Basic Exam Review Sheet

Stepan Malkov[1]

This sheet reviews the key concepts and material assessed on the UCLA Basic Exam. The central focus is on material covered in upper-level undergraduate and first-year graduate courses in linear algebra and real analysis. More importantly, the breadth and scope of content is sufficient to prepare a mathematics student for advanced graduate-level coursework in algebra and analysis. It also includes plenty of examples and problem-solving strategies in order to make it as easy as possible to study for an exam that broadly covers this material.

The following is conventional notation used throughout this sheet:

- $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ - integers, rational numbers, real numbers, and complex numbers, respectively.
- $\mathbb{K}$ - an arbitrary scalar field.
- "iff" - if and only if.

Linear Algebra

Fundamentals

At its core, linear algebra is the study of *vector spaces*, typically taken to be the finite-dimensional Euclidean spaces $\mathbb{R}^d$, and the action of linear transformations on elements of those spaces. We begin with some fundamental concepts in the study of linear algebra.

Definition 1. *A* ***vector space*** $(V, \mathbb{K})$ *is a set* V *of vectors (denoted in bold) along with a field* $\mathbb{K}$ *of scalars that satisfy the following axioms:*

- *(Identity)* $\exists\, \mathbf{0} \in V$,

$$\mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v} = \mathbf{v}, \quad \forall\, \mathbf{v} \in V.$$

- *(Associativity)*

$$\mathbf{v_1} + (\mathbf{v_2} + \mathbf{v_3}) = \mathbf{v_1} + \mathbf{v_2} + \mathbf{v_3} = (\mathbf{v_1} + \mathbf{v_2}) + \mathbf{v_3}, \quad \forall \mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3} \in V.$$

- *(Commutativity)*

$$\mathbf{v_1} + \mathbf{v_2} = \mathbf{v_2} + \mathbf{v_1}, \quad \forall\, \mathbf{v_1}, \mathbf{v_2} \in V.$$

- *(Existence of Inverses)* $\forall\, \mathbf{v} \in V, \exists\, \mathbf{w} = -\mathbf{v} \in V$,

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v} = \mathbf{0}.$$

- *(Scalar Associativity)*

$$k_1(k_2\mathbf{v}) = k_1 k_2 \mathbf{v} = k_1(k_2\mathbf{v}), \quad \forall\, k_1, k_2 \in \mathbb{K}, \mathbf{v} \in V.$$

- *(Distributivity)*

$$(k_1 + k_2)\mathbf{v} = k_1\mathbf{v} + k_2\mathbf{v}, \quad \forall\, k_1, k_2 \in \mathbb{K}, \mathbf{v} \in V.$$

and

$$k(\mathbf{v_1} + \mathbf{v_2}) = k\mathbf{v_1} + k\mathbf{v_2}, \quad \forall\, k \in \mathbb{K}, \mathbf{v_1}, \mathbf{v_2} \in V.$$

Remark 1. *Note that* 0 *and* $\mathbf{0}$ *are distinct objects - the former is the zero scalar, while the latter is the zero vector.*

Vectors are usually denoted in bold or with an arrow, i.e. $\mathbf{v}$ or $\vec{v}$. In the following, we adopt standard practices and omit the notation for convenience, referring simply to a vector v. For clarity, we use $a, b, c, ...$ to refer to scalars and v, w to refer to vectors, using the bold notation where necessary to distinguish between the two.

Let $(V, \mathbb{C})$ be a vector space. We now define the notions of linear independence and basis, which are both crucial to the study of linear transformations.

Definition 2. *A subset* $W \subset V$ *is called* ***linearly independent*** *if for any finite collection of vectors* $\{v_i\}_{i=1}^k \subset W$,

$$\sum_{i=1}^k a_i v_i = 0 \Rightarrow a_i = 0 \quad \forall i.$$

Particularly, $\{v_1, ..., v_k\}$ *are linearly independent if*

$$\sum_{i=1}^k a_i v_i = 0 \Rightarrow a_1 = a_2 = ... = a_k = 0.$$

Definition 3. *The* ***span*** *of* $W \subset V$ *is the set of finite linear combinations of vectors in* W, *i.e.*

$$\textit{span } W = \left\{ \sum_{i=1}^k a_i w_i \,\middle|\, w_i \in W \right\}.$$

Definition 4. *A* ***basis*** *of a vector space* V *is a set of linearly independent vectors which span the space. Particularly,* $\{v_1, ..., v_k\}$ *is a basis for* V *if*

$$\sum_{i=1}^k a_i v_i = 0 \Rightarrow a_i = 0 \quad \forall i,$$

[1] *Department of Mathematics, University of California, Los Angeles*

and

$$v = \sum_{i=1}^{k} b_i v_i \quad \forall\, v \in Vm$$

for some $b_1, ..., b_k \in \mathbb{C}$.

The following is a key theorem (which relies on the Axiom of Choice in the form of Zorn's lemma) demonstrating that every vector space has a basis.

Theorem 1. *(Steinitz Exchange Lemma) Let* $\{v_1, ..., v_m\}$ *be linearly independent and suppose* $V = span\{w_1, ..., w_n\}$*. Then* $m \leq n$ *and* $\{v_1, ..., v_m, w_{i_1}, ..., w_{i_k}\}$ *is a basis for* V*.*

Theorem 2. *Every vector space* V *has a basis.*

Proof. Consider a partial order on linearly independent subsets of V, with the order given by inclusion. By Zorn's lemma, there must therefore exist a maximal subset W. But W must span V - otherwise there would exist $v \notin \text{span}\, W$, and $W \cup \{v\}$ would contain W, contradicting its maximality. □

Similarly, one can show the following.

