single elementary row operation.

2.3 Properties of matrices

$I_n A = A$ for all $A \in F^{n\times p}$ and $BI_n = B$ for all $B \in F^{m\times n}$
$I^{-1} = I$ and $(A^{-1})^{-1} = A$
If B and C are both inverses of A, then $B = BI = B(AC) = (BA)C = IC = C$
If A and B are nonsingular $n \times n$ matrices, then AB is also nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$
The transpose of a lower triangular matrix is upper triangular and vice versa.

A matrix that is both symmetric and triangular must be diagonal.
Diagonal matrix multiplication is commutative.
A matrix that is both skew-symmetric and triangular must be the zero matrix.
The product of two lower (upper) triangular matrices is lower (upper) triangular.
We can **partition** a matrix into **submatrices** with horizontal or vertical lines.

Let A be an $m\times n$ matrix. Also, let A' and the elementary matrix E be obtained, respectively, from A and I_m by applying the same singular elementary row operation. Then, $A' = EA$.

If E is an elementary matrix corresponding to a row operation and F is the elementary matrix corresponding to the inverse row operation, then $FEA = A$ for all A

Every elementary matrix is nonsingular. Moreover, the inverse of an elementary matrix corresponding to a row operation is the elementary matrix corresponding to the inverse of that row operation and hence is of the same type.

Inverse row operations

type	operation	inverse
I	$ⓘ \leftrightarrow ⓙ$	$ⓘ \leftrightarrow ⓙ$
II	$ⓘ \leftarrow c * ⓘ$	$ⓘ \leftarrow \frac{1}{c} * ⓘ$
III	$ⓘ \leftarrow ⓘ + c * ⓙ$	$ⓘ \leftarrow ⓘ - c * ⓙ$

Let A be a square matrix. Then the following conditions are equivalent:

1. A is nonsingular.
2. LE$(A, 0)$ has only the trivial solution.
3. The reduced row echelon form of A is equal to I.
4. A is a product of elementary matrices.

Let A, B be $n \times n$ matrices. If $AB = I$, then $BA = I$. Moreover, $A^{-1} = B$ and $B^{-1} = A$

To find an inverse, take the matrix $\begin{bmatrix} A \mid I \end{bmatrix}$ and apply elementary row operations until the identity matrix is on the left. The matrix on the right is the inverse.

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

1. A is nonsingular.
2. LE(A, b) has a unique solution for every n-vector b.
3. LE(A, b) has a unique solution for some n-vector b.

If a square matrix A can be reduced to upper triangular form by using only row operation III,
then it can be written as a product of a lower and an upper triangular matrix.
Such a factorization is called **LU factorization**.

3 Determinants

Let $N \in \mathbb{N}$. A **permutation** of n distinct objects is a reordering of these objects.
We can label these objects by numbers $\{1, 2...n\}$ and consider a permutation as a bijective function whose domain and codomain are both the set $\{1, 2...n\}$.
The set of all permutations is called the **symmetric group of degree** n
and is denoted by S_n with cardinality $n!$.

Let $\sigma \in S_n$. We can say that (i, j) is an **inversion pair** of σ if $i < j$ and $\sigma(i) > \sigma(j)$
The **sign of a permutation** σ is defined by

$$\operatorname{sign}\sigma = (-1)^{\text{the \# of inversion pairs of } \sigma} = \begin{cases} 1 & \text{if the \# of inversion pairs of } \sigma \text{ is even} \\ -1 & \text{if the \# of inversion pairs of } \sigma \text{ is odd} \end{cases}$$

The **determinant** of A is defined by

$$\det A = \sum_{\sigma \in S_n} \left((\operatorname{sign}\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \right)$$

Let A be an $n \times n$ matrix. We denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting the ith and jth column by A_{ij}. The scalar $\det A_{ij}$ is the (i, j) **minor** of A and the minor multiplied by $(-1)^{i+j}$ is the (i, j) **cofactor** of A.

