Math 217 List of Definitions Needed For Final

Professor Karen Smith [1]

These are the definitions you must be able to state precisely, below in **boldface.** Important Theorems and Techniques are also discussed.

Words defined here: Subspace of a vector space V. For vectors in a vector space: span, relation, linearly independent, basis. Dimension. Linear transformation $V \xrightarrow{T} W$, and for this linear transformation: the rank, kernel, image of T. Isomorphism $V \xrightarrow{T} W$. The $\mathcal{B}$-coordinates of a vector in V. The change of basis matrix $S_{\mathcal{B}\to\mathcal{A}}$. The $\mathcal{B}$-matrix $[T]_{\mathcal{B}}$ of a transformation. Similar matrices. Inner product. Length (or magnitude) of a vector in an inner product space, length of a vector in $\mathbb{R}^n$. Perpendicularity of two vectors in an inner product space (or in $\mathbb{R}^n$). An orthonormal set of vectors (in $\mathbb{R}^n$ or in an inner product space). The orthogonal complement of a subspace V of $\mathbb{R}^n$. Orthogonal transformation of $\mathbb{R}^n$, orthogonal matrix, symmetric matrix. The least squares solutions to $A\vec{x} = \vec{b}$. Multilinearity property of determinant, alternating property of determinant, multiplicative property of determinant, eigenvector, eigenvalue, eigenspace, characteristic polynomial, algebraic multiplictiy, geometric multiplicity, eigenbasis, diagonalizable, Spectral Theorem.

Chapter 4

Bases and Dimension.

Definition 1. A **subspace** of a vector space V is a subset W which satisfies

(1) The zero vector is in W.
(2) If $\vec{x}, \vec{y} \in W$, then also $\vec{x} + \vec{y} \in W$;
(3) If $\vec{x} \in W$ and k is any scalar, then also $k\vec{x} \in W$.

Definition 2. The **span** of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is

$$Span\,\{\vec{v}_1, \ldots, \vec{v}_n\} \;=\; \{c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n \,|\, c_i \in \mathbb{R}\}.$$

Definition 3. A **relation** on a set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is any expression of the form

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = 0,$$

where the c_i are scalars.

[1] Special thanks to Mark Spencer, Jordan Katz, Dominic Russell, Anthony Zheng and Miranda Riggs for help with this document.

Definition 4. The set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is **linearly independent** if whenever
$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = 0,$$
for some scalars c_i, then $c_1 = c_2 = \cdots = c_n = 0$.

Definition 5. A **basis** for V is a set[2] of vectors which are both *linearly independent* and *span* V.

Theorem 6. *Let V be a vector space.*

(1) *A set of vectors is a basis if and only if it is a minimal spanning set.*
(2) *A set of vectors is a basis if and only if it is a maximal linearly independent set.*

Definition 7. The **dimension** of a vector space is the number (possibly infinite) of elements in a basis.

Book Theorem 3.3.4: Let V be a vector space of dimension m. Then

(1) We can find at most m linearly independent vectors in V.
(2) We need at least m vectors to span V.
(3) Any set of m linearly independent vectors in V is a basis.
(4) Any set of m vectors which spans V is a basis.

Transformations.

Definition 8. Let V and W be vector spaces. A **linear transformation** is a mapping $V \xrightarrow{T} W$ that satisfies

(1) $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$ for all vectors $\vec{x}, \vec{y} \in V$.
(2) $T(k\vec{x}) = kT(\vec{x})$ for all vectors $\vec{x} \in V$ and all scalars k.

Definition 9. An **isomorphism** $V \xrightarrow{T} W$ is a bijective linear transformation.

Recall that **bijective** means the map is both **surjective** and **injective. Surjective** means the image of T is the whole target. **Injective** means that if $x \neq y$ are two elements in the source, then $T(x) \neq T(y)$.

Definition 10. The **image** of a linear transformation $V \xrightarrow{T} W$ is
$$\text{im}T = \{\vec{w} \in W \,|\, T(\vec{v}) = \vec{w} \, \textit{for some} \; \vec{v} \in V\}.$$

[2]More precisely, a basis is an *ordered set* of vectors which both span and are linearly independent: The basis $(\vec{e}_1, \vec{e}_2)$ for $\mathbb{R}^2$ is strictly speaking different from the basis $(\vec{e}_2, \vec{e}_1)$.

Definition 11. The **rank** of a transformation $V \xrightarrow{T} W$ is the dimension of the image.

Definition 12. The **kernel** of a linear transformation $V \xrightarrow{T} W$ is

$$\ker T = \{\vec{v} \in V \,|\, T(\vec{v}) = \vec{0}\}.$$

Proposition 13. *A linear transformation is injective if and only if the kernel is zero. So $V \xrightarrow{T} W$ is an isomorphism if and only if $\ker T = 0$ and im $T = W$.*

Definition 14. The **standard** matrix of a linear transformation $\mathbb{R}^m \xrightarrow{T} \mathbb{R}^n$ is the unique $n \times m$ matrix A such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^m$.
This is the same as the **matrix of T in the standard basis.**[3]
In this case:

- The rank of T is the same as the rank of the matrix A.
- The image of T is the span of the columns of A.
- T is an isomorphism if and only if A is invertible, which is the same as $\det A \neq 0$.
- The kernel of A is the solution set of $A\vec{x} = \vec{0}$.

