MATH2022: Linear and Abstract Algebra Notes

Kenny Y.

Contents

1 Groups

1.1 Basic Definitions of Field and Group

Definition 1.1.1. For $n \in \mathbb{N}$, $n \geq 1$, the set $\mathbb{Z}_n$ or $\mathbb{Z}/n\mathbb{Z}$ is the set of remainders when dividing by n, i.e.

$$\mathbb{Z}_n = \{0, \ldots, n-1\}.$$

Operations allowed on $\mathbb{Z}_n$ are $+$, $-$, $\times$. If $a, b \in \mathbb{Z}_n$ with $b \neq 0$, we can define a/b as the number $x \in \mathbb{Z}_n$ such that $bx = 1$ in $\mathbb{Z}_n$ if it exists.

Definition 1.1.2 (Field). A field $(F, +, \cdot)$ is a set F with an $+$ and a $\cdot$ with the following properties for all $a, b, c \in F$.

1. **Associativity:** $(a+b)+c = a+(b+c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
2. **Commutativity:** $a+b = b+a$ and $a \cdot b = b \cdot a$.
3. **Distribution:** $a \cdot (b+c) = a \cdot b + a \cdot c$.
4. **Identities:** There is $0 \in F$ and $1 \in F$, and such that $a+0 = a$, $a \cdot 1 = a$.
5. **Inverses:** There is $-a \in F$ for all $a \in F$ with $a + (-a) = 0$, and there is $a^{-1} \in F$ for all $a \in F$, $a \neq 0$, with $a \cdot a^{-1} = 1$.

Theorem 1.1.3. $\mathbb{Z}_n$ is a field if and only if n is a prime.

Definition 1.1.4 (Group). A group $(G, *)$ is a set G and an operation $*$ with the following properties for all $a, b, c \in G$.

1. **Associativity:** $(a * b) * c = a * (b * c)$.
2. **Identity:** There is an element $e \in G$ such that $a * e = e * a = a$ for all $a \in G$.
3. **Inverses:** For all $a \in G$ there is an element $b \in G$ such that $a * b = b * a = e$.

Note: We do not assume commutativity or abelian.

1.2 The Symmetric Group

Definition 1.2.1 (Permutation). A permutation of a set X is a bijection $f : X \to X$.

Definition 1.2.2 (Inverse Permutation). Given a permutation $f : X \to X$, the inverse permutation $f^{-1} : X \to X$ is the inverse map defined by $f^{-1}(w) = x$ if $f(x) = w$.

Example 1.2.3. $X = \{a, b, c, d\}$ and $f : X \to X$ defined by $a \mapsto b$, $c \mapsto a$, $b \mapsto c$, $d \mapsto d$, then f^{-1} is defined by $a \mapsto c$, $b \mapsto a$, $c \mapsto b$, $d \mapsto d$.

Theorem 1.2.4.

1. If f and g are permutations, then $f \circ g$ is also a permutation.
2. Composition of maps is associative.
3. The identity map $id_X : X \to X$, $x \mapsto x$ is a permutation.
4. Composing f and f^{-1} gives the identity map id_X.

Definition 1.2.5 (Symmetric Groups). The symmetric group on n elements, denoted by S_n or $\mathrm{Sym}(n)$, is group of permutations on the set $\{1, 2, 3, \ldots, n\}$.

Fact: $|S_n| = n!$.

1.3 Transpositions

Definition 1.3.1 (Transpositions). A transposition $\tau \in S_n$ is a permutation which only swaps two elements, that is

$$\tau = (a\ b)$$

for some a, b. We call it a simple transposition if it has a cycle notation

$$\tau = (i\ i+1).$$

Proposition 1.3.2. Any cycle can be written as a product of transpositions.

Proof. $(a_1\ a_2\ \ldots\ a_k) = (a_1\ a_2)(a_1\ a_3)\ldots(a_1\ a_k)$.

Corollary 1.3.3. As a permutation is a product of disjoint cycles, any permutation is a product of transpositions.

Example 1.3.4. $\alpha = (1\ 3\ 5)(2\ 4\ 6) = (1\ 3)(1\ 5)(2\ 4)(2\ 6)$.

Definition 1.3.5 (Parity). We say a permutation α is *even* (*odd*) if it can be written as a product of an even (odd) number of transpositions.

Theorem 1.3.6. Every permutation α in S_n is either even or odd and not both.

Remark 1.3.7.

1. A cycle $(a_1\ a_2\ \dots\ a_k)$ has parity

$$\begin{cases} \text{even}, & k \text{ is odd} \\ \text{odd}, & k \text{ is even} \end{cases}$$

2. Transpositions are self-inverse: $\tau = \tau^{-1} = (a\ b)$.
3. A permutation α and its inverse α^{-1} have same parity.

1.4 Subgroups

Definition 1.4.1 (Subgroup). Let $(G, *)$ be a group. Let $H \subseteq G$ be a subset. Then we say H is a subgroup of G if $(H, *)$ is a group with the same operation. We denote this by $H \leq G$.

Definition 1.4.2 (The Alternating Group). The set $\mathrm{Alt}(n)$ or A_n is defined by

$$\{\alpha \in S_n \mid \alpha \text{ even}\}.$$

We call this the alternating group on n elements.

Proposition 1.4.3. $A_n \leq S_n$.

Fact:

$$|A_n| = \frac{1}{2}|S_n| = \frac{1}{2}(n!)$$

1.5 Conjugation in Group

Definition 1.5.1 (Conjugate). Let G be a group, and $a, b \in G$. The conjugate of a by b is the element

$$b^{-1}ab \in G.$$

Sometimes denoted as a^b (does not mean a to the power of b).