Let A be a $n \times n$ square matrix. Then, the following statements hold:

1. $\det A^T = \det\ A$
2. If A has a zero row, then $\det A = 0$
3. Suppose that $n \geq 2$. If A has two identical rows, then $\det A = 0$.
4. If A is a triangular matrix, then $\det A = \prod_{i=1}^{n} a_{ii}$
5. Iff $\det A \neq 0$, A is nonsingular. Iff $\det A = 0$, A is singular.

Suppose that $n \geq 2$. Let A be an $n \times n$ matrix. Let $i, j \in \{1, 2, ..., n\}$ such that $i \neq j$. Then

$$\sum_{k=1}^{n} (-1)^{i+k} a_{jk} \det A_{ik} = 0$$

If A and B are both $n \times n$ matrices, $\det(AB) = \det(A)\det(B)$ and $\det(cA) = c^n \det A$

3.1 Calculating a determinant

Cofactor expansion along row i or along column j:
Let A be an $n \times n$ matrix. if $n = 1$, then $\det(A) = a_{11}$. If $n \geq 2$, then

$$\det A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik} = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj} \qquad \forall i, j \in \{1, 2, ..., n\}$$

$\det I = 1$
Determinants of 2×2 and 3×3 matrices:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \qquad \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - gec - hfa - idb$$

Suppose that $n \geq 2$ and A is an $n \times n$ matrix.
Let E be the elementary matrix corresponding to the row operation $(i) \leftrightarrow (j)$.
Then, $\det(EA) = -\det A$ and $\det E = -1$.
Let E be the elementary matrix corresponding to the row operation $(i) \leftarrow c * (i)$.
Then, $\det(EA) = c \det A$ and $\det E = c$.
Let E be the elementary matrix corresponding to the row operation $(i) \leftarrow (i) + c * (j)$.
Then, $\det(EA) = \det A$ and $\det E = 1$.

Every square matrix A can be transformed to row echelon form, that is $R = E_k E_{k-1} ... E_1 A$ where R is in row echelon form and E_i's are all elementary matrices.

- If the last row of R is zero, then $\det(A) = 0$
- Otherwise, A is nonsingular and $\det(A) = [\det(E_k)\det(E_{k-1}) ... \det(E_1)]^{-1}$

Because $\det A = \det A^T$, we can also use **elementary column operations.**

For calculating a determinant, elementary row operations are the fastest method (except for 2×2 and 3×3), while cofactor expansion is useful for proofs.

3.2 Adjoint matrix

If $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, then the **adjoint** of A is given by

$$\text{adj}(A) = \begin{bmatrix} \text{Cof}(a_{11}) & \text{Cof}(a_{12}) & \cdots & \text{Cof}(a_{1n}) \\ \text{Cof}(a_{21}) & \text{Cof}(a_{22}) & \cdots & \text{Cof}(a_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cof}(a_{n1}) & \text{Cof}(a_{n2}) & \cdots & \text{Cof}(a_{nn}) \end{bmatrix}$$

where each cofactor $\text{Cof}(a_{ij})$ can be expressed as the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the i-th row and j-th column, multiplied by $(-1)^{i+j}$.

$A(\text{adj}\, A) = \det A * I$. If $\det A \neq 0$, $A^{-1} = \dfrac{1}{\det A} \text{adj}\, A$

Suppose that $A \in \mathbb{F}^{n \times n}$ is nonsingular. Let $b \in \mathbb{F}^n$ and A_i be the matrix obtained from A by replacing the ith column by b. Then, x given by $x_i = \dfrac{\det A_i}{\det A}$ for $i \in \{1, 2, \ldots, n\}$ is the unique solution of LE(a, b)

4 Eigenvalues and eigenvectors

Population dynamics can be expressed as $A^t x$ where A is a matrix describing the population changes, t is the time and x is a vector with the situation at $t = 0$.

