Linear Algebra I
Summary of Lectures

Dr Nicholas Sedlmayr

Contents

Notation

iff	if and only if
$\Rightarrow$	if then
$\equiv$	defined as
$\therefore$	therefore
$\because$	because
$\square$	end of proof
$\mathbb{R}$	set of real numbers
$\mathbb{C}$	set of complex numbers
$\mathbb{I}$	identity matrix
$\forall$	the universal quantifier, for all
$\exists$	the existential quantifier, there exists
$\in$	is an element of
$\subset$	is a subset of

1 Matrices

- Definition of an $n \times m$ matrix, $\mathbf{A} = (a_{ij})$ with n row and m columns. Addition of matrices $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$.
  - Associativity, commutativity and existence of a zero matrix ($\mathbf{0}$) for addition.
- Definition of multiplication of a matrix by a scalar: $\lambda\mathbf{A} = (\lambda a_{ij})$.
- Definition: Matrix multiplication of an $n \times m$ matrix $\mathbf{A}$ and an $m \times k$ matrix $\mathbf{B}$ is an $n \times k$ matrix $\mathbf{C} = \mathbf{AB} = (c_{ij})$ where $c_{ij} = \sum_{r=1}^{n} a_{ir}b_{rj}$.
  - Associativity, existence of a zero matrix ($\mathbf{0}$) and an identity matrix $\mathbb{I}$, distributivity.
  - No commutativity!
- A matrix $\mathbf{A}$ can have a right inverse $\mathbf{AB} = \mathbb{I}$ and a left inverse $\mathbf{CA} = \mathbb{I}$.
  - Proposition 1.1: If a square matrix has either a left or right inverse then they have a unique inverse from both the left and right.
  - If a non-square matrix has both a left and right inverse then they are the same and the inverse is unique.
  - Proposition 1.2: If $\mathbf{A}$ and $\mathbf{B}$ are invertible square matrices then $\mathbf{AB}$ is also invertible and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
- The transpose of an $n \times m$ matrix is written as $\mathbf{A}^T$, which is an $m \times n$ matrix found by transposing the rows and columns of $\mathbf{A}$.
  - Proposition 1.3: For two $n \times n$ matrices $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ and $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$.
- Elementary row operations perform simple operations on the rows of an $n \times m$ matrix $\mathbf{A}$ and can be written as an $n \times n$ matrix $\mathbf{R}$ with the operation preformed by the multiplication $\mathbf{RA}$. ρ_i is used to refer to row i. There are three of them:
  - $\rho_j := \rho_j + \lambda\rho_i$, add λ copies of row i to row j;
  - $\rho_i := \lambda\rho_i$, multiple row i by λ with $\lambda \neq 0$; and
  - $\text{swap}(\rho_i, \rho_j)$, swap rows i and j.
- Echelon form: A matrix where each row starts with a sequence of zeros, and the number of zeros in this sequences increases from row to row from top to bottom until the final row is reached or all remaining rows are composed entirely of zeros is said to be in echelon form.
  - All matrices can be put into echelon form using elementary row operations.
- Rank: If $\mathbf{A}$ can be converted to the echelon form matrix $\mathbf{B}$ and $\mathbf{B}$ has k non-zero rows then the rank of $\mathbf{A}$ is $\operatorname{rk}\mathbf{A} = k$.

- An $n \times k$ matrix $\mathbf{A}$ with $n \leq k$ and $\operatorname{rk} \mathbf{A} = n$ can be converted to a matrix $\mathbf{B}$ where the left $n \times n$ block is the identity matrix using elementary row operations.

- Augmented matrix: If we have an $n \times m$ matrix $\mathbf{A}$ and an $n \times k$ matrix $\mathbf{A}$ the augmented matrix $(\mathbf{A}|\mathbf{B})$ is the $n \times (m+k)$ matrix created by writing $\mathbf{A}$ and $\mathbf{B}$ next to each other.

- Calculating the inverse: For an $n \times n$ matrix $\mathbf{A}$ of rank n we can calculate the inverse using the augmented matrix $(\mathbf{A}, \mathbb{I})$. Apply row operations, $\mathbf{R}$, such that $\mathbf{R}(\mathbf{A}, \mathbb{I}) = (\mathbb{I}, \mathbf{B})$, clearly $\mathbf{B} = \mathbf{R}$ is the left inverse of $\mathbf{A}$: $\mathbf{BA} = \mathbb{I}$.

- Solving systems of linear equations. A system of linear equations on the variables $x_1, x_2, \ldots x_n$ can be written as $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is a $k \times n$ matrix, $\mathbf{x} = (x_1, x_2, \ldots x_n)^T$ and $\mathbf{b}$ is a $k \times 1$ column vector. By putting the augmented matrix $(\mathbf{A}|\mathbf{b})$ into echelon form one can read off the solution.

  - Proposition 1.4: An equation $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is a $k \times n$ matrix, $\mathbf{x} = (x_1, x_2, \ldots x_n)^T$ and $\mathbf{b}$ is a $k \times 1$ column vector has at least one solution iff $\operatorname{rk} \mathbf{A} = \operatorname{rk}(\mathbf{A}|\mathbf{b})$. It has exactly one solution if $\operatorname{rk} \mathbf{A} = n$

- Trace: This operation is defined only for square matrices. A square $n \times n$ matrix $\mathbf{A} = (a_{ij})$ has a trace $\operatorname{tr} \mathbf{A} = \sum_{i=1}^{n} a_{ii}$. I.e. it is the sum over all diagonal entries of the matrix. It has the following properties:

  - For two square $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, $\operatorname{tr} \mathbf{AB} = \operatorname{tr} \mathbf{BA}$.
  - If $\mathbf{P}$ is a square $n \times n$ invertible matrix and $\mathbf{A}$ is a square $n \times n$ matrix then $\operatorname{tr} \mathbf{P}^{-1}\mathbf{AP} = \operatorname{tr} \mathbf{A}$.
  - $\operatorname{tr}(\mathbf{A} + \mathbf{B}) = \operatorname{tr} \mathbf{A} + \operatorname{tr} \mathbf{B}$.
  - $\operatorname{tr} \mathbf{A}^T = \operatorname{tr} \mathbf{A}$.
  - For a scalar λ then $\operatorname{tr} \lambda \mathbf{A} = \lambda \operatorname{tr} \mathbf{A}$.

- Determinant: This operation is defined only for square matrices and we will define it via "expansion by the first row". For a 1×1 matrix $\mathbf{A} = (a_1 1)$ we have $\det \mathbf{A} = |\mathbf{A}| = a_{11}$. For an $n \times n$ matrix $\mathbf{A} = (a_{ij})$ it is

$$\begin{aligned}
\det \mathbf{A} &= \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \\
&= a_{11} \begin{vmatrix} a_{22} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n3} & \ldots & a_{nn} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n3} & \ldots & a_{nn} \end{vmatrix} \\
&\quad + \ldots + (-1)^{n-1} a_{1n} \begin{vmatrix} a_{21} & a_{22} & \ldots & a_{2,n-1} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{n,n-1} \end{vmatrix}.
\end{aligned}$$

  - Proposition 1.5: For two $n \times n$ square matrices $\mathbf{A}$ and $\mathbf{B}$, $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$. It can be proved using the following lemmas.