Theorem 3. *Every basis of* V *has the same cardinality.*

Proof. Let $\{v_i\}_{i\in I}, \{w_j\}_{j\in J}$ be two bases for V. This implies that v_i is a *finite* linear combination of some $\{w_{j_k}\}$ for $j_k \in J_i$ for each i. Thus, $V = \text{span}\,\{v_i\} = \text{span}\,\bigcup_j \{w_{j_k}\}$. Since no proper subset of $\{w_j\}$ can span V, $\bigcup_i J_i = J$. Then, since each J_i has at least one element j_i, there exists an injective map $i \to j_i$, i.e. $|I| \leq |J|$. By symmetry, $|J| \leq |I|$, so $|I| = |J|$. □

Critically, this allows us to define the notion of dimension for a vector space.

Definition 5. *The* ***dimension*** *of a vector space* $dim\, V = |V|$ *is the cardinality of a basis* $\mathcal{B}$ *of* V*. If* $|V| < \infty$*, the vector space is called* ***finite-dimensional****. Otherwise, it* V *is known as an* ***infinite-dimensional*** *vector space.*

Remark 2. *It can be easily shown that every finite-dimensional vector space of dimension* d *is isomorphic to* $\mathbb{K}^d$*.*

Example 1.

- *As mentioned, every finite-dimensional vector space is isomorphic to* $\mathbb{K}^d$*.*
- *Continuous functions on a metric space* (X, d) *are an infinite-dimensional vector-space.*
- *Matrices* $M_{m\times n}(\mathbb{C})$ *form a vector space of dimension* mn*.*
- *If* $W \subset V$ *is a* ***subspace*** *(i.e. a subset of* V *that is a vector space), then* V/W *is a vector space of dimension* $|V| - |W|$*.*

Linear Transformations

One arrives at the crux of linear algebra when considering the action of a particular set of functions called *linear transformations* acting on the vector space V.

Definition 6. *A* ***linear transformation*** $L : V \to W$ *between vector spaces is a function satisfying the following properties:*

- $L(av) = aL(v) \quad \forall a \in \mathbb{K}, v \in V,$
- $L(v_1 + v_2) = L(v_1) + L(v_2) \quad \forall v_1, v_2 \in V.$

The ***kernel*** *or* ***null space*** *of* L *is defined as*

$$ker\, L = \{v \in V | Lv = 0\}$$

and the ***image*** *or* ***range*** *of* L *is defined to be*

$$L(V) = im\ L = \{w \in W | \exists\, v \in V, Lv = w\}.$$

$dim\ ker\ L$ *is called the* ***nullity*** *and* $dim\ im\ L$ *is called the* ***rank*** *of* L*. Moreover, let* $Col(A)$ *and* $Nul(A)$ *be the span of the columns and rows of* A*, respectively. If* A *is an* $n \times m$ *matrix and* $rank(A) = \min(n, m)$*,* A *is said to be* ***full rank***

Given a matrix A, it is true that $\text{Col}(A) = \text{im}\ A$.

Lemma 4. *For a linear transformation* $L : V \to W$ *with* $|V| = |W|$*, the following are equivalent:*

1. *L is an isomorphism.*
2. *L is injective, i.e.* $ker\ L = 0$*.*
3. *L is surjective, i.e.* $L(V) = W$*.*
4. *L sends a basis of V to a basis of W.*

Then, L *is called* ***non-singular****.*

A linear transformation $L : V \to V$ represented by A is invertible if A^{-1} exists, and $AA^{-1} = A^{-1}A = I$.

Lemma 5. *The following are equivalent:*

1. $Av = w$ *has a unique solution given by* $v = A^{-1}w$*.*
2. $det\ A \neq 0$*.*
3. *A is an isomorphism of vector spaces.*
4. *A has full rank.*

Reduced Row Echelon Form

Definition 7. *A process known as* ***Gaussian elimination*** *allows one to use three* ***row/column operations*** *on a matrix (switching two rows/columns, adding a multiple of one to another, or multiplying by a nonzero constant) to bring a matrix to* ***reduced row echelon form****, which takes the form*

$$\begin{bmatrix} 1 & 0 & a_1 & ... & ... \\ 0 & 1 & a_2 & ... & ... \\ 0 & 0 & 0 & 1 & ... \end{bmatrix}$$

The matrix is in this form if it has a leading 1 in each row and if every other value in the column with a leading 1 is zero. This form allows one to write a parametrization of the solutions to $Ax = b$, or equivalently, construct a basis for the kernel of A. Moreover, the columns with leading 1s are the columns which span the image of A.

Matrix of a Linear Transformation

For $L \in \text{Hom}(V, W)$, and $\{v_i\}_{i=1}^n = \mathcal{B}, \{w_j\}_{j=1}^m = \mathcal{E}$ being bases of V, W, respectively, we note that

$$Lv_i = \sum_{j=1}^{m} a_{ij} w_j,$$

meaning that the matrix $\mathcal{M}_\mathcal{B}^\mathcal{E} = (a_{ij})$. In other words, the i-th **column** of the matrix is the coefficients of the basis vectors $\{w_j\}$ of Lv_i.

