Test 1 Review

MATH-2251
Aidan Richer

This document contains all relevant definitions, theorems, and methods for Test 1 from chapters 8-10.

1 Chapter 8: Diagonalization of Matrices

Eigenvalues and Eigenvectors
We begin by defining the notion of an eigenvalue and an eigenvector for both linear maps and $n \times n$ matrices.

Def 8.1.1. Let V be an n dimensional vector space over $\mathbb{F}$ and let $T : V \to V$ be a linear map.

1. A scalar $\lambda \in \mathbb{F}$ is said to be an **eigenvalue** of T if and only if there exists a nonzero vector $\vec{v} \in V$, such that
$$T(\vec{v}) = \lambda\vec{v}$$
2. Every nonzero vector $\vec{v} \in V$ such that $T(\vec{v}) = \lambda\vec{v}$ is said to be an **eigenvector** of T corresponding λ and the λ-eigenspace of T is defined to be $\{\vec{v} \in V \mid T(\vec{v}) = \lambda\vec{v}\}$ which is simply the nullspace of $T - \lambda I_V$.
$$\mathcal{N}(T - \lambda I_V) = \{\vec{v} \in V \mid (T - \lambda I_V)\vec{v} = \vec{0}\}$$
3. The generalized λ-eigenspace of T is defined to be
$$V_\lambda = \{\vec{v} \in V \mid \text{there exists a positive integer n with } (T - \lambda I_V)^n\vec{v} = \vec{0}\}$$

Def 8.1.2. Let A be an $n \times n$ matrix over $\mathbb{F}$.

1. A scalar $\lambda \in \mathbb{F}$ is said to be an **eigenvalue** of A if and only if there exists a nonzero vector $\vec{v} \in \mathbb{F}^n$, such that
$$A\vec{v} = \lambda\vec{v}$$
2. Every nonzero vector $\vec{v} \in \mathbb{F}^n$ such that $A\vec{v} = \lambda\vec{v}$ is said to be an **eigenvector** of A corresponding λ and the λ-eigenspace of A is defined to be $\{\vec{v} \in V \mid A\vec{v} = \lambda\vec{v}\}$ which is simply the nullspace of the matrix $\lambda I_n - A$.
$$\mathcal{N}(\lambda I_n - A) = \{\vec{v} \in V \mid (\lambda I_n - A)\vec{v} = \vec{0}\}$$
3. The generalized λ-eigenspace of A is defined to be
$$V_\lambda = \{\vec{v} \in V \mid \text{there exists a positive integer } n \text{ with } (\lambda I_n - A)^n\vec{v} = \vec{0}\}$$

Def 8.1.3.

1. If V is a finite dimensional vector space over $\mathbb{F}$ and $T : V \to V$ is a linear map then T is said to be **diagonable** provided there exists a basis $\mathcal{B}$ of V such that $[T]_\mathcal{B}$ is diagonal.
2. An $n \times n$ matrix A over $\mathbb{F}$ is said to be **diagonable over** $\mathbb{F}$ provided it is similar to a diagonal matrix - i.e. there is an invertible matrix $P \in \mathcal{M}_{n\times n}(\mathbb{F})$ such that
$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$
where $\lambda_i \in \mathbb{F}$.

Eigenvalues Related to the Characteristic Polynomial

Def 8.3.1. If $A \in \mathcal{M}_{n\times n}(\mathbb{F})$ the **characteristic polynomial** of A is defined to be

$$\chi_A(x) = \det(xI_n - A)$$

Theorem 8.3.1. Fundamental Theorem of Algebra
Every polynomial $f(x) \in \mathbb{C}[x]$ has at least one root in $\mathbb{C}$. In other words, every complex polynomial can be factored into a product of linear polynomials in $\mathbb{C}[x]$. Alternately, the only irreducible polynomials in $\mathbb{C}[x]$ have degree one.

Theorem 8.5.1.
If $A \in \mathcal{M}_{n\times n}(\mathbb{F})$ and $\{\vec{v_1}, \dots, \vec{v_k}\}$ are eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \dots, \lambda_k \in \mathbb{F}$, then $\{\vec{v_1}, \dots, \vec{v_k}\}$ is linearly independent.