  - Lemma 1.5.1: Any square matrix $\mathbf{A}$ can be written as the product $\mathbf{S}_1\mathbf{S}_2\ldots\mathbf{S}_k$ of 'generalized' elementary row operations $\rho_i := \lambda\rho_i$ and $\rho_i := \rho_i + \lambda\rho_j$ where λ can be zero.
  - Lemma 1.5.2: The determinant of a matrix whose top two rows are identical is zero.
  - Lemma 1.5.3:

$$\begin{vmatrix} a_{11}+\lambda b_{11} & a_{12}+\lambda b_{12} & \ldots & a_{1n}+\lambda b_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} + \lambda \begin{vmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

  - Lemma 1.5.4: The matrix of any generalized elementary row operation can always be written as one of $\mathbf{S}_i\mathbf{T}_j\mathbf{R}\mathbf{T}_j\mathbf{S}_i$, $\mathbf{T}_j\mathbf{R}\mathbf{T}_j$, $\mathbf{S}_i\mathbf{R}\mathbf{S}_i$, or $\mathbf{R}$, where $\mathbf{R}$ is the matrix of a generalized row operation acting only on rows 1 and 2, and $\mathbf{S}_i$, $\mathbf{T}_j$ are the matrices of swap(ρ_1, ρ_i) and swap(ρ_2, ρ_j) respectively.
  - Lemma 1.5.5: If $\mathbf{R}$ is the matrix of the swap operation swap(ρ_i, ρ_j) then $\det \mathbf{RA} = -\det \mathbf{A}$.

- Proposition 1.6: For an $n \times n$ matrix $\mathbf{A}$ the following are equivalent:
  (a) $\mathbf{A}^{-1}$ exists,
  (b) $\det \mathbf{A} \neq 0$, and
  (c) $\operatorname{rk} \mathbf{A} = n$.

- Proposition 1.7: For an $n \times n$ matrix $\mathbf{A}$:
  (a) $\det(\mathbf{A}^T) = \det \mathbf{A}$,
  (b) If $\mathbf{A}$ is invertible $\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1}$,
  (c) $\det \mathbb{I} = 1$, and
  (d) if $\mathbf{A}$ has a row (or column) entirely composed of zeros $\det \mathbf{A} = 0$.

- Proposition 1.8: The determinant of an upper triangular matrix $\mathbf{A}$ is equal to the product of its diagonal entries.

- Proposition 1.9: The determinants of row operations are:
  (a) If $\mathbf{A}_{ij,\lambda}$ is the matrix for $\rho_i := \rho_i + \lambda\rho_j$ then $\det \mathbf{A}_{ij,\lambda} = 1$,
  (b) If $\mathbf{T}_{i,\lambda}$ is the matrix for $\rho_i := \lambda\rho_i$ then $\det \mathbf{T}_{i,\lambda} = \lambda$, and
  (a) If $\mathbf{S}_{ij}$ is the matrix for swap(ρ_i, ρ_j) then $\det \mathbf{S}_{ij} = -1$.

- Proposition 1.10: If we swap any two rows in a determinant then the determinant changes sign, see Lemma 1.5.5. It follows that if a matrix has two identical rows then its determinant is zero. (This is equally true for a matrix with two identical columns.)

- Proposition 1.11: A determinant can be expanded along any row (or column). The sign associated with any entry a_{ij} is $(-1)^{i+j}$.

- Minors and cofactors: Let $\mathbf{A} = (a_{ij})$ be an $n \times n$ matrix and let b_{ij} be the determinant of the $(n-1) \times (n-1)$ matrix obtained form $\mathbf{A}$ by deleting row i and column j. Furthermore let $c_{ij} = (-1)^{i+j} b_{ij}$. Then
  - $\mathbf{B} = (b_{ij})$ is the matrix of minors of $\mathbf{A}$,
  - $\mathbf{C} = (c_{ij})$ is the matrix of cofactors of $\mathbf{A}$, and
  - $\operatorname{adj} \mathbf{A} = \mathbf{C}^T$ is the adjugate matrix of $\mathbf{A}$.

2 Vector Spaces

- Definition 2.1: A vector space. A vector space V over a field F (see definition 2.3) is a set containing:

  - a special zero vector $\mathbf{0}$;
  - an operation of addition of two vectors $\mathbf{u} + \mathbf{v} \in V$, for $\mathbf{u}, \mathbf{v} \in V$; and
  - multiplication of a vector V with a number $\lambda \in F$ with $\lambda \mathbf{v} \in V$.

  The vector space must be closed under both of these operations and must satisfy the following laws $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $\lambda, \mu \in F$:

  (1) associativity $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;
  (2) commutativity $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
  (3) $\mathbf{u} + \mathbf{0} = \mathbf{u}$;
  (4) $\mathbf{v} + (-1)\mathbf{v} = \mathbf{0}$;
  (5) $\lambda(\mu \mathbf{u}) = (\lambda\mu)\mathbf{v}$;
  (6) distributivity $\lambda(\mathbf{u} + \mathbf{v}) = \lambda\mathbf{u} + \mu\mathbf{v}$; and
  (7) distributivity $(\lambda + \mu)\mathbf{u} = \lambda\mathbf{u} + \mu\mathbf{u}$.

- Proposition 2.2: $\forall \mathbf{v} \in V$ and $\forall \lambda \in F$:

  (a) $\mathbf{v} = 1\mathbf{v}$;
  (b) $0\mathbf{v} = \mathbf{0}$; and
  (c) $\lambda\mathbf{0} = \mathbf{0}$.

- Definition 2.3: A field is a set F containing distinct elements 0 and 1 with two binary operations $+$ and $\cdot$ satisfying the axioms $\forall a, b, c \in F$:

  (1) $a + b = b + a$;
  (2) $(a + b) + c = a + (b + c)$;
  (3) $a + 0 = a$;
  (4) $\forall a \ \exists -a$ such that $a + (-a) = 0$;
  (5) $a \cdot b = b \cdot a$;
  (6) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;
  (7) $a \cdot 1 = a$;
  (8) $\forall a \neq 0 \ \exists a^{-1}$ such that $a \cdot a^{-1} = 1$; and
  (9) $a \cdot (b + c) = a \cdot b + a \cdot c$;

  If a field F is finite its order is the number of elements in F.

- Note that as a field also satisfies all axioms of a vector space a field F is also itself a vector space $V = F$ over the field F and all properties of a vector space apply.

- Theorem 2.4: For each prime p and each positive integer n, there is a unique field of order p^n. Additionally, every finite field is of this form.

- Definition 2.5: Given a vector space V over F, a subspace of V is a subset $W \subset V$ which contains the zero vector of V and is closed under the operations of addition and scalar multiplication.

- Lemma 2.5.1: Let $W \subset V$ be nonempty, where V is a vector space over F. Then W is a subspace of V iff $\mathbf{v} + \lambda\mathbf{u} \in W$ for each $\mathbf{v}, \mathbf{u} \in W$ and each scalar λ.

- Definition 2.6: Given a vector space V over F, and given a subset of V $A = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots \mathbf{u}_n\}$,

$$W = \{\lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \lambda_3\mathbf{u}_3 + \dots \lambda_n\mathbf{u}_n : \lambda_1, \lambda_2, \dots \lambda_n \in F\}$$

  is the subspace of V spanned by A. The elements of W are called linear combinations of vectors from A and the subspace W is denoted as span A.

- Definition 2.7: If A is an infinite subset of V, where V is a vector space over F, we define span A to be the set of all linear combinations of finite subsets of A.

- Definition 2.8: A set $A \subset V$ of vectors in a vector space V over F is linearly dependent if there are $n \in \mathbb{N}$ vectors $a_1, a_2, \dots a_n$ and scalars $\lambda_1, \lambda_2, \dots \lambda_n$ not all zero such that

$$\lambda_1 a_1 + \lambda_2 a_2 + \dots \lambda_n a_n = 0\,.$$

  Otherwise A is linearly independent.

  - For a finite set $A = \{a_1, a_2, \dots a_n\}$ it is linearly independent iff $\forall$ scalars $\lambda_1, \lambda_2, \dots \lambda_n \in F$

$$\lambda_1 a_1 + \lambda_2 a_2 + \dots \lambda_n a_n = 0 \Rightarrow \lambda_1 = \lambda_2 = \dots \lambda_n = 0.$$

  - If A is infinite it is linearly independent iff every subset of A is linearly independent.
  - By convention the empty set is linearly independent.