Let A be a square matrix.
A scalar λ is said to be an **eigenvalue** of A if there is a nonzero vector x such that $Ax = \lambda x$.
The vector x is said to be an **eigenvector** corresponding to λ. We call the pair (λ, x) an **eigenpair**.
$Ax = \lambda x$ is a nonlinear equation in unknowns λ and x. It can also be written as $(\lambda I - A)x = 0$.

λ is an eigenvalue of A if and only if:

1. $\text{LE}(A - \lambda I, 0)$ has a nontrivial solution
2. $A - \lambda I$ is singular
3. $\det(\lambda I - A) = 0$

If A is an $n \times n$ matrix, then
$\det(\lambda I - A) = \lambda^n + p_{n-1}\lambda^{n-1} + \ldots + p_1\lambda + p_0$ where p_k is a scalar for each $k \in \{0, 1, \ldots, n-1\}$.

A variation of the **fundamental theorem of algebra**:
Let $n \in \mathbb{N}$. Every polynomial of the form $\lambda^n + p_{n-1}\lambda^{n-1} + \ldots + p_1\lambda + p_0$
where $p_k \in \mathbb{F}$ for each $k \in \{0, 1, \ldots, n-1\}$ has n complex roots.

The polynomial $\det(\lambda I - A)$ is called the **characteristic polynomial** of A: $p_A(\lambda)$.
The equation $p_A(\lambda) = 0$ is called the **characteristic equation** of A.
Let (λ, x) be an eigenpair of a real matrix. Then, $(\overline{\lambda}, \overline{x})$ is also an eigenpair of the same matrix.
Eigenvalues and eigenvectors can be complex.

How to find eigenvectors:

1. Calculate the polynomial $\det(\lambda I - A)$.
2. The roots of the polynomial are the eigenvalues λ.
3. For each eigenvalue, solve $(\lambda I - A)x = 0$.

The **trace** tr A is equal to the sum of its diagonal entries.
tr A is equal to the sum of its eigenvalues and det A is equal to the product of its eigenvalues.

4.1 Linear difference equations

A **linear difference equation** is an equation of the form $x(k+1) = Ax(k)$
where $k \in \{0, 1, 2, \ldots\}, x(k) \in \mathbb{F}^n$, and $a \in \mathbb{F}^{n \times n}$. Given $x(0) = x_0$, $x(k) \in A^k x_0$ for all $k \in \mathbb{N}$.
Let A be an $n \times n$ matrix and let (λ, x) be an eigenpair of A. Then, $A^k x = \lambda^k x$ for all $k \in \mathbb{N}$.

5 Diagonalizability

Let A and B be $n \times n$ matrices.
We say that B is **similar** to A if there exists a nonsingular S such that $B = SAS^{-1}$.
If B is similar to A, A is similar to A. Then, we can simply say that A and B are similar.
Similar matrices have the same characteristic polynomial and eigenvalues.
A square matrix is **diagonalizable** if it is similar to a diagonal matrix. ($A = XDX^{-1}$)

Let $k \in \mathbb{N}$ and let $x_1, x_2, \ldots, x_k$ be vectors in $\mathbb{F}^n$.
We say that $x_1, x_2, \ldots, x_k$ are **linearly independent** if

$$c_1x_1 + c_2x_2 + \ldots + c_kx_k = 0 \implies c_1 = c_2 = \ldots = c_k = 0$$

Otherwise, we say that $x_1, x_2, \ldots, x_k$ are **linearly dependent**.

A square matrix is nonsingular if and only if its columns are linearly independent.
An $n \times n$ matrix is diagonalizable if and only if it has n linearly independent eigenvectors.

If a is diagonalizable, that is $AX = XD$ for a nonsingular matrix X and a diagonal matrix D, then column vectors of X are eigenvectors of A and diagonal values of D are eigenvalues of A

Let $A \in \mathbb{F}^{n\times n}$ and let $k \in \mathbb{N}$ with $k \leq n$. If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues ($\lambda_i \neq \lambda_j$ for $i \neq j$) of A with corresponding eigenvectors $x_1, x_2, \ldots, x_k$, then $x_1, x_2, \ldots, x_k$ are linearly independent.
Every $n \times n$ matrix that has n distinct eigenvalues is diagonalizable.
If A is diagonalizable, $A^k = XD^kX^{-1}$. Diagonal matrices behave like a scalar when multiplying.