Theorem 1.5.2.

1. $(ab)^c = a^c b^c$
2. $(a^b)^{-1} = (a^{-1})^b$
3. $(a^b)^c = a^{bc}$

Proposition 1.5.3. Let $\alpha = (a_1\ \dots\ a_k)$ be a cycle, and β be any permutation. Then

$$\beta^{-1}\alpha\beta = \alpha^\beta = (\beta(a_1)\ \beta(a_2)\ \dots\ \beta(a_k))$$

is the cycle where $\beta(a_i)$ means $a_i \mapsto \beta(a_i)$.

Corollary 1.5.4. If $\alpha = \sigma_1\sigma_2\cdots\sigma_l$ is a disjoint cycle expression for α, then

$$\alpha^\beta = \sigma_1^\beta\sigma_2^\beta\cdots\sigma_l^\beta$$

by our general properties of conjugation.

Theorem 1.5.5.

1. For any $\alpha, \beta \in S_n$, α and α^β have the same *parity*.
2. For any $\alpha, \beta \in S_n$, α and α^β have the same *cycle type*.

Definition 1.5.6 (Cycle type). We call a cycle $\sigma = (a_1\ a_2\ \ldots\ a_k)$ a k-cycle. For $\alpha \in S_n$, if $\alpha = \sigma_1\sigma_2\cdots\sigma_l$ a disjoint cycle expression, and σ_i is a k_i-cycle, we say α has a cycle type $[k_1, k_2, \ldots, k_l]$. Convention: $k_1 \geq k_2 \geq \cdots \geq k_l$.

Theorem 1.5.7. $\alpha, \gamma \in S_n$ have the same cycle type if and only if $\gamma = \alpha^\beta$ for some β.

1.6 Dihedral Groups

Definition 1.6.1 (Dihedral Group). The dihebral group, denoted D_n $(n \geq 3)$ is the group of symmetries of a regular n-gon.

Definition 1.6.2 (D_n in General). For a regular n-gon, we get basic symmetries $r =$ rotation by $2\pi/n$, $s =$ reflection. Then

$$\begin{aligned} D_n &= \{1, r, \ldots, r^{n-1}, s, sr, \ldots, sr^{n-1}\} \\ &= \langle r, s \mid r^n = 1, s^2 = 1, srs = r^{-1} \rangle, \end{aligned}$$

and $|D_n| = 2n$.

1.7 Cyclic Subgroups

Definition 1.7.1 (Cyclic Subgroup). Let G be a group, and $g \in G$. For $k \in \mathbb{Z}$, let

$$g^k = \begin{cases} kg, & k > 0 \\ e, & k = 0 \\ kg^{-1}, & k < 0 \end{cases}$$

Then $\langle\, g \,\rangle = \{g^k \mid k \in \mathbb{Z}\}$ is the cyclic subgroup of G generated by g.

Fact: $\langle\, g \,\rangle$ is abelian.

Definition 1.7.2 (Additive Group). When G is a group whose operation is addition, like $\mathbb{Z}$, $\mathbb{Z}_n$, we say G is an *additive group*.

$$\langle\, g \,\rangle = \{kg \mid k \in \mathbb{Z}\}$$

Theorem 1.7.3. If $k \in \mathbb{Z}_n$, then k generates $\mathbb{Z}_n$, i.e. $\langle\, k \,\rangle = \mathbb{Z}_n$, if and only if k and n are relatively prime, i.e. coprime, $\gcd(k, n) = 1$.

Definition 1.7.4. We say a general group G is *cyclic* if $G = \langle g \rangle$ for some $g \in G$, call g a *generator* of G.

1.8 Group Isomorphism

Definition 1.8.1. For $(G, \cdot)$ and $(H, *)$, we say G is isomorphic to H if there is a bijection $\varphi : G \to H$ such that for all $a, b \in G$,

$$\varphi(a \cdot b) = \varphi(a) * \varphi(b).$$

We denote this by $G \cong H$.

In general, if $G = \langle\, g \,\rangle$ is a cyclic group, then either $G \cong (\mathbb{Z}_n, +)$ or $G \cong (\mathbb{Z}, +)$ with an isomorphism $\varphi : g$, generator $\mapsto$ generator in $\mathbb{Z}_n$ or $\mathbb{Z}$.

1.9 Cartesian Products of Cyclic Groups

Definition 1.9.1 (Cartesian Product). Let $(\mathbb{Z}_n, +)$ and $(\mathbb{Z}_m, +)$ be finite groups. Then the Cartesian product

$$\mathbb{Z}_n \times \mathbb{Z}_m = \{(a, b) \,|\, a \in \mathbb{Z}_n,\ b \in \mathbb{Z}_m\}$$

is also a group under component wise addition $(a, b) + (c, d) = (a + c, b + d)$.

Theorem 1.9.2. As additive groups, $\mathbb{Z}_n \times \mathbb{Z}_m \cong \mathbb{Z}_{nm}$ if and only if n, m are relatively prime.