- Proposition 2.9: Suppose $\mathbf{A} = \{a_1, a_2, \dots a_n\} \subset V$ is linearly independent, where V is a vector space over F. Suppose also that $v \in V$ and there are scalars $\lambda_1, \dots \lambda_n$ and $\mu_1, \dots \mu_n$ such that

$$v = \lambda_1 a_1 + \lambda_2 a_2 + \dots \lambda_n a_n$$

  and

$$v = \mu_1 a_1 + \mu_2 a_2 + \dots \mu_n a_n$$

  then $\lambda_1 = \mu_1$, $\lambda_2 = \mu_2$,$\dots \lambda_n = \mu_n$.

- Definition 2.10: A basis of a vector space V is a linearly independent set $B \subset V$ which spans V.

- Theorem 2.11: Let V be a vector space over F, and let $B \subset V$ be linearly independent. Then there is a basis B' of V with $B \subset B'$.

- Theorem 2.12: Suppose span $A = V$ and $B \subset V$ is linearly independent. Then there is a basis B' of V with $B \subset B' \subset A \cup B$.

- Lemma 2.13: The "exchange lemma". Suppose $a_1, a_2, \dots a_n, b$ are vectors in a vector space V, and suppose that

  $$b \in \text{span}(a_1, a_2, \dots a_{n-1}, a_n)$$

  but

  $$b \notin \text{span}(a_1, a_2, \dots a_{n-1}),$$

  then

  $$a_n \in \text{span}(a_1, \dots, a_{n-1}, b).$$

  If in addition $\{a_1, a_2, \dots a_n\}$ is linearly independent then so is $\{a_1, a_2, \dots a_{n-1}, b\}$.

- Theorem 2.14: Suppose S and B are both bases of a vector space V over F and. Then A and B have the same number of elements.

- Definition 2.15: The number of elements of a basis of a vector space V over F is called the dimension of V and is written as $\dim V$.

- Corollary 2.16: If V is a vector space over F and $U \subset V$ is a subspace of V then $\dim U \leq \dim V$. If, additionally, $\dim V$ is finite and $U \neq V$ then $\dim U < \dim V$.

- Corollary 2.17: Suppose that V is a vector space over F, $\dim V$ is finite, and $U \subset V$ is a subspace of V with $\dim U = \dim V$, then $U = V$.

- The coordinates of a vector $v \in V$, with V a vector space over F, with respect to an *ordered* basis $B = \{v_1, v_2, \dots v_n\}$ are $(\lambda_1, \lambda_2, \dots, v_n)^T$ where

  $$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots \lambda_n v_n.$$

- Definition 2.18: Two vector spaces V and W, both over the same field F, are isomorphic if there is a bijection $f : V \to W$ such that

  $$f(u + v) = f(u) + f(v)$$

  and

  $$f(\lambda v) = \lambda f(v),$$

  $\forall u, v \in V$ and $\forall \lambda \in F$. The bijection is said to be an isomorphism from V to W and we write $V \cong W$ or $f : V \xrightarrow{\sim} W$.

- Theorem 2.19: Suppose V is a vector space over $\mathbb{R}$ with finite dimension $n \geq 0$. Then $V \cong \mathbb{R}^n$ as real vector spaces. Similarly if V is a vector space over $\mathbb{C}$ with finite dimension $n \geq 0$. Then $V \cong \mathbb{C}^n$ as complex vector spaces.

3 Inner Product Spaces

- Definition 3.1: If V is a vector space over $\mathbb{R}$, then an inner product on V is a map $(\langle|\rangle)$ from $V \times V$ to $\mathbb{R}$ with the following properties:

  (a) Symmetry: $\langle v|w\rangle = \langle w|v\rangle \ \forall v, w \in V$.

  (b) Linearity: $\langle u|\lambda v+\mu w\rangle = \lambda\langle u|v\rangle+\mu\langle u|w\rangle \ \forall u, v, w \in V$ and $\forall \lambda, \mu \in \mathbb{R}$.

  (c) Positive definiteness:

    (i) $\langle v|v\rangle \geq 0 \ \forall v \in V$, and

    (ii) $\langle v|v\rangle = 0$ iff $v = 0$.

  As the inner product is linear with respect to both variables it is sometimes called bilinear.

- Definition 3.2: A finite dimensional vector space over $\mathbb{R}$ with an inner product defined is called a Euclidean space.

- Definition 3.3: The norm (or length) of a vector$v$ is written as $||v|$ and defined by

$$||v|| = \sqrt{\langle v|v\rangle},$$

  the positive square root of the inner product of v with itself. The distance between two vectors v and w is written as $d(v, w)$ and is $d(v, w) = ||v-w||$.

- Proposition 3.4: $\forall v \in V$, where V is a Euclidean space, and $\forall \lambda \in \mathbb{R}$ then $||\lambda v|| = |\lambda| \cdot ||v||$.

- Proposition 3.5: The "Cauchy-Schwarz inequality" says that $\forall v, w \in V$, where V is a Euclidean space, then

$$|\langle v|w\rangle| \leq ||v|| \cdot ||w|| \, .$$

- Proposition 3.6: The "triangle inequality" says that $\forall v, w \in V$, where V is a Euclidean space, then

$$||v + w|| \leq ||v|| + ||w|| \, .$$

- Definition 3.7: If V is a Euclidean space, and $v, w \in V$, then v and w are said to be orthogonal if $\langle v|w\rangle = 0$. If both v and w are nonzero, then the angle between v and w is defined to be θ, $0 \leq \theta \leq \pi$ and

$$\cos\theta = \frac{\langle v|w\rangle}{||v|| \cdot ||w||} \, .$$

- Definition 3.8: If V is a vector space over $\mathbb{C}$, then a map $(\langle|\rangle)$ from $V \times V$ to $\mathbb{C}$ is an inner product if the following are true:

  (a) Conjugate-Symmetry: $\langle v|w\rangle = \overline{\langle w|v\rangle} \ \forall v, w \in V$.

  (b) Linearity: $\langle u|\lambda v+\mu w\rangle = \lambda\langle u|v\rangle+\mu\langle u|w\rangle \ \forall u, v, w \in V$ and $\forall \lambda, \mu \in \mathbb{R}$.

  (c) Positive definiteness:

    (i) $\langle v|v\rangle \geq 0 \ \forall v \in V$, and

(ii) $\langle v|v\rangle = 0$ iff $v = 0$.

This inner product is sometimes called sesquilinear.

- Definition 3.9: A finite dimensional vector space over $\mathbb{C}$ with an inner product define is called a unitary space.
- A vector space over $\mathbb{R}$ or $\mathbb{C}$, of any dimension, we will refer to as an inner product space.
- Definition 3.10: The norm (or length) of a vector $v \in V$, with V a vector space over $\mathbb{C}$,is written as $||v||$ and defined by

$$||v|| = \sqrt{\langle v|v\rangle},$$

the positive square root of the inner product of v with itself. The distance between two vectors v and w is written as $d(v, w)$ and is $d(v, w) = ||v-w||$.