5.1 Linear differential equations

A system of **linear differential equations** is of the form $x'(t) = Ax(t)$ where $x : \mathbb{R} \to \mathbb{F}^n$ is a vector-valued function and A is an $n \times n$ matrix.
An **initial value problem** amounts to finding a solution to $x'(t) = Ax(t) \quad x(0) = x_0$
Let A be a square matrix. We define its **exponential** by $\sum_{k=0}^{\infty} \frac{A^k}{k!}$.

$$(e^{tA})' = Ae^{tA} \qquad x(t) = e^{tA}x_0 \implies x'(t) = Ax(t) \text{ and } x(0) = x_0$$

If A is diagonalizable, $e^{tA} = Xe^{tD}X^{-1}$

6 Subspaces

Let S be a subset of $\mathbb{F}^n$. We say that S is a **subspace** of $\mathbb{F}^n$ if:

1. S is nonempty.
2. S is closed under scalar multiplication: $\boldsymbol{x} \in S \implies c\boldsymbol{x} \in S$ for all $c \in \mathbb{F}$.
3. S is closed under vector addition: $\boldsymbol{x}, \boldsymbol{y} \in S \implies \boldsymbol{x} + \boldsymbol{y} \in S$.

$\{\mathbf{0}_n\}$ (**zero subspace**) and $\mathbb{F}^n$ are trivial subspaces of $\mathbb{F}^n$. Every subspace of $\mathbb{F}^n$ contains $\{\mathbf{0}_n\}$.
Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \in \mathbb{F}^n$. Their **span** is defined as $\{c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_k\boldsymbol{v}_k : c_i \in \mathbb{F} \text{ for } i \in \{1, 2, \ldots, k\}\}$
$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_k\boldsymbol{v}_k$ is called a **linear combination** of the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$

Let $A \in \mathbb{F}^{m\times n}$. The **range** of A is defined by $R(A) := \{Ax : \boldsymbol{x} \in \mathbb{F}^n\}$
and the **null space** is defined by $N(A) := \{\boldsymbol{x} \in \mathbb{F}^n : A\boldsymbol{x} = \mathbf{0}_m\}$.
$R(A)$ is a subspace of $\mathbb{F}^m$ and $N(A)$ is a subspace of $\mathbb{F}^n$.
The subspace $N(A - \lambda I)$ is called the **eigenspace** of $A \in \mathbb{F}^{n\times n}$ corresponding to the eigenvalue λ.

6.1 Basis of a subspace

Let $S \subseteq \mathbb{F}^n$ be a subspace and $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \in S$. Then $(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k)$ is a **basis** for S if

1. $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are linearly independent
2. $\text{span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k) = S$

This implies that for every vector $\boldsymbol{v}$ in S, $\boldsymbol{v} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_k\boldsymbol{v}_k$ with unique a_i.

Let $\boldsymbol{e}_i$ denote the ith column of I_n. Then, $(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots \boldsymbol{e}_n)$ is a basis for $\mathbb{F}^n$.
Suppose that $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_k \in \mathbb{F}^n$ and $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots \boldsymbol{y}_l \in \mathbb{F}^n$ such that span $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_k) \subseteq (\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots \boldsymbol{y}_l)$.
If $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_k$ are linearly independent, $k \leq l$.