Theorem 8.5.2.
If $A \in \mathcal{M}_{n\times n}(\mathbb{F})$ and $\lambda_1, \dots, \lambda_r \in \mathbb{F}$ are distinct eigenvalues of A then the sum of the eigenspaces

$$\mathcal{N}(\lambda_1 I_n - A) + \cdots + \mathcal{N}(\lambda_r I_n - A)$$

is a direct sum

$$\mathcal{N}(\lambda_1 I_n - A) \oplus \cdots \oplus \mathcal{N}(\lambda_r I_n - A)$$

Diagonalizability of a Matrix
Let $A \in \mathcal{M}_{n\times n}(\mathbb{F})$ then A is diagonable over $\mathbb{F}$ provided the following algorithm is successful.

1. Compute the characteristic polynomial: $\chi_A(\lambda) = \det(\lambda I_n - A)$.
2. Factor $\chi_A(\lambda)$ over $\mathbb{F}$. In order for A to be diagonable over $\mathbb{F}$, $\chi_A(\lambda)$ must factor linearly over $\mathbb{F}$.
3. For each distinct eigenvalue λ_i determine a basis $\mathcal{B}_i = \{\vec{v}_{i1}, \dots, \vec{v}_{il_i}\}$ of the λ_i eigenspace $\mathcal{N}(\lambda_i I_n - A)$ of A.
4. Set
$$\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i = \{\vec{v}_{11}, \dots, \vec{v}_{1\ell_1}, \vec{v}_{21}, \dots, \vec{v}_{2\ell_2}, \dots, \vec{v}_{r1}, \dots, \vec{v}_{r\ell_r}\}.$$
If $\mathcal{B}$ is a basis of $\mathbb{F}^n$, then A is diagonable and
$$P^{-1}AP = \begin{bmatrix} \lambda_1 I_{l_i} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_r I_{l_r} \end{bmatrix}$$
where
$$P = \begin{bmatrix} \vec{v_{11}} & \dots & \vec{v_{1l_i}} & \dots & \vec{v}_{rl_r} \end{bmatrix}$$

2 Chapter 9: Inner Product Spaces

Def 9.1.1. Complex and Real Inner Product Spaces
A **complex** ($\mathbb{C}$) inner product space is a complex vector space V together with a map $< \cdot, \cdot >$: $V \times V \to \mathbb{C}$ which satisfies the following properties:

1. $< \vec{v}, \vec{v} > \geq 0$ for all $\vec{v} \in V$, and equal to 0 exactly when $\vec{v} = \vec{0}$
2. $< \vec{v}, \vec{w} > = \overline{< \vec{w}, \vec{v} >}$ for all $\vec{v}, \vec{w} \in V$
3. $< \vec{u}, \vec{v} + \vec{w} > = < \vec{u}, \vec{v} > + < \vec{u}, \vec{w} >$ for all $\vec{u}, \vec{v}, \vec{w} \in V$
4. $< \alpha\vec{v}, \vec{w} > = \overline{\alpha} < \vec{v}, \vec{w} > = < \vec{v}, \overline{\alpha}\vec{w} >$ for all $\vec{v}, \vec{w} \in V$ and $\alpha \in \mathbb{C}$

A **real** ($\mathbb{R}$) inner product space is a real vector space V together with a map $< \cdot, \cdot >$: $V \times V \to \mathbb{R}$ which satisfies the following properties:

1. $< \vec{v}, \vec{v} > \geq 0$ for all $\vec{v} \in V$, and equal to 0 exactly when $\vec{v} = \vec{0}$
2. $< \vec{v}, \vec{w} > = < \vec{w}, \vec{v} >$ for all $\vec{v}, \vec{w} \in V$
3. $< \vec{u}, \vec{v} + \vec{w} > = < \vec{u}, \vec{v} > + < \vec{u}, \vec{w} >$ for all $\vec{u}, \vec{v}, \vec{w} \in V$
4. $< r\vec{v}, \vec{w} > = r < \vec{v}, \vec{w} > = < \vec{v}, r\vec{w} >$ for all $\vec{v}, \vec{w} \in V$ and $r \in \mathbb{R}$