- Proposition 3.11: $\forall v \in V$, where V is a unitary space, and $\forall \lambda \in \mathbb{R}$ then $||\lambda v|| = |\lambda| \cdot ||v||$.
- Proposition 3.12: The "Cauchy-Schwarz inequality" says that $\forall v, w \in V$, where V is a unitary space, then

$$|\langle v|w\rangle| \leq ||v|| \cdot ||w|| \,.$$

- Proposition 3.13: The "triangle inequality" says that $\forall v, w \in V$, where V is a unitary space, then

$$||v + w|| \leq ||v|| + ||w|| \,.$$

- Definition 3.14: A bilinear form on a real vector space V is a map $F : V \times V \to \mathbb{R}$ which $\forall u, v, w \in V$ and $\forall \alpha, \beta \in \mathbb{R}$ satisfies
  (a) $\langle \alpha u + \beta v|w\rangle = \alpha\langle u|w\rangle + \beta\langle v|w\rangle$, and
  (b) $\langle u|\alpha v + \beta w\rangle = \alpha\langle u|v\rangle + \beta\langle u|w\rangle$.
- Definition 3.15: A bilinear form on a real vector space V is symmetric if
  (c) $F(u, v) = F(v, u)$ $\forall u, v \in V$.
- Definition 3.16: The matrix $\mathbf{A} = (a_{ij})$ with $a_{ij} = F(e_i, e_j)$ is called the 'matrix of the bilinear form F with respect to the ordered basis $e_1, e_2, \ldots e_n$ of V'. If F is symmetric then $\mathbf{B}$ is symmetric.
- Proposition 3.17: Suppose V is a real vector space with ordered basis $e_1, e_2, \ldots e_n$ and F is a bilinear form defined on V, with matrix $\mathbf{A}$ with respect to this basis. Then for any vectors $v, w \in V$ and their corresponding coordinate forms $\mathbf{v} = (v_1, v_2, \ldots v_n)^T$ and $\mathbf{w} = (w_1, w_2, \ldots w_n)^T$ with respect to the same basis we have

$$F(v, w) = \mathbf{v}^T \mathbf{A} \mathbf{w} \,.$$

- The base change matrix from a basis $e_1, e_2, \ldots e_n$ to $f_1, f_2, \ldots f_n$ is $\mathbf{P} = (p_{ij})$ where $f_i = \sum_{k=1}^{n} p_{ki} e_k$.

- Proposition 3.18: (The base change formula) Given two ordered bases of a Euclidean space V, $e_1, e_2, \dots e_n$ and $f_1, f_2, \dots f_n$ related by the base change matrix $\mathbf{P}$ from basis $e_1, e_2, \dots e_n$ to $f_1, f_2, \dots f_n$, suppose $\mathbf{A}$ and $\mathbf{B}$ are the matrices of the inner product with respect to $e_1, e_2, \dots e_n$ and $f_1, f_2, \dots f_n$. Then $\mathbf{B} = \mathbf{P}^T \mathbf{A} \mathbf{P}$.

- Definition 3.19: A sesquilinear form on a complex vector space V is a map $F : V \times V \to \mathbb{C}$ which $\forall u, v, w \in V$ and $\forall \alpha, \beta \in \mathbb{C}$ satisfies

  (a) $\langle \alpha u + \beta v | w \rangle = \bar{\alpha} \langle u | w \rangle + \bar{\beta} \langle v | w \rangle$, and

  (b) $\langle u | \alpha v + \beta w \rangle = \alpha \langle u | v \rangle + \beta \langle u | w \rangle$.

- Definition 3.20: A sesquilinear form on a complex vector space V is conjugate-symmetric if

  (c) $F(u, v) = \overline{F(v, u)}$ $\forall u, v \in V$.

- Definition 3.21: The matrix $\mathbf{B} = (a_{ij}$ with $a_{ij} = F(e_i, e_j)$ is called the 'matrix of the bilinear form F with respect to the ordered basis $e_1, e_2, \dots e_n$ of the complex vector space V'. If F is conjugate-symmetric then $\mathbf{B}$ is conjugate-symmetric, i.e. $\bar{\mathbf{B}}^T = \mathbf{B}$.

- Proposition 3.22: Suppose V is a complex inner product space with ordered basis $e_1, e_2, \dots e_n$ and F is a sesquilinear form defined on V, with matrix $\mathbf{A}$ with respect to this basis. Then for any vectors $v, w \in V$ and their corresponding coordinate forms $\mathbf{v} = (v_1, v_2, \dots v_n)^T$ and $\mathbf{w} = (w_1, w_2, \dots w_n)^T$ with respect to the same basis we have

$$F(v, w) = \overline{\mathbf{v}}^T \mathbf{A} \mathbf{w} \,.$$

- Proposition 3.23: (The base change formula) Given two ordered bases of a complex inner product space V, $e_1, e_2, \dots e_n$ and $f_1, f_2, \dots f_n$ related by the base change matrix $\mathbf{P}$ from basis $e_1, e_2, \dots e_n$ to $f_1, f_2, \dots f_n$, suppose $\mathbf{A}$ and $\mathbf{B}$ are the matrices of the inner product with respect to $e_1, e_2, \dots e_n$ and $f_1, f_2, \dots f_n$. Then $\mathbf{B} = \overline{\mathbf{P}}^T \mathbf{A} \mathbf{P}$.

4 Orthogonal Bases

- Definition 4.1: Two vectors v and w in an inner product space are orthogonal if $\langle v|w\rangle$=0. The set of vectors $\{v_1, v_2, \ldots\}$ is said to be orthogonal, and the vectors $v_1, v_2 \ldots$ in the set are said to be mutually orthogonal if each pair of distinct vectors v_i, v_l with $i \neq l$ are said to be an orthogonal pair, $\langle v_i|v_l\rangle = 0$.

- Definition 4.2: A set $\{w_1, w_2, \ldots\}$ of vectors in an inner product space is said to be orthonormal if $\langle w_i|w_j\rangle = \delta_{ij}$. If the orthonormal set is a basis then it is called an orthonormal basis.

- Proposition 4.3: If V is an inner product space over $\mathbb{R}$ or $\mathbb{C}$, $v_1, v_2, \ldots v_n \in V$, $v_i \neq 0\, \forall i = 1 \ldots n$, and the v_i are mutually orthogonal then $\{v_1, v_2, \ldots v_n\}$ is a linearly independent set.

- Lemma 4.4: If u, v are any two vectors in an inner product space V with $v \neq 0$ then the vector
$$w = u - \frac{\langle v|u\rangle}{\langle v|v\rangle} v$$
is orthogonal to v.

- Lemma 4.5: If V is an inner product space, $u, v_1, v_2, \ldots v_k \in V$ and $v_1, v_2, \ldots v_k$ are mutually orthogonal non-zero vectors then
$$w = u - \sum_{i=1}^{k} \frac{\langle v_i|u\rangle}{\langle v_i|v_i\rangle} v_i$$
is orthogonal tro $v_1, v_2, \ldots v_k$.

- Theorem 4.6: (The Gram-Schmidt process) If $\{v_1, \ldots v_n\}$ is a basis of a finite dimensional inner product space V, then $\{w_1, \ldots w_n\}$ obtained by
$$\begin{aligned} w_1 &= v_1 \\ w_2 &= v_2 - \tfrac{\langle w_1|v_2\rangle}{\langle w_1|w_1\rangle} w_1 \\ &\vdots \\ w_k &= v_k - \textstyle\sum_{i=1}^{k-1} \tfrac{\langle w_i|v_k\rangle}{\langle w_i|w_i\rangle} w_i \\ &\vdots \end{aligned}$$
is an orthogonal basis of V.

- Corollary 4.7: Any finite dimensional inner product space V has an orthonormal basis.

- Definition 4.8: Two real vector spaces V, W with forms $F : V \times V \to \mathbb{R}$ and $G : W \times W \to \mathbb{R}$ respectively are isomorphic if there is a bijection $f : V \to W$ such that
$$\begin{aligned} f(u + v) &= f(u) + f(v)\,, \\ f(\lambda v) &= \lambda f(v) \text{ and} \\ F(u, v) &= G(f(u), g(v))\,, \end{aligned}$$

$\forall u, v \in V$ and $\forall \lambda \in \mathbb{R}$.
Similarly two complex vector spaces V, W with forms $F : V \times V \to \mathbb{C}$ and $G : W \times W \to \mathbb{C}$ respectively are isomorphic if there is a bijection $f : V \to W$ such that

$$\begin{aligned} f(u+v) &= f(u) + f(v)\,, \\ f(\lambda v) &= \lambda f(v) \text{ and} \\ F(u,v) &= G(f(u), g(v))\,, \end{aligned}$$

$\forall u, v \in V$ and $\forall \lambda \in \mathbb{C}$.

- Corollary 4.9: Let V be a Euclidean vector space of dimension n. Then V is isomorphic to $\mathbb{R}^n$ with the standard inner product as an inner product space. Similarly each unitary vector space V of dimension n is isomorphic to $\mathbb{C}^n$ with the standard inner product as an inner product space.