Let $\boldsymbol{x}_1, \dots \boldsymbol{x}_k \in \mathbb{F}^n$ be linearly independent. Then $k \leq n$ and $\boldsymbol{x}_1, \dots \boldsymbol{x}_k$ is a basis for $\mathbb{F}^n$.
Let $S \neq \{\mathbf{0}_n\}$ be a subspace of $\mathbb{F}^n$ and let $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots \boldsymbol{v}_k \in S$ be linearly independent vectors s.t. $\text{span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \dots \boldsymbol{v}_k) \neq S$. Then, there exists $\boldsymbol{v}_{k+1} \in S$ s.t. $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots \boldsymbol{v}_{k+1}$ are linearly independent.
Every nonzero subspace has a basis. Every two bases for a given subspace have equal cardinality.
Every set of lin. ind. vectors in a nonzero subspace can be completed into a basis of that subspace.
Let $S \subseteq \mathbb{F}^n$ be a subspace. If $S = \{\mathbf{0}_n\}$, then its **dimension** is 0.
If $S \neq \{\mathbf{0}_n\}$, its dimension is the cardinality of a basis for S. Notation: $\dim S$
Let $A \in \mathbb{F}^{m \times n}$. The **rank** of A is defined by $\text{rank}\, A := \dim R(A)$
and the **nullity** of A is defined by $\text{null}\, A := \dim N(A)$
The rank is the number of nonzero rows.
Rank-nullity theorem: Let $A \in \mathbb{F}^{m \times n}$. Then, $\text{rank}\, A + \text{null}\, A = n$

7 Hermitian matrices

The **conjugate** of a matrix is defined entry-wise: $[\overline{A}]_{ij} = \overline{a}_{ij}$. $A^* := \overline{A}^T$
An $n \times n$ matrix A is called **Hermitian** is $A^* = A$. $A^*B^* = (AB)^*$
The diagonal entries and eigenvalues of Hermitian matrices are real numbers.

7.1 Scalar products

For two vectors $\boldsymbol{z}, \boldsymbol{w} \in \mathbb{F}^n$, the product $\boldsymbol{z}^*\boldsymbol{w} = \overline{z}_1 w_1, \overline{z}_2 w_2, \dots, \overline{z}_n w_n$ is called the **scalar product** of $\boldsymbol{z}$ and $\boldsymbol{w}$. The scalar product of two real vectors is always a real number.
Two vectors $\boldsymbol{z}, \boldsymbol{w} \in \mathbb{F}^n$ are said to be **orthogonal** if $\boldsymbol{z}^*\boldsymbol{w} = 0$. (notation: $\boldsymbol{z} \perp \boldsymbol{w}$)
$\boldsymbol{z^*z} \geq 0$ $\boldsymbol{z * z} = 0 \iff z = 0$
The **Euclidean length** of a vector is defined by

$$||\boldsymbol{z}|| := \sqrt{\boldsymbol{z}^*\boldsymbol{z}} = \sqrt{|z_1|^2 + |z_2|^2 + \ldots + |z_n|^2}$$

7.2 Orthonormal sets

Let $k, n \in \mathbb{N}$ with $k \leq n$. A set of nonzero vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k\} \subset \mathbb{F}^n$ is said to be an **orthonormal set** if $\boldsymbol{x}_i \perp \boldsymbol{x}_j$ and $||\boldsymbol{x}_i|| = 1$ for all $i, j \in \{1, 2, \dots, k\}$ with $i \neq j$.
The vectors in an orthonormal set are linearly independent.

Let $k, n \in \mathbb{N}$ with $k < n$ and also let $\{\boldsymbol{x_1, x_2, \dots, x_k}\} \subset \mathbb{F}$ be an orthonormal set.
Then, there exists $\boldsymbol{x}_{k+1}$ such that $\{\boldsymbol{x_1, x_2, \dots, x_{k+1}}\}$ is an orthonormal set.
Every vector with length 1 is a part of some orthonormal set with n vectors.