The Rank-Nullity Theorem:[4] *Let $V \xrightarrow{T} W$ be a linear transformation, where V is finite dimensional. Then*

$$\dim \operatorname{im} T + \dim \ker T = \dim V.$$

The rank nullity theorem easily implies the following useful facts:

Book Theorem 4.2.4: Let $V \xrightarrow{T} V$ be a linear transformation of **finite** dimensional vector spaces of the **same** dimension. The following are equivalent:

(1) T is surjective
(2) T is injective
(3) T is an isomorphism.

CAUTION: The theorem 4.2.4 above is **false** for infinite dimensional vector spaces!

Theorem 15. *Two finite dimensional vector spaces V and W are isomorphic if and only if they have the same dimension.*

[3]CAUTION: If the source and target of a linear transformation are not just coordinate spaces $\mathbb{R}^n$, then there is no obvious choice of a matrix. Only after picking a basis $\mathcal{B}$ can we get a matrix, called the $\mathcal{B}$-matrix. More soon.

[4]Remember $\dim \operatorname{im} T$ is called the rank of T. An old word for $\dim \ker T$ is nullity. Thus the name.

Caution: This theorem says that any two vector spaces of dimension 17 are isomorphic. However, *does not say* that every linear map between them is an isomorphism—only that *some map* between them is. For example, the linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$ projecting onto some line is **not** an isomorphism, though of course the source and target are isomorphic here.

Coordinates and $\mathcal{B}$-matrices.

Definition 16. Coordinates: Let V be a finite dimensional vector space with basis $\mathcal{B} = (\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n)$. The **$\mathcal{B}$-coordinates** of $\vec{v}$ are the unique scalars a_i such that

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n,$$

arranged into a column vector

$$[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Definition 17. Change of Basis Matrix: Let $\mathcal{A}$ and $\mathcal{B}$ be two bases for V. The map sending

$$[\vec{x}]_{\mathcal{B}} \mapsto [\vec{x}]_{\mathcal{A}}$$

for each $\vec{x} \in V$ is a linear transformation. Its standard matrix $S_{\mathcal{B}\to\mathcal{A}}$ is called the **change of basis matrix** from $\mathcal{B}$ to $\mathcal{A}$.

You must know how to compute and interpret change of basis matrices! Here's the formula: if $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$, then

$$S_{\mathcal{B}\to\mathcal{A}} = [[\vec{v}_1]_{\mathcal{A}} \ \ [\vec{v}_2]_{\mathcal{A}} \ \ \ldots \ \ [\vec{v}_n]_{\mathcal{A}}].$$

This is the unique matrix such that

$$S_{\mathcal{B}\to\mathcal{A}}[\vec{v}]_{\mathcal{B}} = [\vec{v}]_{\mathcal{A}}$$

for all $\vec{v} \in V$.

Proposition 18. $S_{\mathcal{A}\to\mathcal{B}} = [S_{\mathcal{B}\to\mathcal{A}}]^{-1}$.

Definition 19. $\mathcal{B}-$matrix: Let $V \xrightarrow{T} V$ be a linear transformation. Let $\mathcal{B}$ be a basis for V. The map

$$[\vec{x}]_{\mathcal{B}} \mapsto [T(\vec{x})]_{\mathcal{B}},$$

for all $\vec{x} \in V$, is a linear transformation. Its matrix is called the **$\mathcal{B}-$matrix** of T.

You must know how to compute and interpret $\mathcal{B}$-matrices! If $\mathcal{B} = (\vec{v}_1, \dots, \vec{v}_n)$, the $\mathcal{B}$-matrix is the $n \times n$ matrix

$$[T]_{\mathcal{B}} = \left[[T(\vec{v}_1)]_{\mathcal{B}} \quad [T(\vec{v}_2)]_{\mathcal{B}} \quad \dots \quad [T(\vec{v}_n)]_{\mathcal{B}} \right].$$

This is the unique matrix such that

$$[T]_{\mathcal{B}} \, [\vec{v}]_{\mathcal{B}} = [T(\vec{v})]_{\mathcal{B}}$$

for every vector $\vec{v} \in V$.

The $\mathcal{B}$-matrix lets us compute $T(\vec{v})$ for any $\vec{v}$ as follows: the $\mathcal{B}$-coordinate column vector of $T(\vec{v})$ is the column obtained by multiplying the $\mathcal{B}$-coordinates of $\vec{v}$ by the $\mathcal{B}$-matrix $[T]_{\mathcal{B}}$.