- Proposition 4.10: Suppose that $\{e_1, e_2, \ldots e_n\}$ is an orthonormal basis of a Euclidean space V. Then for any $v \in V$:

$$v = \sum_{i=1}^{n} \langle e_i | v \rangle e_i\,.$$

- Proposition 4.11: (Pythagoras' theorem) Suppose $e_1, e_2, \ldots e_n$ is an orthonormal basis of a Euclidean space V. Then for all $v \in V$

$$||v||^2 = \sum_{i=1}^{n} \langle e_i | v \rangle^2\,.$$

- Corollary 4.12: (Parseval's identity) If $e_1, e_2, \ldots e_n$ is an orthonormal basis of a Euclidean space V, and $v, w \in V$, then

$$\langle v | w \rangle = \sum_{i=1}^{n} \langle v |\, \mathrm{e}_i\, \rangle \langle e_i | w \rangle\,.\,.$$

- Proposition 4.13: (Bessel's inequality) If $e_1, e_2, \ldots e_k$ is an orthonormal set of vectors in a real inner product space V, and $v \in V$, then

$$\sum_{i=1}^{k} \langle e_i | v \rangle^2 \leq ||v||^2\,.$$

- Proposition 4.14: If $e_1, e_2, \ldots e_n$ is an orthonormal basis of a complex inner prodoct space V, and $v, w \in V$, then:
  (a) $v = \sum_{i=1}^{n} \langle e_i | v \rangle e_i$,
  (b) $||v||^2 = \sum_{i=1}^{n} \langle e_i | v \rangle^2$, (Pythagoras' theorem) and
  (c) $\langle v | w \rangle = \sum_{i=1}^{n} \langle v |\, \mathrm{e}_i\, \rangle \langle e_i | w \rangle = \sum_{i=1}^{n} \overline{\langle e_i | v \rangle} \langle e_i | w \rangle$ (Parseval's identity).

- Proposition 4.15: (Bessel's inequality) If $e_1, e_2, \ldots e_k$ is an orthonormal set of vectors in a complex inner product space V, and $v \in V$, then

$$\sum_{i=1}^{k} |\langle e_i | v \rangle|^2 \leq ||v||^2\,.$$

- Definition 4.16: If U and W are subspaces of a vector space V then the sum of U and W is defined as

$$U + W = \{u + w : u \in U, w \in W\}.$$

- Proposition 4.17: $U + W$ is a subspace of a vector space V if U and W are subspaces of V.

- The union of two sets is $A \cup B = \{x : x \in A \vee x \in B\}$. I.e. the elements in either A or B. The intersection of two sets is $A \cap B = \{x : x \in A \wedge x \in B\}$. I.e. the elements in both A or B.

- Definition 4.18: If V is a vector space and U is a subspace of V, then W is called a complement to U in V if

  (a) W is a subspace of V,

  (b) $V = U + W$, and

  (c) $U \cap W = \{0\}$.

  When these conditions are met we write $V = U \oplus W$, and say that V is the direct sum of U and W.

- Definition 4.19: If V is an inner product space and U is a subspace of V we define

$$U^{\perp} = \{v \in V : \langle u|v\rangle = 0 \, \forall u \in U\}.$$

  This is called the orthogonal complement of U in V, or "U perp" for short.

- Lemma 4.20: If V is an inner product space, U is a subspace of V, and U has a basis $\{u_1, \dots u_k\}$, then

$$U^{\perp} = \{v \in V : \langle u_i|v\rangle = 0 \, \forall i = 1, \dots k\}.$$

- Proposition 4.21: If V is an inner product space, and U is a finite dimensional subspace of V, then

  (a) $U^{\perp}$ is a subspace of V,

  (b) $U \cap U^{\perp} = \{0\}$, and

  (c) $U + U^{\perp} = V$.

- Proposition 4.22: If $V = U \oplus W$ then $\dim(V) = \dim(U) + \dim(W)$.

- Corollary 4.23: If V is a finite dimensional inner product space, and U is a subspace of V, then

  (a) $\dim(U) + \dim(U^{\perp}) = \dim(V)$, and

  (b) $(U^{\perp})^{\perp} = U$.

5 Linear Transformations

- Definition 5.1: If V and W are two vectors spaces over the same field F, then a linear transformation from V to W (also called a linear map or homomorphism) is a map $f : V \to W$ satisfying

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v)\,, \quad \forall\, u, v \in V \text{ and } \forall\, \lambda, \mu \in F\,.$$

  The space of V is called the domain of F and the space of W is called the co-domain.

- Lemma 5.2: A linear transformation $f : V \to W$ satisfies
  (a) $f(0) = 0$,
  (b) $f(\lambda u) = \lambda u$,
  (c) $f(-u) = -f(u)$,
  (d) $f(u + v) = f(u) + f(v)$, and
  (e) $f(\sum_{i=1}^{n} \lambda_i u_i) = \sum_{i=1}^{n} \lambda_i f(u_i)$.

- Definition 5.3: Given $f : V \to W$ as in definition 5.1, the image (or range) of f is $\{f(v) : v \in V\}$. This is written as $f(V)$ or $\mathrm{im}(f)$.
  The kernel (or nullspace) of f is $\{v \in V : f(v) = 0\}$, written $\ker(f)$

- Proposition 5.4: If $f : V \to W$ is a linear transformation, then $\mathrm{im}(f)$ is a subspace of W and $\ker(f)$ is a subspace of V.

- Proposition 5.5: A linear transformation $f : V \to W$ is injective iff $\ker(f)$ is the zero subspace $\{0\}$ of V.

- Definition 5.6: The rank of f is the dimension of $\mathrm{im}(f)$, written $\mathrm{r}(f)$.
  The nullity of f is the dimension of $\ker(f)$, written $\mathrm{n}(f)$.

- Theorem 5.7: The rank-nullity formula. If $f : V \to W$ is a linear transformation then

$$\mathrm{r}(f) + \mathrm{n}(f) = \dim(V)\,.$$

- Proposition 5.8: If $f : V \to W$ is a linear transformation of finite dimensional vector spaces V, W over the same field F then
  (a) f is injective iff $n(f) = 0$, and
  (b) f is surjective iff $r(f) = \dim(W)$.

- Corollary 5.9: If $f : V \to W$ is a linear transformation of finite dimensional vector spaces V, W over the same field F then
  (a) f is injective iff $r(f) = \dim(V)$, and
  (b) f is surjective iff $n(f) = \dim(V) - \dim(W)$.

- Let $f_{\mathbf{A}} : F^n \to F^m$ be $f_{\mathbf{A}}(\mathbf{v}) = \mathbf{A}\mathbf{v}$ where $\mathbf{v} \in F^n$ and $\mathbf{A}$ is an $m \times n$ matrix over the field F. Then
  (a) $\mathrm{im}(\mathbf{A}) = \{\mathbf{A}\mathbf{v} : \mathbf{v} \in F^n\}$,
  (b) $\ker(\mathbf{A}) = \{\mathbf{v} \in F^n : \mathbf{A}\mathbf{x} = 0\}$, and

(c) $r(\mathbf{A})$ and $n(\mathbf{A})$ are the rank and nullity of $\mathbf{A}$, i.e. the dimensions of $\text{im}(\mathbf{A})$ and $\ker(\mathbf{A})$ respectively.

- For a linear transformation the codomain is a vector space and therefore operations of addition and scalar multiplication must be defined. Any set S of functions from any set X to a vector space W, such that S is closed under addition and scalar multiplication forms a vector space.

- Let $\mathcal{L}[V, W]$ be the set of all linear transformations from a vector space V to a vector space W over the same field F. The zero element in $\mathcal{L}[V, W]$ is the map that takes every element to the zero in W.