Let $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_k \in \mathbb{F}^n$ and $\boldsymbol{x}_{k+1} \in \mathbb{F}^n$ be vectors such that $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_k\}$ is an orthonormal set and the vectors $\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_k, \boldsymbol{x}_{k+1}$ are linearly independent.
Then, the vector $\boldsymbol{z} = \boldsymbol{x}_{k+1} - \displaystyle\sum_{i=1}^{k} (\boldsymbol{y}_i^* x_{k+1}) \boldsymbol{y}_i$ is nonzero.
Moreover, if $\boldsymbol{y}_{k+1} = \dfrac{\boldsymbol{z}}{||\boldsymbol{z}||}$, then the following statements hold:

1. $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_{k+1}\}$ is an orthonormal set.
2. $\boldsymbol{y}_{k+1} \in \text{span}(\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_k, \boldsymbol{x}_{k+1})$.
3. $\text{span}(\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_{k+1}) = \text{span}(\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_k, \boldsymbol{x}_{k+1})$

Gram-Schmidt process

Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m$ be linearly independent vectors.
Define the vectors $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_m$ recursively by $\boldsymbol{y}_1 = \dfrac{\boldsymbol{x}_1}{||\boldsymbol{x}_1||}$ and $\boldsymbol{y}_k = \dfrac{\boldsymbol{z}}{||\boldsymbol{z}||}$
where $\boldsymbol{z} = \boldsymbol{x}_{k+1} - \sum_{i=1}^{k} (\boldsymbol{y}_i^* x_{k+1}) \boldsymbol{y}_i$ for all k with $2 \leq k \leq m$.
Then, $\{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_l\}$ is an orthonormal set
and for every l with $1 \leq l \leq m$ $\quad \text{span}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_l) = \text{span}(\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_l)$

7.3 Unitary matrices

A matrix $U \in \mathbb{F}^{n \times n}$ is called **orthogonal** if $U^T U = I$ and **unitary** if $U^* U = I$.
Therefore, the inverses of orthogonal and unitary matrices are U^T and U^* respectively.
Let $U \in \mathbb{F}^{n \times n}$. Then, U is a unitary matrix iff the set of column vectors is an orthonormal set.
Let $V, W \in \mathbb{F}^{n \times n}$ be unitary matrices. Then, VW is also unitary.

Schur's theorem: For every matrix $A \in \mathbb{F}^{n \times n}$. there exists a unitary matrix $U \in \mathbb{F}^{n \times n}$ such that $U^* A U$ is an upper triangular matrix. $\quad A = UTU^*$ is called a **Schur decomposition**.
If $A \in \mathbb{R}^{n \times n}$ has only real eigenvalues, there exists an orthogonal matrix $U \in \mathbb{R}^{n \times n}$ such that $U^T A U$ is a real upper triangular matrix.

Let $\lambda_1, \lambda_2, \ldots, \lambda_l$ be the distinct eigenvalues of $A \in \mathbb{F}^{n \times n}$. Then the characteristic polynomial of A can be factorized as

$$p_A(\lambda) = \prod_{i=1}^{l} (\lambda - \lambda_i)^{\alpha_i}$$

where $\alpha_1 \geq 1$ denote the algebraic multiplicity of λ_i. Note the $\sum_{i=1}^{l} \alpha_i = n$.

Let $\lambda_1, \lambda_2, \ldots, \lambda_l$ be the distinct eigenvalues of $A \in \mathbb{F}^{n \times n}$ with respective algebraic multiplicities $\alpha_1, \alpha_2, \ldots, \alpha_l$. Then, there exists a Schur decomposition such that

$$T = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1l} \\ 0 & T_{22} & \cdots & T_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & T_{ll} \end{bmatrix}$$

where $T_{ii} \in \mathbb{F}^{\alpha_i \times \alpha_i}$ are upper triangular matrices whose diagonal entries are all λ_i.

$A \in \mathbb{F}^{n \times n}$ is **unitarily diagonalizable** if there is a unitary matrix U s.t. $U^* A U$ is diagonal.
Spectral theorem: Let $A \in \mathbb{F}^{n \times n}$ be a Hermitian matrix.
Then there exists a unitary matrix U such that $U^* A U$ is a real diagonal matrix.
For every $n \times n$ Hermitian matrix, there exist n eigenvectors that form an orthonormal set.
In particular, for every $n \times n$ real symmetric matrix, there exist n real eigenvectors that form an orthonormal set.