2 Matrices and Linear Transformations

2.1 Elementary Matrices, Invertibility and Determinants

Theorem 2.1.1 (Row-switching transformations). The elementary row operation $R_i \leftrightarrow R_j$ is performed on A by multiplying the elementary matrix

$$E = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

on its left. Coefficient wise,

$$[E]_{k,l} = \begin{cases} 0 & k \neq i, k \neq j, k \neq l \\ 1 & k \neq i, k \neq j, k = l \\ 0 & k = i, l \neq j \\ 1 & k = i, l = j \\ 0 & k = j, l \neq i \\ 1 & k = j, l = i \end{cases}$$

Theorem 2.1.2 (Row-scaling transformations). The elementary row operation $R_i \leftarrow \lambda R_i$ is performed on A by multiplying the elementary matrix

$$E = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

on its left. Coefficient wise,

$$[E]_{k,l} = \begin{cases} 0 & k \neq l \\ 1 & k = l, k \neq i \\ \lambda & k = l, k = i \end{cases}$$

Theorem 2.1.3 (Row-addition transformations). The elementary row operation $R_i \leftarrow R_i + \lambda R_j$ $(i \neq j)$ is performed on A by multiplying the elementary matrix

$$E = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & \lambda & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Coefficient wise,

$$[E]_{k,l} = \begin{cases} 0 & k \neq l, k \neq i, l \neq j \\ 1 & k = l \\ \lambda & k = i, l = j \end{cases}$$

Theorem 2.1.4. A matrix is invertible if and only if it can be row reduced to the identity.

Definition 2.1.5 (Determinant). For $A = [a_{ij}] \in \mathrm{Mat}_n(F)$, choose a row i. Then

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

where A_{ik} is the $(n-1) \times (n-1)$ matrix with row i and column k removed.

Proposition 2.1.6. If A is a triangular matrix, i.e.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} a_{11} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n-11} & \cdots & a_{n-1n-1} & 0 \\ a_{n1} & \cdots & a_{nn-1} & a_{nn} \end{pmatrix},$$

then

$$\det A = \prod_{k=1}^{n} a_{kk} = a_{11} a_{22} \cdots a_{nn}.$$

Theorem 2.1.7. For $A, B \in \mathrm{Mat}_n(F)$.

(i) $\det(AB) = \det A \cdot \det B$

(ii) $\det A = \det(A^T)$

(iii) A is invertible if and only if $\det A \neq 0$

Theorem 2.1.8 (Elementary Row Operations and Determinants).

1. Swapping two rows *multiplies the determinant by* -1.
2. Scaling a row by λ *multiplies the determinant by* λ.
3. Adding a multiple of a row to another row *does not change the determinant.*

2.2 Rotation and Reflection Matrices

Definition 2.2.1 (Rotation and Reflection Matrices). Let $\theta \in \mathbb{R}$. Define R_θ and T_θ as follows

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad T_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

The corresponding matrix transformations are given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix},$$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto T_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta + y\sin\theta \\ x\sin\theta - y\cos\theta. \end{pmatrix}$$

R_θ rotates the point (x, y) counterclockwise by θ radians. T_θ reflects across the line making angle $\theta/2$ with positive x-axis.

Theorem 2.2.2 (Properties).

1. $R_\theta = R_{2\pi} = I$
2. $R_\theta R_\psi = R_{\theta+\psi}$
3. $R_\theta^n = R_{n\theta}$
4. $R_\theta^{-1} = R_{-\theta}$
5. $T_\theta^2 = I$, so $T_\theta = T_\theta^{-1}$
6. $T_\psi^{-1} R_\theta T_\psi = T_\psi R_\theta T_\psi = R_\theta^{-1} = R_{-\theta}$
7. $R_\theta T_\psi R_\theta = T_\psi$
8. $T_\theta T_\psi = R_{\theta-\psi}$

Theorem 2.2.3. $\langle R_{2\pi/n}, T_{2\pi/n} \rangle \cong D_n$.

Theorem 2.2.4. The set $\{R_\theta, T_\theta \,|\, \theta \in \mathbb{R}\}$ is isomorphic to the *symmetry group of a circle.*

2.3 Linear Transformations

Definition 2.3.1 (Linear Maps). For $n, m \geq 0$, a function $L : F^n \to F^m$ is called *linear* or a *linear map* from F^n to F^m if for all $\mathbf{v}, \mathbf{w} \in F^n$ and $\lambda \in F$,

1. $L(\mathbf{v} + \mathbf{w}) = L(\mathbf{v}) + L(\mathbf{w})$,
2. $L(\lambda \mathbf{v}) = \lambda L(\mathbf{v})$.

Theorem 2.3.2.

1. For any linear map $L : F^n \to F^m$, $L(\mathbf{0}) = \mathbf{0}$.
2. $L : F^n \to F^m$ is linear if and only if

$$L(\lambda_1 \mathbf{v} + \lambda_2 \mathbf{w}) = \lambda_1 L(\mathbf{v}) + \lambda_2 L(\mathbf{w})$$

   for any $\lambda, \mu \in F$ and $\mathbf{v}, \mathbf{w} \in F^n$.

Proposition 2.3.4. Let $L_1 : F^n \to F^m$, $L_2 : F^m \to F^p$ be two linear maps. Then

$$L_2 \circ L_1 : F^n \to F^p$$

is also a linear map.

Definition 2.3.5 (Identity Map). The identity map $\mathrm{id} : F^n \to F^n$ is given by $\mathrm{id}(\mathbf{v}) = \mathbf{v}$.

Definition 2.3.6 (Invertible Map). We say a linear map $L : F^n \to F^n$ is invertible if there is some linear map $T : F^n \to F^n$ such that

$$L \circ T = T \circ L = id : F^n \to F^n.$$

Then we write $T = L^{-1}$.

Proposition 2.3.7. For $M \in \mathrm{Mat}_{m \times n}(F)$, the associated map $L_M : F^n \to F^m$ is *linear*.

Proposition 2.3.8. If $L : F^n \to F^m$ is linear, then there exists a unique $M \in \mathrm{Mat}_{m \times n}(F)$ representing L. In which, the jth column of M is equivalent to $L(\mathbf{e}_j)$, where $\mathbf{e}_j$ is the jth standard basis vector in F^n.