Definition 8. *The **row (column) rank** of a matrix/linear transformation is the maximum number of linearly independent rows (columns) of a matrix. Equivalently, the rank of a linear transformation ϕ is the dimension of the image of ϕ.*

*Two matrices/linear transformations A, B are **similar** if there exists an invertible matrix P such that $B = PAP^{-1}$. Similarity can be easily shown to be an equivalence relation.*

Remark 3. *Note that*

$$P = \mathcal{M}_\mathcal{E}^\mathcal{B}(I)$$

satisfies $P^{-1}\mathcal{M}_\mathcal{B}^\mathcal{B}(\phi)P = \mathcal{M}_\mathcal{E}^\mathcal{E}(\phi)$, implying that similar matrices are precisely those that represent a particular linear transformation with respect to two different bases.

Theorem 6. *(Rank-Nullity Theorem) Let $T : V \to W$ be a linear transformation. Then dim im T + dim ker T = dim V.*

Proof. Let $\{v_1, ..., v_k\}$ be a basis for ker T. Then, by the Steinitz exchange lemma, there exist vectors $v_{k+1}, ..., v_n$ such that $\{v_1, ..., v_n\}$ form a basis for V. It is left to show that $v_{k+1}, ..., v_n$ form a basis for im T. Clearly, $\{T(v_1), ..., T(v_n)\} = \{T(v_{k+1}), ..., T(v_n)\}$ span im T. Moreover, they are linearly independent, as $\sum_{i=k+1}^{n} a_i v_i = v \in \ker T$ implies $a_i = 0$ (since $v_{k+1}, ..., v_n \notin \ker T$) and are linearly independent). Thus, $\{T(v_1), ..., T(v_n)\}$ form a basis for im T, proving the Rank-Nullity Theorem.
Alternative Proof: Let $r = \dim \text{im T}, n = \dim V$. If the rank is maximal, we are done. If not, there are $n -$ rank A free variables $t_1, ...t_{n-r}$ in the solution to $Ax = 0$, and let $x_1, ..., x_{n-r}$ be the solutions obtained by taking one variable to be 1 and the rest to be zero. Note that this implies that $\{x_1, ..., x_{n-r}\}$ are linearly independent, and moreover, they span ker A. Thus, they form a basis for ker$A$, and dim ker $T = n - r$. □

Remark 4. *By the first isomorphism theorem and the splitting lemma, we can actually obtain a stronger statement that V = ker $T \oplus$ im T^*.*

Dual Space

Definition 9. *Define $V^* = Hom(V, \mathbb{K})$ to be the dual space of V. If V has basis $\{v_1, ..., v_n\}$, V^* has the **dual basis** $\{v_1^*, ..., v_n^*\}$, where $v_i^*(v_j) = \delta_{ij}$. One can similarly define the **double dual** $V^{**} = Hom(V^*, \mathbb{K})$.*

For $\phi \in \text{Hom}(V, W), f \in W^*$, $f \circ \phi \in V^*$, so the mapping $f \to f \circ \phi$ induces a map $\phi* \in \text{Hom}(W^*, V^*)$.

Theorem 7. $\mathcal{M}_{\mathcal{E}^*}^{\mathcal{B}^*}(\phi^*) = (\mathcal{M}_\mathcal{B}^\mathcal{E})^T$.

Theorem 8. *The row rank and column rank of a matrix are equal (from now on referred to as the **rank** of the matrix).*

Proof. Pick bases of V, W such that A represents the linear transformation ϕ with respect to these bases. It suffices to show that ϕ and ϕ^* have the same column rank. Then,

$$f \in \ker \phi^* \Leftrightarrow \phi(V) \subseteq \ker f \Leftrightarrow f \in \text{Ann}(\phi(V)).$$

It can then be shown that dim Ann $(\phi(V)) = \dim W - \dim \phi(V)$ and $\dim \ker \phi^* = \dim W^* - \dim \phi^*(W^*)$, demonstrating that $\dim \phi(V) = \dim \phi^*(W^*)$, as needed. □

Tensor Products

Definition 10. *Given vector spaces V, W, a **tensor product space** $V \otimes W$ is a vector space with an associated bilinear function $B : V \times W \to V \otimes W$ mapping $(v, w) \to v \otimes w$. $v \otimes w$ is called the **tensor product** of v and w, and is also called an **simple tensor.** Tensors are defined by three key properties:*

1. $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$.
2. $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$.
3. $a(v \otimes w) = (av) \otimes w = v \otimes (aw)$ *for* $a \in \mathbb{K}$.

*Elements of a tensor product space are called **tensors,** and every tensor can be written as a sum of simple tensors.*

Given a linear map $f \in \text{Hom}(U, V)$, the tensor product $f \otimes W$ is the unique linear map that satisfies $(f \otimes W)(u \otimes w) = f(u) \otimes w$. Importantly, if f, g are two linear transformation represented by matrices A, B, respectively, then their tensor product is a linear transformation represented by the **Kronecker product** of A and B, which is a $|A||B| \times |A||B|$ matrix C given by the block product of A and B, i.e.

$$C = \begin{bmatrix} a_{11}B & a_{12}B & ... \\ a_{21}B & a_{22}B & ... \\ ... & ... & ... \end{bmatrix}.$$

Tensor products can be thought of as "products" of vector spaces in analogy to direct sums, and satisfy the following properties:

1. If $\{v_i\}, \{w_j\}$ are bases for V and W, respectively, $\{v_i \otimes w_j\}$ is a basis for $V \otimes W$. Particularly, if dim $V = n$ and dim $W = m$, dim $V \otimes W = nm$.
2. $U \otimes (V \otimes W) \cong (U \otimes V) \otimes W$.
3. $V \otimes W \cong W \otimes V$.
4. $U \otimes (V \oplus W) \cong (U \otimes V) \oplus (U \otimes W)$.
5. $\text{Hom}(U \otimes V, W) \cong \text{Hom}(U, \text{Hom}(V, W))$.