A square matrix A is **normal** if $A^* A = AA^*$. Every Hermitian matrix is normal.
A matrix is unitarily diagonalizable if and only if it is normal.

Let A be an $n \times n$ matrix:
A is Hermitian $\implies$ A is normal $\iff$ A is unitarily diagonalizable $\implies$ A is diagonalizable $\iff$ A has n lin. ind. eigenvectors $\impliedby$ A has n distinct eigenvalues

Suppose that A is diagonalizable. Let λ be an eigenvalue of A with algebraic multiplicity α.
Then, there exist α linearly independent eigenvectors corresponding to λ.

Let A be Hermitian and let $(\lambda_i, \boldsymbol{x}_i)$ be eigenpairs of A with $i \in \{1, 2\}$. If $\lambda_1 \neq \lambda_2$, then $\boldsymbol{x}_1 \perp \boldsymbol{x}_2$.

Let A be an $n \times n$ Hermitian matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_l$ be distinct eigenvectors with multiplicities $\alpha_1, \alpha_2, \ldots, \alpha_l$. Let $i \in \{1, 2, \ldots, l\}$ and let $x_1^i, x_2^i, \ldots, x_{a_i}^i$ be linearly independent eigenvectors corresponding to λ_i. The Gram-Schmidt process yields an orthonormal set of eigenvectors $y_1^i, y_2^i, \ldots, y_{a_i}^i$. Then,

$$\begin{bmatrix} y_1^1 & y_2^1 & \cdots & y_{\alpha_1}^1 \mid y_1^2 & y_2^2 & \cdots & y_{\alpha_2}^2 \mid \cdots \mid y_1^l & y_2^l & \cdots & y_{\alpha_l}^l \end{bmatrix}$$

is a unitary diagonalizer for A.

8 Jordan matrices

Definition of **geometric multiplicity**: $\gamma_i = \dim N(\lambda_i I - A)$.
A matrix has exactly $\sum_{i=1}^{l} \gamma_i$ linearly independent eigenvectors.
The geometric multiplicity is less than or equal to its algebraic multiplicity.
A square matrix is diagonalizable iff algebraic and geometric multiplicities are the same for every λ.

Let $k \geq 1$. A $k \times k$ **Jordan block** with eigenvalue λ is the **upper bidiagonal matrix**, with λ as the diagonal entries and 1 as the entry above every diagonal entry.
Bdiag$(M_1, M_2, \ldots, M_l)$ is the $m \times m$ matrix with matrices $M_i \in \mathbb{F}^{m \times m}$ on the diagonal where $m = \sum_{i=1}^{l} m_i$.
An $n \times n$ **Jordan matrix** with eigenvalue λ is a block diagonal matrix

$$J(\lambda) = \text{Bdiag}(J_{k_1}(\lambda), J_{k_2}(\lambda), \ldots, J_{k_m}(\lambda)) \text{ where } n = \sum_{i=1}^{m} k_i$$

Definition of the **Jordan canonical form** J of A:
Let $\lambda_1, \lambda_2, \ldots, \lambda_l$ be the distinct eigenvalues of $A \in F^{n \times n}$ with respective algebraic multiplicities $\alpha_1, \alpha_2, \ldots, \alpha_l$ and geometric multiplicities $\gamma_1, \gamma_2, \ldots, \gamma_l$. Then, there exists a matrix J given by

$$J = \text{Bdiag}(J(\lambda_1), J(\lambda_2), \ldots, J(\lambda_l))$$

where $J(\lambda_i) \in \mathbb{F}^{\alpha_1 \times \alpha_1}$ is a Jordan matrix of the form

$$J = \text{Bdiag}(J_{k_1^i}(\lambda_1), J_{k_2^i}(\lambda_2), \ldots, J_{k_{\gamma_i}^i}(\lambda_l))$$

with $k_1^i \geq k_2^i \geq k_{\gamma_i}^i$ and $\sum_{j=0}^{\gamma_i} k_j^i = \alpha_i$ such that $A \sim J$.