Theorem 2.3.9. If $L_1 : F^n \to F^m$ is represented by $M_1 \in \mathrm{Mat}_{m,n}(F)$ and $L_2 : F^m \to F^p$ is represented by $M_2 \in \mathrm{Mat}_{p \times m}(F)$, then $L_2 \circ L_1 : F^n \to F^p$ is represented by $M_2 M_1$.

Theorem 2.3.10. If $L : F^n \to F^n$ is invertible, and L is represented by M, then

1. M is invertible,
2. L^{-1} is represented by M^{-1}.

3 Eigenvalues and Eigenvectors

3.1 Eigenvalues, Eigenvectors and Eigenspaces

Definition 3.1.1 (Eigenvalues and Eigenvectors). Let $M \in \mathrm{Mat}_n(F)$. Let $\mathbf{v}$ be a nonzero column vector in F^n. We say that $\mathbf{v}$ is an eigenvector for M if there is some scalar λ such that

$$M\mathbf{v} = \lambda\mathbf{v}.$$

If so, we say λ is an *eigenvalue* of M and $\mathbf{v}$ is an *eigenvector* associated to λ.

Definition 3.1.2 (Eigenspace). The λ-eigenspace of M is the set of solutions

$$\{\mathbf{v} \,|\, (M - \lambda I) = \mathbf{0}\}.$$

Proposition 3.1.3. $\lambda \in F$ is an eigenvalue of M if and only if $\det(M - \lambda I) = 0$.

Definition 3.1.4 (Characteristic Equation). Let $M \in \mathrm{Mat}_n(F)$. The determinant $\det(M - \lambda I)$ is a polynomial in λ of degree n, which we call the *characteristic polynomial of M*. The equation $\det(M - \lambda I) = 0$ is called the *characteristic equation of M*.

Corollary 3.1.5. The eigenvalues of M are given by the roots (or zeros) of the characteristic polynomial of M.

Definition 3.1.6 (Algebraic Multiplicity). For a matrix M, we say an eigenvalue λ of M has algebraic multiplicity k if λ is a root of the characteristic polynomial k times.

Definition 3.1.7 (Geometric Multiplicity). For a matrix M, we say an eigenvalue λ of M has geometric multiplicity k if the λ-eigenspace has k parameters. That is, the solution to $(M - \lambda I) = \mathbf{0}$ had k free variables.

Proposition 3.1.8. If M is a square matrix with an eigenvalue λ, then

$$\text{geometric multiplicity of } \lambda \leq \text{algebraic multiplicity of } \lambda.$$

3.2 Diagonalisation

Definition 3.2.1 (Diagonalisable Matrix). For M, an $n \times n$ matrix, we say M is diagonalisable if there exists an invertible matrix P and diagonal matrix D such that

$$M = PDP^{-1}.$$

If so, the diagonal entries of D are the eigenvalues of M and the columns of P are the corresponding eigenvectors of M.

Theorem 3.2.2. A matrix M is diagonalisable if and only if for all eigenvalues λ,

$$\text{algebraic multiplicity } \lambda = \text{geometric multiplicity } \lambda.$$

Definition 3.2.3 (Matrix Polynomial). Let $M \in \mathrm{Mat}_n(F)$, $p(x) = \sum_{k=0}^{n} a_k x^k$ be a polynomial in a variable x with coefficients $a_0, a_1, \ldots, a_n \in F$. The matrix polynomial $p(M)$ is given by

$$p(M) = a_0 I + a_1 M + a_2 M^2 + \cdots + a_n M^n.$$

Theorem 3.2.4 (Cayley-Hamilton Theorem). Let $M \in \mathrm{Mat}_n(F)$ with characteristic polynomial $\chi_M(\lambda) = \det(M - \lambda I)$. Then,

$$\chi_M(M) = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Theorem 3.2.5. If M is invertible, then there is some polynomial $p(x)$ such that

$$M^{-1} = p(M).$$

3.3 Application: Stochastic Matrices

Definition 3.3.1 (Probability Vector and Stochastic Matrices). A *probability vector* is a vector $\mathbf{v}$ with non-negative entries that sum to 1. A *stochastic matrix* is a matrix whose columns are probability vectors. If both columns and rows sum to 1, we call that matrix *doubly stochastic*.

Theorem 3.3.2. Let A be a stochastic matrix. Then 1 is an eigenvalue of A, so there is *at least one steady-state probability vector* $\mathbf{v}$ with $A\mathbf{v} = \mathbf{v}$. And, all eigenvalues λ of A have $|\lambda| \leq 1$. If A is regular stochastic, meaning A^k has strictly positive entries for some $k \geq 1$, then

- There is a *unique steady state probability vector* $\mathbf{v}$.
- $\lim_{n\to\infty} A^n = \begin{pmatrix} \mathbf{v} & \mathbf{v} & \cdots & \mathbf{v} \end{pmatrix}$
- For any probability vector $\mathbf{x}$,
$$\lim_{n\to\infty} A^n \mathbf{x} = \mathbf{v}.$$

3.4 Matrix Exponential and Systems of Differential Equations

Definition 3.4.1 (Matrix Exponential). For $M \in \mathrm{Mat}_n(\mathbb{R})$, the *matrix exponential* e^M is defined as

$$e^M := \sum_{k=0}^{\infty} \frac{1}{k!} M^k = I_n + M + \frac{1}{2!} M^2 + \frac{1}{3!} M^3 + \cdots$$

This converges to a well-defined matrix in $\mathrm{Mat}_n(\mathbb{R})$.

Theorem 3.4.2.