Def 9.1.2. Hermitian Conjugation
Let $A = (a_{ij}) \in \mathcal{M}_{m \times n}(\mathbb{C})$. Then the **hermitian conjugation** of A, denoted A^H or A^*, is the $n \times m$ matrix obtained from A by conjugating each entry and then taking the transpose. $A^H = A^* = (\overline{a}_{ij})^T = (\overline{a}_{ji})$.

The hermitian transpose satisfies the following properties:

1. $(A + B)^H = A^H + B^H$
2. $(A^H)^H = A$
3. $(AB)^H = B^H A^H$
4. $(\alpha A)^H = \overline{\alpha} A^H$
5. $A^H = A^T$ when $A \in \mathcal{M}_{n \times n}(\mathbb{R})$

Def 9.1.4. Norm, Orthogonal, and Orthonormal Inner Product Spaces
Let V be an inner product space.

1. The **norm** of a vector $\vec{v} \in V$, denoted $||\vec{v}||$, is defined to be $\sqrt{<\vec{v}, \vec{v}>}$.
2. Two vectors $\vec{v}, \vec{u} \in V$ are said to be **orthogonal** provided $<\vec{v}, \vec{u}> = 0$.
3. If $\{\vec{v_1}, \ldots, \vec{v_k}\} \subseteq V$, then $\{\vec{v_1}, \ldots, \vec{v_k}\}$ is said to be an **orthogonal** set of vectors provided $\vec{v_1}, \ldots, \vec{v_k}$ are mutually orthogonal vectors - i.e. $<\vec{v_i}, \vec{v_j}> = 0$ for $i \neq j$ and $\{\vec{v_1}, \ldots, \vec{v_k}\}$ is orthonormal provided it is orthogonal and each $\vec{v_i}$ is a unit vector - i.e. $||\vec{v_i}|| = 1$.

Def 9.1.6. Unitary and Orthogonal Matrices
An $n \times n$ complex matrix U is said to be **unitary** provided $U^H = U^{-1}$. An $n \times n$ real matrix Q is said to be **orthogonal** provided $Q^T = Q^{-1}$.

Theorem 9.3.1. Gram-Schmidt Orthogonalization Process
Given any linearly independent set of vectors $\{\vec{v_1}, \ldots, \vec{v_m}\}$ in an inner product space $(V, <\cdot, \cdot>)$ there exists an orthonormal set of vectors $\{\vec{u_1}, \ldots, \vec{u_m}\}$ such that $span\{\vec{v_1}, \ldots, \vec{v_k}\} = span\{\vec{u_1}, \ldots, \vec{u_k}\}$ for all $k = 1, \ldots, m$.

GSOP Alogrithm
For practical implementations of the GSOP we often break the process into two steps.

1. Adjust the angles - construct an orthogonal set $\{\vec{w_1}, \ldots, \vec{w_m}\}$ from $\{\vec{v_1}, \ldots, \vec{v_m}\}$ as follows.
$$\vec{w_1} = \vec{v_1}$$
$$\vec{w_2} = \vec{v_2} - \frac{<\vec{w_1}, \vec{v_2}>}{<\vec{w_1}, \vec{w_1}>}\vec{w_1}$$
$$\vdots$$
$$\vec{w_m} = \vec{v_m} - \frac{<\vec{w}_{m-1}, \vec{v_m}>}{<\vec{w}_{m-1}, \vec{w}_{m-1}>}\vec{w}_{m-1} - \cdots - \frac{<\vec{w_1}, \vec{v_m}>}{<\vec{w_1}, \vec{w_1}>}\vec{w_1}$$
2. Adjust the lengths - normalize the vectors $\vec{w}_1, \ldots, \vec{w}_m$ produced in step 1 to produce an orthonormal set, $\{\vec{u}_1, \ldots, \vec{u}_m\}$ with
$$\vec{u}_i = \frac{\vec{w}_i}{||\vec{w}_i||}$$

Def 9.5.1. U-perp
If U is a subspace of an inner product space $(V, <\cdot, \cdot>)$, then the **orthogonal complement** of U in V, denoted $U^{\perp}$, is defined by

$$U^{\perp} = \{\vec{v} \in \mathbb{R}^n \mid\ <\vec{u}, \vec{v}> = \vec{0} \text{ for all } \vec{u} \in U\}$$

We often read $U^{\perp}$ as U-perp.