Proof. We prove (1), i.e. that if dim $V = n$ and dim $W = m$, dim $V \otimes W = nm$. Note that there is a **canonical isomorphism**

$$V^* \otimes W \cong \text{Hom } (V, W),$$

given by the map $f : V^* \otimes W \to \text{Hom}(V, W)$ as

$$f(\phi \otimes v)(u) = \phi(u)v$$

and its inverse map $g : \text{Hom}(V, W) \to V^* \otimes W$ given by

$$g(u) = \sum_i e_i^* \otimes u(e_i).$$

Indeed, the two maps are inverses of each other, as

$$f(g(u))(v) = \sum_i f(e_i^* \otimes u(e_i))(v) = \sum_i e_i^*(v)u(e_i)$$
$$= u\left(\sum_i e_i^*(v)e_i\right) = u(v).$$

and

$$g(f(\phi \otimes v)) = \sum_i e_i^* \otimes f(\phi \otimes v)(e_i) = \sum_i e_i^* \otimes \phi(e_i)v =$$
$$\sum_i \phi(e_i)e_i^* \otimes v = \phi \otimes v.$$

Since dim Hom$(V, W) = nm$, the result follows, proving (1). ☐

Eigenvectors and Eigenvalues

Throughout this section, let $L : V \to V$ be a linear transformation (i.e. L is an **endomorphism**) on a finite-dimensional vector space V represented by a matrix A with respect to some basis.

Definition 11. *A nonzero **eigenvector** v of L (A, respectively) with **eigenvalue** λ is a vector that satisfies*

$$Av = \lambda v.$$

*The subspace of eigenvectors of V with eigenvalue λ forms an **eigenspace** $ker(A - \lambda I) = E_\lambda \subset V$. We also define the **minimal polynomial** $m_A(x)$ to be the unique monic polynomial of smallest degree such that $m_A(A) = 0$. Moreover, we define the **characteristic polynomial** $c_A(x) = det(A - xI)$.*

By Lemma 4, we conclude that λ is an eigenvalue of A iff det $(A - \lambda I) = 0$.

Definition 12. *The **algebraic multiplicity** of an eigenvalue λ is its multiplicity as a root of $c_A(x)$. The **geometric multiplicity** is the dimension of the eigenspace dim E_λ = dim $ker(A - \lambda I)$.*

Lemma 9. *The geometric multiplicity is strictly less than or equal to the geometric multiplicity.*

Proof. Let $v_1, ..., v_k$ be eigenvectors for a given eigenvalue λ and change basis to those eigenvectors. Then, the linear transformation takes the form

$$\begin{bmatrix} \lambda I_k & B \\ 0 & C \end{bmatrix}.$$

Then, $\chi_A(x) = \chi(I_k)\chi(C) = (x - \lambda)^k \chi(C)$. Thus, the algebraic multiplicity of λ is at least k. ☐

Moreover, the following relations hold between eigenvalues, eigenvectors, and minimal/characteristic polynomials.

Theorem 10. *The following are equivalent:*

1. $m_A(\lambda) = 0$.
2. $c_A(\lambda) = 0$.
3. *λ is an eigenvalue of A.*
4. *$\lambda I - A$ is singular.*
5. *det $(\lambda I - A) = 0$.*

Proof. The equivalence of (1) and (2) follows from the remark below, since $m_A(x)$ generates I_L. (2) and (3) are equivalent by the argument above. In fact, the minimality of $m_A(x)$ guarantees that $m_A(x)|c_A(x)$. (3) and (4) are equivalent since v is a nonzero element of $\ker(\lambda I - A)$, and (4) and (5) are equivalent by our results on determinants. ☐

Remark 5. *More formally, let $I_L = \{p \in \mathbb{K}[t] | p(L) = 0\}$. Then, I_L is a proper ideal in $\mathbb{K}[t]$, so since $\mathbb{K}[t]$ is a PID, the minimal polynomial is defined as the monic generator of the ideal $I_L = Ann(V)$.*

Remark 6. *Note that the field $\mathbb{K}$ plays a crucial role here. If $\mathbb{K}$ is not algebraically closed, then the minimal polynomial might not split into linear factors.*

Theorem 11. *Similar matrices have the same minimal and characteristic polynomials, by the converse is not true.*

Proof.

$$\det(PAP^{-1} - \lambda I) = \det(P)\det(A - \lambda I)\det(P)^{-1} = \det(A - \lambda I).$$

☐

Theorem 12. *Let A and B be matrices which commute, i.e. $AB = BA$. Then,*

1. *A and B share at least one common eigenvector.*
2. *Suppose that A has all distinct eigenvalues. Then, A and B share a common basis of eigenvectors, i.e. there exists a matrix P such that both PAP^{-1} and PBP^{-1} are diagonal (such matrices are called **simultaneously diagonalizable**).*