- $\mathcal{L}[V, V]$ allows additional operations, 'composition of functions', $(f \cdot g)(x) = f(g(x))$. $\forall f, g, h \in \mathcal{L}[V, V]$ and $\forall \lambda, \mu \in F$ we have

  (1) $(f + g) + h = f + (g + h)$,
  (2) $f + g = g + f$,
  (3) $0 + f = f + 0$ (there is a zero element),
  (4) $\lambda(\mu f) = (\lambda\mu)f$,
  (5) $(\lambda + \mu)f = \lambda f + \mu f$,
  (6) $0f = 0$,
  (7) $f + (-1)f = 0$,
  (8) $(f \cdot g) \cdot h = f \cdot (g \cdot h)$,
  (9) $\mathbb{I} \cdot f = f \cdot \mathbb{I} = f$ ($\mathbb{I}$ is the identity map $\mathbb{I}(x) = x$),
  (10) $f \cdot (g + h) = f \cdot g + f \cot h$, and
  (11) $((g + h)) \cdot f = g \cdot f + h \cdot f$.

- If $f : V \to W$ is a linear map, $v_1, v_2, \ldots v_n$ is a basis for V and $w_1.w_2, \ldots w_m$ is a basis for W then

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

  with a_{ij} some scalars. The matrix $\mathbf{A} = (a_{ij})$ is called the matrix of f with respect to the ordered bases $v_1, v_2, \ldots v_n$ of V and $w_1, w_2, \ldots w_m$ of W. In coordinate form with $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots \lambda_n)^T$ the coordinates of $v \in V$ with respect to the basis of V and $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots \mu_m)^T$ the coordinates of $f(v)$ with respect to the basis of W then

$$\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\lambda}.$$

  – If we change the basis $v_1, v_2, \ldots v_n$ of V to $v'_1, v'_2, \ldots v'_n$ related by the base change matrix $\mathbf{P} = (p_{ij})$ so that $v'_j = \sum_{i=1}^{n} p_{ij} v_i$ then the matrix of f with respect to the ordered bases $v'_1, v'_2, \ldots v'_n$ of V and $w_1, w_2, \ldots w_m$ of W is $\mathbf{AP}$.

  – If we change the basis $w_1, w_2, \ldots w_m$ of W to $w'_1, w'_2, \ldots w'_m$ related by the base change matrix $\mathbf{Q} = (q_{ij})$ so that $w'_j = \sum_{i=1}^{m} q_{ij} w_i$ then the matrix of f with respect to the ordered bases $v_1, v_2, \ldots v_n$ of V and $w'_1, w'_2, \ldots w'_m$ of W is $\mathbf{Q}^{-1}\mathbf{A}$.

  - Together we have $\mathbf{Q}^{-1}\mathbf{A}\mathbf{P}$ if we change the basis of both the domain and codomain.

- Proposition 5.10: Let V be a vector space with ordered basis B given by $v_1, v_2, \ldots v_n$ and B' given by $v'_1, v'_2, \ldots v'_n$. Let $\mathbf{P} = (p_{ij})$ be the base change matrix so that $v'_j = \sum_{i=1}^{n} p_{ij} v_i$. Suppose $f : V \to V$ is a linear map which has matrix $\mathbf{A}$ with respect to the ordered basis B and matrix $\mathbf{B}$ with respect to the ordered basis B'. Then

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\,.$$

- Definition 5.11: If $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ then $\mathbf{A}$ and $\mathbf{B}$ are called similar matrices. If $\mathbf{B} = \mathbf{P}^{T}\mathbf{A}\mathbf{P}$ then $\mathbf{A}$ and $\mathbf{B}$ are called congruent matrices.

- Proposition 5.12: Let $f, g \in \mathcal{L}[V, W]$ where V, W are finite dimensional vector spaces, and let f, g have matrix representations $\mathbf{A}, \mathbf{B}$ respectively with respect to some ordered bases A, B of V, W respectively. Then

  (a) The matrix representation of λf with respect to A, B is the scalar product $\lambda\mathbf{A}$ of the matrix $\mathbf{A}$.

  (b) The matrix representation of the sum$f + g$ with respect to A, B is the matrix sum $\mathbf{A} + \mathbf{B}$.

- Proposition 5.13: (Composition of functions) Let U, V, W be finite dimensional vector spaces, with ordered bases A, B, C respectively, and let $g : U \to V$ and $f : V \to W$ be linear maps. If f and g are represented by the matrices $\mathbf{A}$ and $\mathbf{B}$ with respect to A, B, C then $f \cdot g$ is represented (with respect to the same bases) by the matrix product $\mathbf{A}\mathbf{B}$.

- In the special case $\mathcal{L}[V, V]$, where V is an n-dimensional vector space over F, then given an ordered basis B of V we have a map from $\mathcal{L}[V, V]$ to the set $M_{n,n}(F)$ of $n \times n$ matrices with entries taken from F, taking f to the matrix representing f (which is unique once B is specified). Additionally:

  - Every matrix is the matrix of some transformation f.
  - The zero transformation $\leftrightarrow$ the zero matrix.
  - The identity transformation $\leftrightarrow$ the identity matrix.
  - Scalar multiplication $\leftrightarrow$ scalar multiplication.
  - Addition of transformations $\leftrightarrow$ addition of matrices.

  I.e. $\mathcal{L}[V, V]$ and $M_{n,n}(F)$ are isomorphic. Moreover composition in $\mathcal{L}[V, V]$ and matrix multiplication in $M_{n,n}(F)$ are isomorphic.

- Polynomials over a field F are expressions like

$$f(x) = \sum_{r=0}^{n} a_r x^r \text{ with } a_r \in F\,.$$

  The degree of the polynomial is the largest r for which $a_r \neq 0$ and is written $\deg(f)$.

- Proposition 5.14: (Division algorithm) If $f(x)$ and $g(x)$ are two polynomials, and $g(x)$ is not the zero polynomial then there exist polynomials $q(x)$ and $r(x)$ such that

$$f(x) = g(x)q(x) + r(x)$$

  and either $r(x)$ is the zero polynomial (i.e. $a_r = 0 \ \forall r$) or else $\deg(r) < \deg(g)$.

- It is generally true for polynomials $p(x), q(x)$ in a single free variable that whenever there is a polynomial identity $p(x) = q(x)$ and an $n \times n$ matrix $\mathbf{A}$ that $p(\mathbf{A}) = q(\mathbf{A})$ holds. For polynomials in more than one variable that is not the case.

- Proposition 5.15: If $p(x), q(x)$ are polynomials and $\mathbf{A}$ is an $n \times n$ matrix then $p(\mathbf{A})q(\mathbf{A}) = q(\mathbf{A})p(\mathbf{A})$.

- Proposition 5.16: If $f \in \mathcal{L}[V, V]$ where V is a finite dimensional vector space, and if $p(x), q(x)$ are polynomials, then $p(f)q(f) = q(f)p(f)$.

- Theorem 5.17: (Remainder theorem) Suppose $p(x)$ is a polynomial of degree at least 1 with coefficients from $\mathbb{R}$ or $\mathbb{C}$ and $\alpha \in \mathbb{C}$. Then α is a root of $p(x)$ iff $(x - \alpha)$ divides $p(x)$ exactly.

- Corollary 5.18: A polynomial $p(x)$ of degree $d \geq 1$ has at most d roots.

- Every polynomial in $\mathbb{C}$ has its maximum number of roots, counting multiplicities. For any $p(x)$ of degree $d \geq 1$ $\exists \alpha_1, \alpha_2, \ldots \alpha_d, c \in \mathbb{C}$ such that

$$p(x) = c \prod_{j=1}^{d} (x - \alpha_j) \, .$$

  The field of complex numbers is algebraically closed. The field of real numbers is not algebraically closed, but $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$. I.e.:

  (1) $\mathbb{C}$ is algebraically closed.

  (2) Every element $\xi \in \mathbb{C}$ satisfies a polynomial equation over $\mathbb{R}$.