1. If $B = P^{-1}AP$, then $e^B = P^{-1}e^A P$.
2. If $AB = BA$, then $e^{A+B} = e^A e^B = e^B e^A$.

Proposition 3.4.3. If $D = \begin{pmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_n \end{pmatrix}$ is diagonal, then $e^D = \begin{pmatrix} e^{a_1} & 0 & \cdots & 0 \\ 0 & e^{a_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{a_n} \end{pmatrix}$.

Corollary 3.4.4. If M is diagonalisable, and $M = PDP^{-1}$, then $e^M = Pe^D P^{-1}$.

Definition 3.4.5 (Vector-Valued Function). A *vector-valued function* is a function of the form

$$\mathbf{v}(t) = \begin{pmatrix} v_1(t) \\ \vdots \\ v_n(t) \end{pmatrix}$$

for $t \in \mathbb{R}$, where each $v_i(t)$ is a differentiable function on $\mathbb{R}$. Then we define

$$\mathbf{v}'(t) = \begin{pmatrix} v_1'(t) \\ \vdots \\ v_n'(t) \end{pmatrix}.$$

Definition 3.4.6. For a vector-valued function $\mathbf{v}(t)$, a *first-order constant coefficient homogeneous system of linear differential equations* is given by

$$\mathbf{v}'(t) = A\mathbf{v}(t)$$

for some $A \in \mathrm{Mat}_n(\mathbb{R})$.

Theorem 3.4.7. The system $\mathbf{v}'(t) = A\mathbf{v}(t)$ has solutions

$$\mathbf{v}(t) = e^{tA}\mathbf{v}(0)$$

4 The Structure of Abstract Vector Spaces

4.1 Vector Spaces

Definition 4.1.1 (Vector Space). Fix a field F of scalars. Roughly a vector space V over the field F is an *abelian group* with a *"compatible" scalar multiplication*. More precisely, we have two operations:

- If $\mathbf{v}, \mathbf{w} \in V$, then $\mathbf{v} + \mathbf{w} \in V$ (addition).
- If $c \in F$ and $\mathbf{v} \in V$, then $c\mathbf{v} \in V$ (multiplication by scalars).

Axiom 1 (Scalar multiplication association). If $c, d \in F$ and $\mathbf{v} \in V$, then $c(d\mathbf{v}) = (cd)\mathbf{v}$.

Axiom 2 (Distribution). If $c, d \in F$ and $\mathbf{v}, \mathbf{w} \in V$, then

$$c(\mathbf{v} + \mathbf{w}) = c\mathbf{v} + c\mathbf{w},$$

$$(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}.$$

Axiom 3 (Scalar identity). $1 \in F$, $\mathbf{v} \in V$, then $1\mathbf{v} = \mathbf{v}$.

Axiom 4 (Group Axioms for $(V, +)$).
Addition $+$ is commutative, i.e. $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$.
Zero vector: $\mathbf{0} + \mathbf{v} = \mathbf{v} = \mathbf{v} + \mathbf{0}$.
Additive inverses: $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

Definition 4.1.2 (Subspace). Let V be a vector space over F. A subset $W \subseteq V$ is called a subspace of V if

- $W \neq \emptyset$ (nonempty),
- W is itself a vector space over F using operations from V, i.e. if $\mathbf{v}, \mathbf{w} \in W$ and $\lambda \in F$, then

$$\mathbf{v} + \mathbf{w} \in W \quad \text{and} \quad \lambda\mathbf{v} \in W.$$

4.2 Spans of Vectors

Definition 4.2.1 (Linear Combination). Let X be a subset of vectors in a vector space V. A linear combination of vectors from X is an expression of the form

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_k\mathbf{v}_k$$

where $\lambda_i \in F$, $\mathbf{v}_i \in X$.

Definition 4.2.2 (Span). Let V be a vector space, and X be a subset of vectors in V. The *span of* X, denoted $\operatorname{span}(X)$ or $\langle X \rangle$ is defined to be

- if $X = \emptyset$, set $\operatorname{span}(X) = \{\mathbf{0}\}$;
- if $X \neq \emptyset$, then
$$\operatorname{span}(X) = \{\lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k \mid \lambda_i \in F, \mathbf{v}_i \in X\}.$$

We call $\operatorname{span}(X)$ the *subspace of* V *spanned by* X. If $X = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, we often write

$$\operatorname{span}(X) = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \quad \text{or} \quad \langle X \rangle = \langle \mathbf{v}_1, \ldots, \mathbf{v}_n \rangle.$$

Theorem 4.2.3. Let V be a vector space over F, X a subset of V. If $W \subseteq V$ is a subspace of V which contains X, then $\operatorname{span}(X) \subseteq W$ is a subspace of W.

4.3 Row and Column Space of a Matrix

Definition 4.3.1 (Row and Column Space). Let $M \in \mathrm{Mat}_{m\times n}(F)$, $\{R_1, \ldots, R_m\}$ be rows of M, and $\{C_1, \ldots, C_n\}$ be columns of M. The row space of M is given by the span

$$\mathrm{Row}(M) = \operatorname{span}\{R_1, \ldots, R_m\} \subseteq \mathrm{Mat}_{1\times n}(F) \approx F^n.$$

The column space of M is given by the span

$$\mathrm{Col}(M) = \operatorname{span}\{C_1, \ldots, C_n\} \subseteq \mathrm{Mat}_{m\times 1}(F) \approx F^m.$$

Theorem 4.3.2. Let $A, B \in \mathrm{Mat}_{m\times n}(F)$. Then $\mathrm{Row}(A) = \mathrm{Row}(B)$ if and only if they are row equivalent (same RREF); $\mathrm{Col}(A) = \mathrm{Col}(B)$ if and only if $\mathrm{Row}(A^T) = \mathrm{Row}(B^T)$.