Proof. To prove (1), pick an eigenvector v of A, and note that the subspace $V = \text{span}\ (v, Bv, B^2v, ...)$ (which consists of eigenvectors of A from the commutativity condition) is invariant under B. Thus, V contains an eigenvector of B, which is also an eigenvector of A, showing that the two matrices share an eigenvector. We now show (2). Consider a decomposition of V into eigenspaces of A:

$$V = E_{\lambda_1} \oplus ... \oplus E_{\lambda_n}.$$

Since the matrices commute, each eigenspace is stable under the action of B, and since each eigenspace is one-dimensional, this implies that E_{λ_i} is also an eigenspace of B. It thus follows that there exists a shared basis of eigenvectors. ☐

Remark 7. *This theorem requires care with the assumptions. It is not true without the assumption of distinct eigenvalues - for instance if $A = I$, A is diagonalizable but the statement of the theorem is not necessarily true. Similarly, A and B may not share any eigenvalues (take $A = I, B = 2I$, for instance) in common unless the distinct eigenvalue condition on A is specified. Finally, the commutativity condition does not imply that both matrices share a basis of generalized eigenvectors (i.e. the matrices are not necessarily simultaneously Jordanized).*

Listed below are additional important properties of eigenvalues and eigenvectors:

1. tr $(A) = \sum_{i=1}^n a_{ii} = \sum_i^n \lambda_i$.
2. det $(A) = \prod_{i=1}^n \lambda_i$
3. The eigenvalues of A^k are λ_i^k for $k \in \mathbb{Z}$.

4. A is invertible iff every eigenvalue is nonzero.

Remark 8. *Note that since $c_A(x)$ consists of eigenvalues of A as its roots (with the appropriate multiplicity), $c_A(x) = x^n - tr\,(A)x^{n-1} + ... + (-1)^n det\,(A)$.*

Definition 13. *If $v \in ker(A - \lambda I)^m$ but $v \notin ker(A - \lambda I)^{m-1}$ for some $m > 1$, v is known as a* ***generalized eigenvector*** *of rank m. The set spanned by all generalized vectors for a given eigenvalue λ is called the* ***generalized eigenspace*** *F_λ.*

Range-Null Space Decomposition

Given a linear transformation T on a finite-dimensional vector space V, we note that the following chain conditions hold:

$$\ker T \subseteq \ker T^2 \subseteq ... \subseteq \ker T^k \subseteq ...$$

and

$$\text{im } T \supseteq \text{im } T^2 \supseteq ... \supseteq \text{im } T^k \supseteq ...,$$

and since V is finite-dimensional, this implies that there exists an integer k past which both chains stabilize, which is called the **index** of T. Then, one can make the following important claim:

Theorem 13. *(Range-Nullspace Decomposition)*

$$V = ker\, T^k \oplus im\, T^k.$$

Proof. It can be easily shown that the intersection

$$\ker T^k \cap \text{im } T^k = \{0\}.$$

Then, if B is a basis for $\ker T^k$ and C is a basis for $\text{im } T^k$, by the rank nullity theorem, $B \cup C$ is a basis for V since $|B| + |C| = \dim V$. Thus, $V = \ker T^k \oplus \text{im } T^k$. □

Remark 9. *In fact, one can show that if m is the minimal integer such that $ker\, T^m = ker\, T^{m+1}$, then the chain stabilizes at m, and an analogous condition holds for $im\, T$.*

Remark 10. *As a consequence, while there may not always exist a basis for V consisting of eigenvectors of T, there always exists a basis for V consisting of generalized eigenvectors for T.*

Diagonalizability

Definition 14. *A matrix A describing a linear transformation T is* ***diagonalizable*** *if it is similar to a diagonal matrix, i.e. $PDP^{-1} = A$, where D is a diagonal matrix. The columns of P form a basis of V consisting of eigenvectors of T, and the diagonal entries of D are the eigenvalues of A.*

The following properties characterize diagonalizable matrices:

- $V = \bigoplus_\lambda E_\lambda$.
- The geometric and algebraic multiplicity of each eigenvalue coincide.
- $m_A(x)$ is a product of distinct linear factors over $\mathbb{K}$.
- A set of matrices $\{A_i\}$ is simultaneously diagonalizable iff $A_i A_j = A_j A_i$ for all i, j.

Remark 11. *Typically, one considers whether a matrix is diagonalizable over $\mathbb{C}$. A real matrix A is diagonalizable over $\mathbb{R}$ iff both its eigenvalues are real, as otherwise the eigenvectors will be complex.*

Bilinear Forms and Inner Products

Of particular interest in linear algebra is the study of bilinear forms $B : V \to V \to \mathbb{K}$, which are functions linear in each argument. It can be shown that with respect to a certain basis, they can be represented as $B(v, w) = \langle v, w \rangle = v^T A w$ for some matrix A. We differentiate between bilinear forms over $\mathbb{R}$ and $\mathbb{C}$. Over $\mathbb{R}$, **symmetric forms** satisfy the following properties:

1. $\langle av, w \rangle = \langle v, aw \rangle = a\langle v, w \rangle$.
2. $\langle v, w \rangle = \langle w, v \rangle$.
3. $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$.

Over $\mathbb{C}$, **Hermitian forms** satisfy the following properties:

1. $\langle av, w \rangle = a\langle v, w \rangle$.
2. $\langle v, aw \rangle = \overline{a}\langle v, w \rangle$.
3. $\langle v, w \rangle = \overline{\langle w, v \rangle}$.
4. $\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle$.