6 Eigenvalues

- Definition 6.1: Let $\mathbf{A}$ be an $n \times n$ matrix over a field F. Then a column vector $\mathbf{x} \in F^n$ is called an eigenvector of $\mathbf{A}$, with eigenvalue $\lambda \in F$, if $\mathbf{x} \neq 0$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

- Theorem 6.2: A scalar λ is an eigenvalue of an $n \times n$ matrix $\mathbf{A}$ iff the matrix $\mathbf{A} - \lambda\mathbb{I}_n$ has nullity $n(\mathbf{A} - \lambda\mathbb{I}_n) > 0$.

- Theorem 6.3: Suppose λ is an eigenvalue of an $n \times n$ matrix $\mathbf{A}$. Then the eigenvectors of $\mathbf{A}$ having eigenvalue λ are the non-zero vectors in $\ker(\mathbf{A} - \lambda\mathbb{I}_n) = \{\mathbf{x} : (\mathbf{A} - \lambda\mathbb{I}_n)\mathbf{x} = 0\}$.

- Theorem 6.4: Every $n \times n$ matrix $\mathbf{A}$ over $F = \mathbb{R}$ or $F = \mathbb{C}$ has an eigenvalue λ in $\mathbb{C}$ and an eigenvector $\mathbf{x} \in \mathbb{C}^n$ with eigenvalue λ.

- Definition 6.5: If $f : V \to V$ is a linear map, where V is a vector space over a field F, and $0 \neq v \in V$ with $f(v) = \lambda v$ for some $\lambda \in F$, then v is an eigenvector of f, with eigenvalue λ.

- Proposition 6.6: If $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{P}$ are $n \times n$ matrices related by $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ then $\mathbf{B}$ and $\mathbf{A}$ have the same eigenvalues.

- Lemma 6.7: Suppose that $f : V \to V$ is a linear transformation of an n-dimensional vector space V over a field F. If f has nullity of at least one then there is a basis $v_1, v_2, \ldots v_n$ such that

$$f(v_j) \in \text{span}(v_1, v_2, \ldots v_{n-1}) \quad \forall j = 1, \ldots n\,.$$

- Proposition 6.8: Let $V = \mathbb{C}^n$ be the n-dimensional vector space over $\mathbb{C}$, and suppose f is a linear transformation from V to V. Then there is a basis of V such that, with respect to this basis, the matrix of f is upper triangular.

- Proposition 6.9: If $\mathbf{A}$ is an upper triangular matrix then the diagonal entries in $\mathbf{A}$ are precisely the eigenvalues of $\mathbf{A}$.

- Theorem 6.10: If $\mathbf{A}$ is any upper triangular $n \times n$ matrix with entries from $\mathbb{R}$ or $\mathbb{C}$ and $\lambda_1, \lambda_2, \ldots \lambda_n$ are the diagonal entries of $\mathbf{A}$, including repetitions, then the matrix

$$(\mathbf{A} - \lambda_1\mathbb{I})(\mathbf{A} - \lambda_2\mathbb{I}) \ldots (\mathbf{A} - \lambda_n\mathbb{I})\,,$$

  is the zero matrix.

- Lemma 6.11: If $\mathbf{A}$ is an upper triangular matrix with eigenvalue λ then $\det[\mathbf{A} - \lambda\mathbb{I}] = 0$.

- Proposition 6.12: If $\mathbf{A}$ is an $n \times n$ matrix over $\mathbb{R}$ or $\mathbb{C}$ with eigenvalue $\lambda \in \mathbb{C}$ then $\det[\mathbf{A} - \lambda\mathbb{I}] = 0$.

- Let $\mathbf{Y}(x) = (y_1(x), y_2(x), \dots y_n(x)^T)$ and $\mathbf{Y}'(x) = \mathbf{AY}(x)$ be a system of first order linear differential equations with constant coefficients given by $\mathbf{A} = (a_{ij})$ with $a_{ij} \in \mathbb{C}$. A general solution is given by

$$\mathbf{Y}(x) = \sum_{i=1}^{r} b_i \, \mathrm{e}^{\lambda_i x} \, \mathbf{Y}_i$$

  where $b_i \in \mathbb{C}$ and λ_i are the r eigenvalues of $\mathbf{A}$ with corresponding eigenvectors $\mathbf{Y}_i$. Higher order equations (with constant coefficients) can also be solved in this way by introducing new functions. For example consider $\mathbf{Y}''(x) = \mathbf{A}_1\mathbf{Y}(x) + \mathbf{A}_2\mathbf{Y}'(x)$. Let $\mathbf{Y}_2(x) = \mathbf{Y}'(x)$ and $\mathbf{Y}_1(x) = \mathbf{Y}(x)$ then we have

$$\begin{pmatrix} \mathbf{Y}_1'(x) \\ \mathbf{Y}_2'(x) \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbb{I}_n \\ \mathbf{A}_1 & \mathbf{A}_2 \end{pmatrix} \begin{pmatrix} \mathbf{Y}_1(x) \\ \mathbf{Y}_2(x) \end{pmatrix} ,$$

  which can be solved in the same way as before. This generalizes in the obvious way to higher order differential equations.

- Theorem 6.13: If $\lambda_1, \lambda_2, \dots \lambda_r$ are distinct eigenvalues of an $n \times n$ matrix $\mathbf{A}$, with $r \leq n$, with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_r$ then $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_r$ are linearly independent.

- Definition 6.14: An $n \times n$ matrix $\mathbf{A}$ with entries from $\mathbb{R}$ or $\mathbb{C}$ is said to be diagonalizable if there exists an invertible matrix $\mathbf{P}$ such that $\mathbf{B} = \mathbf{P}^{-1}\mathbf{AP}$ with $\mathbf{B}$ being diagonal. $\mathbf{P}$ is said to diagonalize $\mathbf{A}$.

- Theorem 6.15: An $n \times n$ matrix $\mathbf{A}$ with entries from $\mathbb{R}$ or $\mathbb{C}$ is diagonalizable iff $\mathbf{A}$ has n linearly independent eigenvectors.

- In relation to theorem 6.15 we note the following:

  (a) The diagonal entries of $\mathbf{B}$ are the eigenvalues of $\mathbf{B}$ and $\mathbf{A}$.

  (b) The column vectors of a diagonalizing matrix $\mathbf{P}$ are the eigenvectors of $\mathbf{A}$.

  (c) $\mathbf{P}$, and therefore $\mathbf{B}$ is not unique given $\mathbf{A}$. Any reordering of the eigenvectors in $\mathbf{P}$ would work.

  (d) If $\mathbf{A}$ is diagonalizable then $\mathbf{A} = \mathbf{PBP}^{-1}$ and $\mathbf{A}^k = \mathbf{PB}^k\mathbf{P}^{-1}$ with $\mathbf{B}$ diagonal, giving a simple expression for the powers of the matrix $\mathbf{A}$.

- Definition 6.16: Let V be a finite dimensional inner product space, and suppose that f is a linear transformation $f : V \to V$. f is called orthogonal (if V is an inner product space over $\mathbb{R}$) or unitary (if V is an inner product space over $\mathbb{C}$) if

$$\langle u|v\rangle = \langle f(u)|f(v)\rangle \, , \quad \forall u, v \in V \, .$$

- Proposition 6.17: Let V be a finite dimensional inner product space, and suppose $f : V \to V$ is a linear transformation. Let $e_1, e_2, \dots e_n$ be an orthonormal basis of V. Then f preserves the inner product $\langle u|v\rangle$ on V, i.e. is orthogonal or unitary, iff $f(e_1), f(e_2), \dots f(e_n)$ is orthonormal.

- Proposition 6.18: Let V be a finite dimensional inner product space, and let $e_1, e_2, \dots e_n$ be an orthonormal basis of V. Suppose $f : V \to V$ is a linear transformation with matrix $\mathbf{P}$ with respect to the ordered basis $e_1, e_2, \dots e_n$. Then f preserves the inner product (i.e. is orthogonal or unitary as appropriate) iff $\mathbf{P}^{-1}$ exists and $\mathbf{P}^{-1} = \bar{\mathbf{P}}^T$.