4.4 Linear Independence

Definition 4.4.1 (Non-trivial Linear Combination). Given $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ over F. A *non-trivial linear combination* of these vectors is in the form

$$\sum_{i=1}^{k} \lambda_i \mathbf{v}_i = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k$$

for $\lambda_1, \ldots, \lambda_k \in F$ where at least one of $\lambda_i \neq 0$, $i = 1, \ldots, k$.

Definition 4.4.2 (Linearly Dependent). A collection of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ in V is *linearly dependent* if some non-trivial linear combination of then is $\mathbf{0}$, that is

$$\sum_{i=1}^{k} \lambda_i \mathbf{v}_i = \mathbf{0}$$

for $\lambda_1, \ldots, \lambda_k \in F$ where at least one of $\lambda_i \neq 0$, $i = 1, \ldots, k$.

Theorem 4.4.3. A collection of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ in V is linearly dependent if and only if one vector is a linear combination of the others. So

$$\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k\}.$$

Definition 4.4.4 (Linearly Independent). A collection of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ in V over F is said to be *linearly independent* if

$$\sum_{i=0}^{k} \lambda_i \mathbf{v}_i = \mathbf{0}$$

implies that $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$. No $\mathbf{v}_i$ is in the span of the others, so removing $\mathbf{v}_i$ from $\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ produces a smaller subspace.

Theorem 4.4.5.

1. Any collection of vectors containing $\mathbf{0}$ is linearly dependent.
2. Any collection with repeated vectors in linearly dependent.
3. Two vectors $\mathbf{v}$ and $\mathbf{w}$ are linearly dependent if and only if $\mathbf{v} = \lambda \mathbf{w}$, $\lambda \in F$.
4. In F^n, any collection with $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ with $k > n$ is linearly dependent.
5. For a matrix M, its rows are linearly independent if and only if no row of zeros appears while row reducing.
6. The nonzero rows of a matrix M in echelon form are linearly independent and span $\text{Row}(M)$.

4.5 Bases, Dimensions and Coordinates

Definition 4.5.1 (Basis). Let V be a vector space over F. A collection of vectors $B \subset V$ is called a *basis of* V if

1. B spans V, i.e. $V = \text{span}(B)$.
2. B is *linearly independent*.

Theorem 4.5.2. Let $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ be a basis for V. Then for any $\mathbf{v} \in V$, there exists *unique* $\lambda_1, \lambda_2, \dots, \lambda_n \in F$ such that $\mathbf{v} = \sum_{k=1}^{n} \lambda_k \mathbf{b}_k$.

Definition 4.5.3 (Coordinate Vector). Let $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis of a vector space V. Let $\mathbf{v} \in V$ have unique expression $\sum_{k=1}^{n} \lambda_k \mathbf{b}_k$. The *coordinate vector of* $\mathbf{v}$ or the *coordinates of* $\mathbf{v}$ *with respect to the basis* B is given by

$$[\mathbf{v}]_B = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Definition 4.5.4 (Dimension). The *dimension* of a vector space V is the *size of any basis of* V, that is, let B be a basis of a vector space V. Then

$$\dim V = |B|.$$

Theorem 4.5.5. Any two bases B_1 and B_2 of a vector space V satisfy

$$|B_1| = |B_2| = \dim V.$$

Theorem 4.5.6. For any matrix $M \in \mathrm{Mat}_{m\times n}(F)$, we have

$$\dim(\mathrm{Row}(M)) = \dim(\mathrm{Col}(M)).$$

4.6 Rank and Nullity

Definition 4.6.1 (Rank). We call $\dim(\mathrm{Row}(M))$ (equivalently, $\dim(\mathrm{Col}(M))$) the *rank of* M, denoted $\mathrm{rank}(M)$.

Definition 4.6.2 (Null Space). For $M \in \mathrm{Mat}_{m\times n}(F)$, the *null space of* M, denoted $\mathrm{Null}(M)$ or $M^{\perp}$, is given by

$$\mathrm{Null}(M) = \{\mathbf{x} \in F^n : M\mathbf{x} = \mathbf{0}\}.$$

That is, $\mathrm{Null}(M)$ is the solutions to the *homogeneous system of equations*

$$M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{0}.$$

Theorem 4.6.3. For a matrix M, to produce a basis of $\mathrm{Null}(M)$, we can

1. row reduce M to RREF;
2. produce solutions for $M\mathbf{x} = \mathbf{0}$ using "free variable" columns;
3. take the vectors corresponding to each parameter as our basis.

Definition 4.6.4 (Nullity). We call $\dim(\mathrm{Null}(M))$ (or equivalently $\dim(M^{\perp})$) the *nullity of* M and denote it as $\mathrm{nullity}(M)$.

Theorem 4.6.5 (Rank-Nullity Theorem). If $M \in \mathrm{Mat}_{m\times n}(F)$, then

$$\mathrm{rank}(M) + \mathrm{nullity}(M) = n.$$

4.7 Linear Maps between Vector Spaces

Definition 4.7.1. Let V, W be vector spaces over F. A map $L : V \to W$ is said to be *linear* if

- for all $\mathbf{u}, \mathbf{v} \in V$,
$$L(\mathbf{u} + \mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v});$$
- for all $\mathbf{v} \in V$ and $\lambda \in F$,
$$L(\lambda\mathbf{v}) = \lambda L(\mathbf{v}).$$

Corollary 4.7.2. Let $\mathcal{C}^\infty(\mathbb{R})$ be the real vector space of smooth functions from $\mathbb{R}$ to $\mathbb{R}$. Let the derivative map $D : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ be defined by

$$D(f) = f'.$$

Then $D^k = D \circ D \circ \cdots \circ D$ (k times), the kth derivative map, is linear.