Definition 15. *If $\langle v, v \rangle > (\geq)\, 0$ for $v \neq 0$, $\langle \cdot, \cdot \rangle$ is called* ***positive (semi-)definite.*** *If $\langle v, v \rangle < (\leq)\, 0$ for $v \neq 0$, $\langle \cdot, \cdot \rangle$ is called* ***negative (semi-)definite.*** *With respect to a basis $\{v_1, ..., v_n\}$ of V, the matrix A representing a particular bilinear form satisfies $a_{ij} = \langle v_i, v_j \rangle$.*

Definition 16. *A positive definite symmetric (Hermitian) bilinear form is called an* ***inner product,*** *in which case V is called an* ***inner product space.***

Remark 12. *If A, A' are two matrices representing a Hermitian form with respect to bases $\mathcal{B}, \mathcal{E}$, respectively, then if $P = \mathcal{M}^{\mathcal{E}}_{\mathcal{B}}$, $A' = P^*AP$. As a result, matrices that represent the Hermitian product on $\mathbb{C}^n$ take the form P^*P for some invertible matrix P, and matrices that represent the standard dot product over $\mathbb{R}^n$ take the form P^TP. In fact, over $\mathbb{R}^n$ these matrices turn out to be precisely the symmetric positive-definite matrices.*

Orthogonal Complement

Definition 17. *For a subspace $W \subset V$ and an associated bilinear form $\langle \cdot, \cdot \rangle$, the* ***orthogonal complement*** *is defined as $W^\perp = \{v \in V | \langle w, v \rangle = 0\}$. Moreover, vectors v, w such that $\langle v, w \rangle = 0$ are called orthogonal (sometimes denoted as $v \perp w$).*

Lemma 14. *(Parallelogram Law)*

$$2\|x\|^2 + 2\|y\|^2 = \|x + y\|^2 + \|x - y\|^2.$$

A form is called **nondegenerate** if $V^\perp = \{v | v \perp V\} = \{0\}$. Otherwise, it is called **degenerate**. A form is nondegenerate iff the corresponding matrix A is invertible.

Lemma 15. *If $\langle \cdot, \cdot \rangle$ is nondegenerate, $\langle v, w \rangle = \langle v, w' \rangle$ for all $v \in V$ implies $w = w'$.*

In particular, if the bilinear form is nondegenerate on a finite-dimensional vector space, the orthogonal complement satisfies the following properties:

1. $W \cap W^\perp = \{0\}$.
2. $(W^\perp)^\perp = W$.

3. $W \subseteq V \Rightarrow V^\perp \subseteq W^\perp$.

4. $W \oplus W^\perp = V$.

5. $V = W \oplus W^\perp$.

Definition 18. *The* ***orthogonal projection*** *onto* W $\pi : V = W \oplus W^\perp \to W$ *is defined as* $\pi(v) = w$*, where* $v = w + w', w \in W, w' \in W^\perp$.

Theorem 16. *Suppose* $W \subset V$ *is closed. Then, given* $v \in V$*, there exists a unique element* $\pi(v) = w \in W$ *which minimizes* $\|v - w\|$.

Proof. Let $d = \inf \|v - w\|$, and let w_n be a sequence such that $\|v - w_n\| \to d$. Then,

$$\|w_n - w_m\| = 2\|v - w_n\|^2 + \|v - w_m\| - \|w_n + w_m - 2v\|^2.$$

Since $\|w_n + w_m - v\| \geq d$,

$$\|w_n - w_m\| \leq 2\|w_n - v\|^2 + 2\|w_m - v\|^2 - 4d^2 \to 0$$

as $m, n \to \infty$. Now, let w_1, w_2 be two minimal vectors. Then,

$$\begin{aligned}\|w_2 - w_1\|^2 &= 2\|w_1 - v\|^2 + 2\|w_2 - v\|^2 - 4\left\|\frac{w_1 + w_2}{2} - v\right\|^2 \\ &\leq 2d^2 + 2d^2 - 4d^2 = 0,\end{aligned}$$

proving uniqueness. □

Theorem 17. *(Projection Formula) If* $W \subseteq V$*,* $\{w_1, ..., w_m\}$ *is a basis for* W *such that* $\langle \cdot, \cdot \rangle$ *is nondegenerate on* W*, then the projection onto* W *is given by* $\pi(v) = \sum_{i=1}^m c_i w_i$*, where*

$$w_i = \frac{\langle w_i, v\rangle}{\langle w_i, w_i\rangle}.$$

Theorem 18. *For symmetric or Hermitian forms (i.e. symmetric or self-adjoint matrices), there exists an invertible matrix* P *such that* $P^T AP$ *(*$P^\dagger AP$*, respectively) is a diagonal matrix with entries* $1, -1, 0$ *along the diagonal. The number of 1s, -1s, 0s, specified by the triplet* (p, q, r)*, is called the* ***signature*** *of the form. A form is nondegenerate iff* $r = 0$.

Theorem 19. *(Sylvester's Law) The signature of a matrix is independent of the choice of orthogonal basis.*

Theorem 20. *Given a matrix* A*,Let* A_k *be the* $k \times k$ *top-left submatrix of* A*. Then,* A *is positive definite iff* $\det A_k > 0$ *for all* k.

Spectral Theorem

We now consider the various properties of operators with respect to their action on Hermitian forms.