Theorem 9.5.1
Let U be a subspace of an inner product space $(V, < \cdot, \cdot >)$. Then

- $U^{\perp}$ is a subspace of V
- $\dim U^{\perp} = \dim V - \dim U$
- $(U^{\perp})^{\perp} = U$
- $V = U \oplus U^{\perp}$

Def 9.7.1 Orthogonal Projections
We define the **orthogonal projection** of V onto U to be the map

$$\mathrm{proj}_U : V \to U$$

such that for any $\vec{v} \in V$, $\mathrm{proj}_U(\vec{v})$ is the unique vector in U such that $\vec{v} - \mathrm{proj}_U(\vec{v}) \in U^{\perp}$. The **perpendicular component** of V with respect to U is the map

$$\mathrm{perp}_U : V \to U^{\perp}$$

such that $\mathrm{perp}_U(\vec{v}) = \vec{v} - \mathrm{proj}_U(\vec{v})$.

Important Remark:
To better visualize this, if $\{\vec{u}_1, \dots, \vec{u}_n\}$ is any orthonormal basis of V such that $\{\vec{u}_1, \dots, \vec{u}_k\}$ is a basis of U then we have for any vector $\vec{v} \in V$

$$\mathrm{proj}_U(\vec{v}) = < \vec{u}_1, \vec{v} > \vec{u}_1 + \cdots + < \vec{u}_k, \vec{v} > \vec{u}_k$$

and

$$\mathrm{perp}_U(\vec{v}) = < \vec{u}_{k+1}, \vec{v} > \vec{u}_{k+1} + \cdots + < \vec{u}_n, \vec{v} > \vec{u}_n$$

Adding them we get,

$$\vec{v} = \mathrm{proj}_U(\vec{v}) + \mathrm{perp}_U(\vec{v})$$

Fundamental Theorem of Linear Algebra Part II
(swap H for $\perp$ when moving between $\mathbb{C}$ and $\mathbb{R}$)

1. If $A \in \mathcal{M}_{m\times n}(\mathbb{C})$ then
   - $\mathcal{N}(A)^\perp = \mathcal{R}(A^H)$
   - $\mathcal{R}(A^H)^\perp = \mathcal{N}(A)$
   - $\mathcal{N}(A^H)^\perp = \mathcal{R}(A)$
   - $\mathcal{R}(A)^\perp = \mathcal{N}(A^H)$
   - $\mathbb{C}^n = \mathcal{R}(A^H) \oplus \mathcal{N}(A)$
   - $\mathbb{C}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^H)$
2. If $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ then
   - $\mathcal{N}(A)^\perp = \mathcal{R}(A^T)$
   - $\mathcal{R}(A^T)^\perp = \mathcal{N}(A)$
   - $\mathcal{N}(A^T)^\perp = \mathcal{R}(A)$
   - $\mathcal{R}(A)^\perp = \mathcal{N}(A^T)$
   - $\mathbb{C}^n = \mathcal{R}(A^T) \oplus \mathcal{N}(A)$
   - $\mathbb{C}^m = \mathcal{R}(A) \oplus \mathcal{N}(A^T)$

3 Chapter 10: The Spectral Theorem

Def 10.1.1 Unitary and Orthogonal Similarity

1. $\mathbb{C}$: If A and B are $n \times n$ complex matrices then we say that A is **unitarily similar** to B if and only if there exists a unitary matrix U over $\mathbb{C}$ such that $B = U^{-1}AU = U^HAU$. Further A is **unitarily diagonable** if and only if it is unitarily similar to a diagonal matrix.
2. $\mathbb{R}$: If A and B are $n \times n$ real matrices then we say that A is **orthogonally similar** to B if and only if there exists an orthogonal matrix Q over $\mathbb{R}$ such that $B = Q^{-1}AQ = Q^TAQ$. Further A is **orthogonally diagonable** if and only if it is orthogonally similar to a diagonal matrix.