- For a conjugate symmetric sesquilinear form F define on an inner product space V we can define a corresponding linear map $f : V \to V$ by

$$f(v) = \sum_{i=1}^{n} F(e_i, v)e_i \,, \quad \forall v \in V \,,$$

  which is independent of the orthonormal basis $e_1, e_2, \dots e_n$ used. Similarly given a linear map $f : V \to V$ we can define a corresponding conjugate symmetric sesquilinear form F via

$$F(v, w) = \langle v | f(w) \rangle \,, \quad \forall v, w \in V \,.$$

  These are the inverse of each other.

- Proposition 6.19: Suppose that $f : V \to V$ is a linear transformation on an inner product space V,and suppose that F is the corresponding conjugate symmetric sesquilinear form, and that $\{e_1, e_2, \dots e_n\}$ is an orthonormal basis of V. Then f and F are represented by the same matrix with respect to the ordered basis $e_1, e_2, \dots e_n$.

- Definition 6.20: A linear tranformation $f : V \to V$ of an inner product space V (over $\mathbb{R}$ or $\mathbb{C}$) is said to be self-adjoint (or Hermitian) if

$$\langle f(v) | w \rangle = \langle v | f(w) \rangle = \,, \quad \forall v, w \in V \,.$$

- Proposition 6.21: If f is a self-adjoint transformation of an inner product space V, and if $\{e_1, e_2, \dots e_n\}$ is an orthonormal basis of V, then the matrix $\mathbf{A}$ of f with respect to the ordered basis $e_1, e_2, \dots e_n$ is conjugate-symmetric, i.e. $\bar{\mathbf{A}}^T = \mathbf{A}$.

- Proposition 6.22: If f is a linear transformation of an inner product space V, and if $\{e_1, e_2, \dots e_n\}$ is an orthonormal basis of V with respect to which the matrix $\mathbf{A}$ of f is conjugate-symmetric, i.e. $\bar{\mathbf{A}}^T = \mathbf{A}$, then f is self-adjoint.

- Theorem 6.23: If f is a self-adjoint transformation of an inner product space V and λ is an eigenvalue of f then λ is real.

- Theorem 6.24: Any self-adjoint $f : V \to V$ of a finite dimensional inner product space V is diagonalizable.

- Theorem 6.25: Let f be a self-adjoint linear transformation $f : V \to V$, and suppose v_1, v_2 are eigenvectors of f with corresponding eigenvalues λ_1, λ_2. If $\lambda_1 \neq \lambda_2$ then v_1 and v_2 are orthogonal.

7 Quadratic Forms

- Definition 7.1: Given a symmetric bilinear form F on a real vector space V, we define a map $Q : V \to \mathbb{R}$ by $Q(v) = F(v, v)$; Q is called the quadratic form associated with the symmetric bilinear form F.

- Lemma 7.2: Given a symmetric bilinear form F on a real vector space V, and the quadratic form Q associated with F, then

$$F(v, w) = \frac{1}{2} \left(Q(v + w) - Q(v) - Q(w) \right) , \quad \forall v, w \in V .$$

- Definition 7.3: Given a conjugate-symmetric sesquilinear form F on a complex vector space V, we define a map $H : V \to \mathbb{R}$ by $H(v) = F(v, v)$; H is called the Hermitian form associated with the conjugate-symmetric sesquilinear form F.

- Lemma 7.4: Given a conjugate-symmetric sesquilinear form F on a complex vector space V, and the Hermitian form H associated with F, then $\forall v, w \in V$:

$$\begin{aligned} F(v, w) &= \tfrac{1}{2} \left(H(v + w) + \mathrm{i}H(v - \mathrm{i}w) - (1 + \mathrm{i})(H(v) + H(w)) \right) , \\ F(v, w) &= \tfrac{1}{4} \left(H(v + w) - \mathrm{i}H(v - w) + \mathrm{i}H(v - \mathrm{i}w) - \mathrm{i}H(v + \mathrm{i}w) \right) . \end{aligned}$$

- Proposition 7.5: If Q is a quadratic form on a real vector space V, then

$$Q(\lambda x) = \lambda^2 Q(x) , \quad \forall \lambda \in \mathbb{R} , \text{ and } \forall x \in V .$$

  Similarly if H is a Hermitian form on a complex vector space V, then

$$H(\lambda x) = |\lambda|^2 H(x) , \quad \forall \lambda \in \mathbb{C} , \text{ and } \forall x \in V .$$

- Proposition 7.6: Let $V = \mathbb{R}^n$. Then every quadratic form on V is given by a homogeneous function of the coordinates of degree 2. Conversely every homogeneous function of degree 2 of the coordinates is a quadratic form.

- Theorem 7.7: (Sylvester's law of inertia part I.) Let V be an n dimensional real vector space, and let F be a symmetric bilinear form on V. Then there are non-negative integers k and m, and a basis $\{w_1, w_2, \ldots w_n$ of V such that:

$$\begin{aligned} F(w_i, w_j) &= 0 \quad \forall i \neq j , \\ F(w_i, w_i) &= 1 \quad \text{for } i \leq k , \\ F(w_i, w_i) &= -1 \text{ for } k < i \leq k + m , \\ F(w_i, w_i) &= 0 \quad \text{for } k + m < i . \end{aligned}$$

- Lemma 7.8: Let F be a symmetric bilinear form on a real vector space V, and suppose that $F(v, v) = 0 \; \forall v \in V$. Then $F(v, w) = 0 \; \forall v, w \in V$.

- Lemma 7.9: Let F be a bilinear form on a real vector space V and suppose that $w_1, w_2, \ldots w_n$ are vectors from V which are orthogonal with respect to F. For all scalares $\lambda_i \in \mathbb{R}$ if

$$\lambda_1 w_1 + \lambda_2 w_2 + \ldots + \lambda_n w_n = 0$$

  then $\lambda_j = 0 \; \forall j$ such that $F(w_j, w_j) = 0$.

- Lemma 7.10: Let F be a bilinear form on a real vector space V and suppose that $w_1, w_2, \ldots w_k$ are vectors from V which are orthogonal with respect to F, and that $F(w_i, w_i) \neq 0 \; \forall i \leq k$. Then $\forall v \in V \; \exists u \in V$ such that $F(w_i, u) = 0 \; \forall i \leq k$, and v is a linear combination of $w_1, w_2, \ldots w_k, u$. I.e.
$$V = \text{span}(U \cup \{w_1, w_2, \ldots w_k\})$$
where
$$U = \{u \in V : F(w_i, u) = 0, \forall i \leq k\}.$$

- Theorem 7.11: (Sylvester's law of inertia part II.) Let V be an n dimensional real vector space, and let F be a symmetric bilinear form on V, with two diagonal matrix representations $\mathbf{A}$ and $\mathbf{A}'$, as in theorem 7.7, with respect to the bases $e_1, e_2, \ldots e_n$ and $e'_1, e'_2, \ldots e'_n$ of V. If $\mathbf{A}$ has k positive diagonal entries and m negative diagonal entries, and $\mathbf{A}'$ has k' positive diagonal entries and m' negative diagonal entries, then $k = k'$ and $m = m'$.

- Definition 7.12: Let V be an n dimensional real vector space, and let F be a symmetric bilinear form on V, with a diagonal matrix representation $\mathbf{A}$, with k and m as in theorem 7.7.

  $k + m$ is the rank of F.

  $k - m$ is the signature of F.

- Proposition 7.13: With the same notation as definition 7.12, F is positive definite (i.e. is an inner product) iff $k = n$ and $m = 0$.

- Lemma 7.14: Suppose F is a symmetric bilinear form on a real vector space V, and let $v_1, v_2, \ldots v_n$ be a basis of V. If $F(v_i, v_i) > 0 \; \forall i$ and $F(v_i, v_j) > 0 \; \forall i \neq j$ then F is positive definite.