Lemma 21. *Let* T *be an operator and* T^* *be its adjoint. Then,*

1. $\langle Tv, w\rangle = \langle v, T^*w\rangle$ *for all* $v, w \in V$.

2. T *is normal iff* $\langle Tv, Tw\rangle = \langle T^*v, T^*w\rangle$ *for all* $v, w \in V$.

3. T *is Hermitian (self-adjoint) iff* $\langle Tv, w\rangle = \langle v, Tw\rangle$ *for all* $v, w \in V$.

4. T *is unitary iff* $\langle Tv, Tw\rangle = \langle v, w\rangle$ *for all* $v, w \in V$.

We now state the spectral theorems on finite-dimensional vector spaces, which let us obtain results on the diagonalizability of certain matrices.

Theorem 22. *(Spectral Theorem for Normal Operators) Let* A *be a normal matrix. Then, there exists an orthonormal basis of* V *consisting of eigenvectors of* A*, which comprises a unitary matrix* P *such that* P^*AP *is diagonal.*

Remark 13. *Note that self-adjoint, Hermitian, symmetric, skew-symmetric, and orthogonal matrices are all normal, therefore all of them are diagonalizable.*

Rational Canonical Form

For the rest of this section, we need the following lemma.

Lemma 23. *If* $\mathbb{K}$ *is a field then* $\mathbb{K}[x]$ *is a PID.*

Proof. Since $\mathbb{K}$ is a field, $\mathbb{K}[x]$ is a Euclidean domain. It is left to show that for an ideal $I \subset \mathbb{K}[x]$, the nonzero polynomial of lowest degree generates I, which can be done by using the division algorithm. □

Let T be a linear transformation on a vector space V, which thereby makes V an $F[x]$-module. By the Theorem of Finitely Generated Modules over PIDs, any finite-dimensional vector space V is a direct sum of finitely many cyclic $\mathbb{K}[x]$-modules of the form $\mathbb{K}[x]/(a_i(x))$ In other words,

$$V = \mathbb{K}[x]/(a_1(x)) \oplus ... \oplus \mathbb{K}[x]/(a_m(x)),$$

where the **invariant factors** $a_i(x)$ satisfy $a_1(x)|a_2(x)|...|a_m(x)$ and $(a_m(x)) = \mathrm{Ann}(V)$. These elements can be made unique up to unit by being required to be **monic** (i.e. the leading coefficient is 1). It immediately follows that $a_m(x) = m_T(x)$. Now, for any of the direct summands, we construct a basis $1, \bar{x}, \bar{x}^2, ...$, with $a(\bar{x}) = \sum_i b_i \bar{x}^i = 0$. Then, the matrix for *multiplication by* x can be represented as the **companion matrix**

$$\mathcal{C}_{a(x)} = \begin{bmatrix} 0 & 0 & ... & ... & -b_0 \\ 1 & 0 & ... & ... & -b_1 \\ 0 & 1 & ... & ... & -b_2 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & ... & 1 & -b_{k-1} \end{bmatrix}.$$

Then, since V is considered as an $\mathbb{K}[x]$-module, the action of T is equivalent to the action of $\mathcal{C}_{a(x)}$ for each direct summand, implying that in this basis, T is represented by a block-diagonal matrix

$$\begin{bmatrix} \mathcal{C}_{a_0(x)} & & & \\ & \mathcal{C}_{a_1(x)} & & \\ & & ... & \\ & & & \mathcal{C}_{a_m}(x) \end{bmatrix},$$

which is called the **rational canonical form** for the matrix. From this construction, the following properties follow:

1. Every matrix A has a rational canonical form, and that form is unique.

2. A is similar to its rational canonical form. Two matrices are similar iff they have the same rational canonical form.

3. (Cayley-Hamilton Theorem) $c_A(A) = 0$.

4. The characteristic and minimal polynomials $m_{\mathcal{C}_{a(x)}}(x) = c_{\mathcal{C}_{a(x)}}(x) = a(x)$ of the companion matrix of $a(x)$ are both $a(x)$.

5. The characteristic polynomial of A is the product of all invariant factors of A.

Proof. The last statement follows from the fact above and the multiplicativity of the characteristic polynomial for block-diagonal matrices, as well as the fact that similar matrices share the same characteristic polynomial. This immediately proves the Cayley-Hamilton Theorem, since $m_A(x) = a_m(x)$. □

In addition, any matrix $xI - A$ can be brought into the **Smith Normal Form** using elemetary row and column operations (switching two rows/columns, adding a multiple of one to another, and multiplying a row/column by a unit), yielding

$$\begin{bmatrix} 1 & & & & \\ & ... & & & \\ & & a_1(x) & & \\ & & & ... & \\ & & & & a_m(x) \end{bmatrix}$$

Jordan Canonical (Normal) Form

By further breaking down invariant factors into their elementary divisors, we obtain that given a linear transformation represented by a matrix A (since the roots of $a_m(x)$ are the eigenvalues of A) describes the vector space V as a direct sum of $\mathbb{K}[x]$ modules, i.e.

$$V = \bigoplus_i \mathbb{K}[x]/((x - \lambda_i)^k) = \bigoplus_i \ker(A - \lambda_i I)^k,$$

where λ_i are the eigenvalues of A. Similar to the rational canonical form, we let $1, (\bar{x} - \lambda), ..., (\bar{x} - \lambda)^{k-1}$ be a basis for the direct summands, and since $x = \lambda + (x - \lambda)$, one notes that the matrix for multiplication by x in this basis becomes