Theorem 10.1.1. Unitary and Orthogonal Diagonalizability

1. $\mathbb{C}$: An $n \times n$ complex matrix A is unitarily diagonable if and only if there exists an orthonormal basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ of $\mathbb{C}^n$ consisting of eigenvectors of A - i.e. $A\vec{v}_i = \lambda_i \vec{v}_i$ for $i = 1, \ldots, n$ with $\lambda_i \in \mathbb{C}$. In this case the matrix $U = [\vec{v}_1, \ldots, \vec{v}_n]$ is a unitary matrix which diagonalizes A.

2. $\mathbb{R}$: An $n \times n$ real matrix A is orthogonally diagonable if and only if there exists an orthonormal basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ of $\mathbb{R}^n$ consisting of eigenvectors of A - i.e. $A\vec{v}_i = \lambda_i \vec{v}_i$ for $i = 1, \ldots, n$ with $\lambda_i \in \mathbb{R}$. In this case the matrix $U = [\vec{v}_1, \ldots, \vec{v}_n]$ is an orthogonal matrix which diagonalizes A.

Theorem 10.1.2. Schur's Lemma
Every $n \times n$ matrix A over the complex numbers ($\mathbb{C}$) is unitarily similar to an upper triangular matrix with the eigenvalues of A along the diagonal. In the case of an $n \times n$ matrix A over the real numbers ($\mathbb{R}$), A is orthogonally similar to an upper triangular matrix with the eigenvalues of A along the diagonal provided the $\chi_A(\lambda)$ factors linearly over $\mathbb{R}$.

Def 10.1.2. Normal Complex Matrices
An $(n \times n)$ complex matrix ($\mathbb{C}$) A is said to be **normal** if and only if

$$A^H A = A A^H$$

Theorem 10.1.3. The Spectral Theorem
An $n \times n$ complex ($\mathbb{C}$) matrix A is unitarily diagonable if and only if A is normal. Alternately, A is normal if and only if there exists an orthonormal basis $\mathcal{B} = \{\vec{u}_1, \ldots, \vec{u}_n\}$ of $\mathbb{C}^n$ consisting of eigenvectors of A (- i.e. $A\vec{u}_i = \lambda_i \vec{u}_i$) and hence a unitary matrix $U = [\vec{u}_1, \ldots, \vec{u}_n]$ such that

$$A = UDU^H = \lambda_1 \vec{u}_1 \vec{u}_1^H + \cdots + \lambda_n \vec{u}_n \vec{u}_n^H$$

Notice that $\vec{u}_i \vec{u}_i^H$ is the orthogonal projection of $\mathbb{C}^n$ onto $\mathbb{C}\vec{u}_i$.

Def 10.1.3. Special Matrices
There are three special subclasses of normal complex matrices:

1. **hermitian matrices** $(A^H = A) \implies$ all of its eigenvalues are real.
2. **skew hermitian matrices** $(A^H = -A) \implies$ all of its eigenvalues are purely imaginary.
3. **unitary matrices** $(A^H = A^{-1}) \implies$ all of its eigenvalues have modulus one.

Theorem 10.2.1. Principal Axis Theorem
An $n \times n$ real ($\mathbb{R}$) matrix A is orthogonally diagonable over the real numbers if and only if A is symmetric. Moreover, A is symmetric if and only if there exists an orthonormal basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ of $\mathbb{R}^n$ such that

$$A = \lambda_1 \vec{v}_1 \vec{v}_1^T + \cdots + \lambda_n \vec{v}_n \vec{v}_n^T$$

where $A\vec{v}_i = \lambda_i \vec{v}_i$.