Definition 4.7.3 (Linear differential operator). A *linear differential operator* $\mathcal{L} : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ is any linear combination of compositions of D. That is, linear maps $\mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ of the form

$$\mathcal{L}(f) = \sum_{i=0}^{k} \lambda_i D^i(f) = \lambda_k D^k(f) + \lambda_{k-1} D^{k-1}(f) + \cdots + \lambda_1 D(f) + \lambda_0 f$$

for some $\lambda_0, \dots, \lambda_k \in \mathbb{R}$.

Definition 4.7.4 (Isomorphisms of vector spaces). We say a linear map $L : V \to W$ is an *isomorphism* if it is a *bijection*, that is, it is *one-to-one (injective)* and *onto (surjective)*. We say two vector spaces V, W are *isomorphic* and denote this $V \cong W$ if there exists some isomorphism $L : V \to W$.

Theorem 4.7.5. Let V be a finite-dimensional vector space over F, with $\dim V = n$. Then $V \cong F^n$.

4.8 Change of Basis

Definition 4.8.1. Let V have basis $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, W have basis $D = \{\mathbf{d}_1, \dots, \mathbf{d}_m\}$, $L : V \to W$ be a linear map.
(i) The matrix of L with respect to B and D is the $m \times n$ matrix given by the columns $[L(\mathbf{b}_i)]_D$:

$$[L]_D^B = \begin{pmatrix} | & & | \\ [L(\mathbf{b}_1)]_D & \cdots & [L(\mathbf{b}_n)]_D \\ | & & | \end{pmatrix} \in \mathrm{Mat}_{m \times n}(F).$$

(ii) For any $\mathbf{v} \in V$, we have in coordinates

$$[L]_D^B [\mathbf{v}]_B = [L(\mathbf{v})]_D.$$

Theorem 4.8.2. Let V, W, and U be finite-dimensional vector spaces over F, with bases B, D and G respectively . Let $L_1 : V \to W$, $L_2 : W \to U$ be two linear maps. Then the composition $L_2 \circ L_1 : V \to U$ satisfies

$$[L_2 \circ L_1]_G^B = [L_2]_G^D [L_1]_D^B.$$

Theorem 4.8.3. Let V be a finite dimensional vector space over a field F. Let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, $D = \{\mathbf{d}_1, \ldots, \mathbf{d}_n\}$ be two bases of V. The *change of basis matrix* which converts B-coordinates to D-coordinates is given by

$$[\mathrm{id}]_D^B = \begin{pmatrix} | & & | \\ [\mathbf{b}_1]_D & \cdots & [\mathbf{b}_n]_D \\ | & & | \end{pmatrix}$$

so that, for any $\mathbf{v} \in V$, we have

$$[\mathrm{id}]_D^B [\mathbf{v}]_B = [\mathbf{v}]_D.$$

Theorem 4.8.4. Let V be a finite-dimensional vector space with bases B and D. Then

$$[\mathrm{id}]_B^D = ([\mathrm{id}]_D^B)^{-1}.$$

5 Orthogonal Transformations

5.1 Inner Product Spaces

Definition 5.1.1 (Inner Product). Let V be a vector space over $\mathbb{R}$. An *inner product* on V is a mapping

$$\langle , \rangle : V \times V \to \mathbb{R}$$

taking two vectors $\mathbf{v}, \mathbf{w} \in V$ to a number $\langle \mathbf{v}, \mathbf{w} \rangle \in \mathbb{R}$ satisfying the following axioms:

1. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ (commutativity)
2. $\langle \mathbf{v} + \mathbf{x}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{x}, \mathbf{w} \rangle$ (distributivity)
3. $\langle \lambda \mathbf{v}, \mathbf{w} \rangle = \lambda \langle \mathbf{v}, \mathbf{w} \rangle$ (pull out scalars)
4. $\forall \mathbf{v} \in V$, $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$
5. $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \iff \mathbf{v} = \mathbf{0}$

We say that V is an *inner product space* and write it as a pair $(V, \langle , \rangle)$.

Theorem 5.1.2 (Cauchy-Schwarz Inequality). Let $(V, \langle , \rangle)$ be an inner product space. Then for any $\mathbf{v}, \mathbf{w} \in V$, $|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \|\mathbf{w}\|$.

Definition 5.1.3 (Orthogonal). Let $(V, \langle , \rangle)$ be an inner product space. A collection of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset V$ is said to be *orthogonal* if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ for all $i \neq j$.

Definition 5.1.4 (Orthonormal). For an inner product space V, any collection of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} \subset V$ is said to be *orthonormal* if it is orthogonal and $\|\mathbf{u}_i\| = 1$ for $i = 1, \ldots, k$.

Definition 5.1.5 (Orthogonal/Orthonormal Basis). A collection of vector B in an inner product space V is called an *orthogonal (orthonormal)* basis if (i) B is a basis of V, and (ii) B is an orthogonal (orthonormal) collection.