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ & & & ... & \\ & & & & \lambda \end{bmatrix}.$$

Each such $k \times k$ matrix is called a **Jordan block** of size k with eigenvalue λ. Correspondingly, A can be then represented by a block-diagonal matrix of the form

$$\begin{bmatrix} J_1 & & & & \\ & J_2 & & & \\ & & J_3 & & \\ & & & ... & \\ & & & & J_t \end{bmatrix},$$

with each J_i being a Jordan block, which is called the **Jordan canonical (normal) form** for the matrix. The matrix P such that PAP^{-1} is the Jordan normal form of A consists of the generalized eigenvectors of A. From this construction, the following properties follow:

1. Every matrix A has a Jordan canonical form, and that form is unique up to the permutation of Jordan blocks.
2. A is similar to its Jordan canonical form. Two matrices are similar iff they have the same Jordan canonical form.
3. If A is similar to a diagonal matrix D (i.e. A is diagonalizable), D is the Jordan canonical form of A.
4. If $\mathbb{K}$ contains all the eigenvalues of A, then A is diagonalizable iff $m_A(x)$ is **separable** (i.e. has no repeated roots).
5. Counting multiplicities, the diagonal entries of the Jordan form are the eigenvalues of A.
6. The sum of sizes of all Jordan blocks corresponding to an eigenvalue is the algebraic multiplicity of that eigenvalue.
7. The number of Jordan blocks corresponding to an eigenvalue is the geometric multiplicity of that eigenvalue.

Proof. If A is similar to a diagonal matrix, since similar matrices share the same minimal polynomial, and the minimal polynomial of a diagonal matrix has no repeated roots (its roots are, in fact, the distinct elements along the diagonal), then the minimal polynomial for A has no repeated roots. Conversely, the minimal polynomial for the Jordan normal form is the least common multiple of the minimal polynomials for the Jordan blocks. Since the minimal polynomial for a Jordan block of size k is $(x - \lambda)^k$, it follows that each Jordan block must have size 1, i.e. the Jordan form is a diagonal matrix.

Remark 14. *As a consequence, it follows that every finite-dimensional vector space V has a basis of generalized eigenvalues.*

□

Special Matrices and Their Properties

Normal Matrices

Matrices A over $\mathbb{C}$ such that $A^\dagger A = AA^\dagger$ are called **normal**. Their most important property is that they are diagonalizable over $\mathbb{C}$.

Skew-Symmetric Matrices

Matrices A over $\mathbb{R}$ such that $A = -A^T$ are called **skew-symmetric**. They satisfy the following properties:

- Every diagonal element of A, and therefore the trace of A, is zero.
- $\det A^T = (-1)^n \det A$. In particular, for n odd, $\det A = 0$.
- $\langle Av, w \rangle_\mathbb{R} = \langle v, Aw \rangle_\mathbb{R}$.
- A is normal, therefore diagonalizable.

Symmetric Matrices

Matrices A over $\mathbb{R}$ such that $A = A^T$ are called **symmetric**. They satisfy the following properties:

- The entries on the diagonal of A (and therefore also the trace of A) are real.
- $\langle Av, w \rangle_\mathbb{R} = \langle v, Aw \rangle_\mathbb{R}$.
- A is normal, therefore diagonalizable.

Hermitian Matrices

Matrices A over $\mathbb{C}$ such that $A = A^\dagger$ are called **Hermitian** (or **self-adjoint**). They satisfy the following properties:

- The entries on the diagonal of A (and therefore also the trace of A) are real.
- $\langle Av, w\rangle_\mathbb{C} = \langle v, Aw\rangle_\mathbb{C}$.
- $\det A \in \mathbb{R}$.
- A is normal, therefore diagonalizable.

Orthogonal Matrices

Matrices A over $\mathbb{R}$ such that $AA^T = A^T A = I$ are called **orthogonal.** They satisfy the following properties:

- $A^T = A^{-1}$.
- A is normal, i.e. $AA^\dagger = A^\dagger A$.
- $\det A = \pm 1$.
- $(Av) \cdot (Aw) = v \cdot w$. Equivalently, $\langle Av, Aw\rangle_\mathbb{R} = \langle v, w\rangle_\mathbb{R}$.
- The columns of A form an orthonormal basis of $\mathbb{R}^n$.
- A is normal, therefore diagonalizable over $\mathbb{C}$, with eigenvalues satisfying $|\lambda| = 1$.

Unitary Matrices

Matrices A over $\mathbb{C}$ such that $AA^\dagger = A^\dagger A = I$ are called **unitary.** They satisfy the following properties:

- $A^\dagger = A^{-1}$.
- $|\det A| = 1$.
- $(Av) \cdot (Aw) = v \cdot w$. Equivalently, $\langle Av, Aw\rangle_\mathbb{C} = \langle v, w\rangle_\mathbb{C}$.
- A is normal, therefore diagonalizable over $\mathbb{C}$, with eigenvalues satisfying $|\lambda| = 1$.

Proof. The last one can be proven by noting that for an eigenvector v, $\langle v, v\rangle = \langle Av, Av\rangle = |\lambda|^2 \langle v, v\rangle$, i.e. $|\lambda| = 1$. Diagonalizability follows from the spectral theorem for normal matrices. $\square$

Overall, one recognizes Hermitian matrices as complex versions of symmetric matrices and unitary matrices as complex versions of orthogonal matrices.