Let $A \in \mathbb{F}^{n \times n}$ and let λ be an eigenvalue of A with algebraic multiplicity α and geometric multiplicity γ. For $p \in \{1, 2, \ldots, n+1\}$, define

$$w_p := \text{rank}(\lambda I - A)^{p-1} - \text{rank}(\lambda I - A)^p$$

- $w_p \geq 0$
- There exists a κ such that $w_\kappa > 0$ and $w_i = 0$ for $i > \kappa$.
- κ is the **index** of the eigenvalue λ
- $w_1, w_2, \ldots, w_\kappa$ are the **Weyr characteristics** of A for the eigenvalue λ.
- κ is the size of the largest Jordan block with the eigenvalue λ.
- $w_i \geq w_{i+1}$ for $i \in \{1, 2, \ldots, \kappa\}$.

- Define $\rho_i = w_i - w_{i+1}$ for $i \in \{1, 2, \dots, \kappa\}$.
- ρ_i is the number of Jordan blocks with size i.

Similarity is an equivalence relation.
For $i \in \{1, 2, \dots, l\}$, let M_{ii}, N_{ii} be $q_i \times q_i$ matrices such that $M_{ii} \sim N_{ii}$. Then,

$$\text{Bdiag}(M_{11}, M_{22}, \dots, M_{ll}) \sim \text{Bdiag}(N_{11}, N_{22}, \dots, N_{ll})$$

Let λ be a scalar. Then, $N \sim M$ if and only if $(\lambda I + N) \sim (\lambda I + M)$
Let $A \in \mathbb{F}^{n \times n}$ and $B \in \mathbb{F}^{m \times m}$. Suppose that A and B do not have a common eigenvalue.
For every $C \in \mathbb{F}^{n \times m}$, there exists an $X \in \mathbb{F}^{n \times m}$ such that $AX - XB = C$. Let $A \in \mathbb{F}^{n \times n}$, $B \in \mathbb{F}^{m \times n}$ and $C \in \mathbb{F}^{n \times m}$. Suppose that A and B do not have a common eigenvalue. Then, we have

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \sim \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

Let $N \in \mathbb{F}^{n \times n}$ be a strictly upper triangular matrix. Then, there exists a Jordan matrix J with eigenvalue zero of the form $J = \text{Bdiag}(J_{k_1}, J_{k_2}, \dots, J_{k_m})$
where $k \geq k_2 \geq \dots \geq k_m$ and $n = \sum_{i=0}^{m} k_i$ such that $N \sim J$.
Moreover, m is the geometric multiplicity of the zero eigenvalue of N.

Notation: $J_k := J_k(0)$ $J_k = \begin{bmatrix} 0 & I_{k-1} \\ 0 & 0 \end{bmatrix}$
Permuting the rows and columns of a matrix yields a similar matrix.

8.1 Cayley-Hamilton Theorem

For every square matrix A, $p_A = 0$.
λ^k is an eigenvalue of A^k if and only if λ is an eigenvalue of A.
Let A be a square matrix. We say that a polynomial p is an **annihilating polynomial** of A or p **annihilates** A is $p(A) = 0$.
Let A be a square matrix. Then, there exists a unique monic polynomial m_A of minimum positive degree that annihilates A. We call m_A the **minimal polynomial** of A.
Let A be a square matrix and let p be a nonconstant polynomial that annihilates A. Then, m_A divides p.
Let $\lambda_1, \lambda_2, \dots, \lambda_l$ be distinct eigenvalues of $A \in \mathbb{F}^{n \times n}$ with respective indices $\kappa_1, \kappa_2, \dots, \kappa_l$. Then,

$$m_A(\lambda) = \prod_{i=1}^{l} (\lambda - \lambda_1)^{\kappa_i}$$