Theorem 5.1.6. Let $(V, \langle , \rangle)$ be an inner product space and let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be an orthonormal basis. Let $\mathbf{v} \in V$, then

$$\mathbf{v} = \sum_{k=1}^{n} \langle \mathbf{v}, \mathbf{b}_k \rangle \mathbf{b}_k = \langle \mathbf{v}, \mathbf{b}_1 \rangle \mathbf{b}_1 + \langle \mathbf{v}, \mathbf{b}_2 \rangle \mathbf{b}_2 + \cdots + \langle \mathbf{v}, \mathbf{b}_n \rangle \mathbf{b}_n.$$

That is

$$[\mathbf{v}]_B = \begin{pmatrix} \langle \mathbf{v}, \mathbf{b}_1 \rangle \\ \vdots \\ \langle \mathbf{v}, \mathbf{b}_n \rangle \end{pmatrix}.$$

5.2 Subspaces and Orthogonal Projections

Definition 5.2.1. Let $(V, \langle , \rangle)$ be an inner product space. Let $W \subseteq V$ be a subspace, and let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be an *orthonormal basis of* W. Let $\mathbf{v} \in V$. Then the *projection of* $\mathbf{v}$ *onto* W is given by

$$\operatorname{proj}_W \mathbf{v} = \sum_{k=1}^{n} \operatorname{proj}_{\mathbf{b}_k}(\mathbf{v}) = \sum_{k=1}^{n} \langle \mathbf{v}, \mathbf{b}_k \rangle \mathbf{b}_k.$$

Theorem 5.2.2. For an inner product space $(V, \langle , \rangle)$, a subspace $W \subseteq V$, and $\mathbf{v} \in V$, then

(i) $\operatorname{proj}_W \mathbf{v}$ is the closest vector in W to $\mathbf{v}$.

(ii) $\mathbf{v} - \operatorname{proj}_W \mathbf{v}$ is orthogonal to all $\mathbf{w} \in W$.

Application 5.2.3 (Gram-Schmidt Process). Let W be an inner product space with basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Applying the following steps to obtain an *orthonormal basis*.

Step 1: Set $\mathbf{b}_1 = \hat{v}_1$, $W_1 = \operatorname{span}\{\mathbf{b}_1\}$.

Step 2: Set $\mathbf{x}_2 = \mathbf{v}_2 - \operatorname{proj}_{W_1}(\mathbf{v}_2)$, set $\mathbf{b}_2 = \hat{x}_2$, $W_2 = \operatorname{span}\{\mathbf{b}_1, \mathbf{b}_2\}$.

$\vdots$

Step k: Set $\mathbf{x}_k = \mathbf{v}_k - \operatorname{proj}_{W_{k-1}}(\mathbf{v}_k)$, set $\mathbf{b}_k = \hat{x}_k$, $W_k = \operatorname{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$.

$\vdots$

Step n: Set $\mathbf{x}_n = \mathbf{v}_n - \operatorname{proj}_{W_{n-1}}(\mathbf{v}_n)$, set $\mathbf{b}_n = \hat{x}_n$, $W = \operatorname{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$.

5.3 Symmetric Matrices and Jordan Canonical Forms

Definition 5.3.1 (Bilinear form). Let $A \in \operatorname{Mat}_n(\mathbb{R})$. The bilinear form on $\mathbb{R}^n$ associated to A is the map $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ mapping vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ to $\mathbf{x}^T A \mathbf{y}$.

Definition 5.3.2 (Symmetric Matrix). A matrix $A \in \operatorname{Mat}_n(\mathbb{R})$ is called *symmetric* if $A = A^T$.

Theorem 5.3.3. Any inner product on $\mathbb{R}^n$ is given by a bilinear form

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$$

where A is a symmetric matrix with real entries and all eigenvalues of A are strictly positive.

Theorem 5.3.4. Let $A \in \operatorname{Mat}_n(\mathbb{R})$ be symmetric, then all eigenvalues of A are real.

Theorem 5.3.5. Let $A \in \operatorname{Mat}_n(\mathbb{R})$ be symmetric, let λ_i and λ_j be eigenvalues of A such that $\lambda_i \neq \lambda_j$ with eigenvectors $\mathbf{v}$ and $\mathbf{w}$ respectively. Then $\mathbf{v} \cdot \mathbf{w} = 0$.

Theorem 5.3.6 (Spectral Theorem for Symmetric Matrices). Let $A \in \operatorname{Mat}_n(\mathbb{R})$ be a symmetric matrix, then

1. A is diagonalisable;
2. we can find a matrix P and D such that

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

   where the columns of P are orthonormal.

Theorem 5.3.7. Let P be any matrix with orthonormal columns. Then $P^{-1} = P^T$.

Corollary 5.3.8. Let A be a real symmetric matrix then A can be written as

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

for a diagonal matrix D and matrix P with orthonormal columns.

Definition 5.3.9 (Jordan Block). A Jordan block $J_{n,\lambda}$ for a scalar λ is an $n \times n$ upper triangular matrix of the form

$$J_{n,\lambda} = \begin{pmatrix} \lambda & 1 & & & & \\ & \lambda & 1 & & & \\ & & \lambda & & & \\ & & & \ddots & & \\ & & & & \lambda & 1 \\ & & & & & \lambda \end{pmatrix}$$

where each M_i is a square matrix.

Theorem 5.3.10 (Block diagonal matrix). A block diagonal matrix is a square matrix of the form

$$M = \begin{pmatrix} M_1 & & & \\ & M_2 & & \\ & & \ddots & \\ & & & M_n \end{pmatrix}$$

where each M_k is a square matrix.

Theorem 5.3.11 (Jordan Canonical Forms). Let $A \in \mathrm{Mat}_n(\mathbb{C})$. Then there are matrices P, J such that

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

and J is a block diagonal matrix of Jordan blocks $J_{k,\lambda}$ corresponding to eigenvalues λ of A. We call J the *Jordan Canonical Form of A*.

$$
\begin{pmatrix}
\lambda_1 & 1 & & & & & & & \\
& \lambda_1 & 1 & & & & & & \\
& & \lambda_1 & & & & & & \\
& & & \lambda_2 & 1 & & & & \\
& & & & \lambda_2 & & & & \\
& & & & & [\lambda_3] & & & \\
& & & & & & \ddots & & \\
& & & & & & & \lambda_n & 1 \\
& & & & & & & & \lambda_n
\end{pmatrix}
$$