Theorem 20. *Let $V \xrightarrow{T} V$ be a linear transformation, where V is finite dimensional. Suppose that $\mathcal{B}$ and $\mathcal{A}$ are two different bases for V. Then*

$$[T]_{\mathcal{B}} = [S_{\mathcal{B} \to \mathcal{A}}]^{-1} \, [T]_{\mathcal{A}} \, S_{\mathcal{B} \to \mathcal{A}} = S_{\mathcal{A} \to \mathcal{B}} \, [T]_{\mathcal{A}} \, S_{\mathcal{B} \to \mathcal{A}}$$

where $S = S_{\mathcal{B} \to \mathcal{A}}$ is the change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$ coordinates.

Definition 21. We say that two $n \times n$ matrices A and B are **similar** when there exists an invertible $n \times n$ matrix S such that $B = S^{-1}AS$.

So Theorem 20 says that the matrices of T in different bases are all similar to each other.

The change of basis matrix $S_{\mathcal{B} \to \mathcal{A}}$ is especially easy to find when $\mathcal{A}$ is an especially nice basis. One example is when $V = \mathbb{R}^n$ and $\mathcal{A}$ is the standard basis. For example, if $\mathcal{B} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$ and $\mathcal{E} = (\vec{e}_1, \dots, \vec{e}_n)$, then

$$S_{\mathcal{B} \to \mathcal{E}} = \left[\vec{v}_1 \quad \vec{v}_2 \quad \dots \quad \vec{v}_n \right].$$

Chapter 5

Definition 22. The length $||\vec{v}||$ (or magnitude) of a vector $\vec{v}$ in $\mathbb{R}^n$ is $(\vec{v} \cdot \vec{v})^{1/2}$.

Definition 23. Vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$ are **perpendicular** (or **orthogonal**) if $\vec{v} \cdot \vec{w} = 0$.

Definition 24. Vectors $\{\vec{u}_1, \dots, \vec{u}_n\}$ are an **orthonormal set** if $\vec{u}_i \cdot \vec{u}_j = 0$ unless $i = j$, in which case $\vec{u}_i \cdot \vec{u}_i = 1$. Equivalently, they are unit length and mutually perpendicular.

Theorem 25. *An orthonormal set of vectors is linearly independent.*

Definition 26. The **orthogonal complement** of a subspace V of $\mathbb{R}^n$ is the subspace

$$V^{\perp} = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} \cdot \vec{v} = 0 \text{ for all } \vec{v} \in V\}.$$

Theorem 27. *Let V be a subspace of $\mathbb{R}^n$. Then $V \cap V^{\perp} = 0$ and $\dim V + \dim V^{\perp} = n$. Furthermore, $(V^{\perp})^{\perp} = V$.*

Definition 28. Orthogonal Projection: Every vector in $\vec{x}$ in $\mathbb{R}^n$ can be written uniquely as $\vec{x}^{||} + \vec{x}^{\perp}$, where $\vec{x}^{||} \in V$ and $\vec{x}^{\perp} \in V^{\perp}$. The linear transformation taking each vector $\vec{x}$ to $\vec{x}^{||}$ is called the **projection onto** V.

Check your understanding: be sure you see why projection onto V has image V and kernel $V^{\perp}$. Also: be sure you see that this map fixes all the elements of V.

Make sure you understand this definition geometrically. We can imagine the vector $\vec{x}^{\perp}$ as a perpendicular line segment "dropped from the head of $\vec{x}$ down to V." The difference, $\vec{x}^{||}$, lies in V. This $\vec{x}^{||}$ is the projection of $\vec{x}$ onto V. Put differently, the *projection of $\vec{x}$ to V is the closest vector in V to $\vec{x}$.*

You must also know how to compute the projections. Here's how:

Theorem 29. *Let V be a subspace of $\mathbb{R}^n$ and $(\vec{u}_1, \ldots, \vec{u}_d)$ be an orthonormal basis for V. Then the orthogonal projection onto V is*

$$\vec{x} \mapsto proj_V(\vec{x}) = (\vec{x} \cdot \vec{u}_1)\vec{u}_1 + (\vec{x} \cdot \vec{u}_2)\vec{u}_2 + \cdots + (\vec{x} \cdot \vec{u}_d)\vec{u}_d.$$

Theorem 30. *Every subspace of $\mathbb{R}^n$ has an orthonormal basis.*

A good technique for finding an orthonormal basis is **Gram Schmidt Orthonormalization**. This technique takes any basis to an orthonormalized basis. You need to know how to do this and what, geometrically, is going on as you do each step.

You should also know how to compute a QR factorization.

Orthogonal Transformations.

Definition 31. A linear transformation $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^n$ is **orthogonal** if $||T(\vec{x})|| = ||\vec{x}||$ for all $\vec{x} \in \mathbb{R}^n$.

Theorem 32. *The transformation $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^n$ is orthogonal if and only if $T(\vec{x}) \cdot T(\vec{y}) = \vec{x} \cdot \vec{y}$ for all $\vec{x}, \vec{y} \in \mathbb{R}^n$.*

The theorem says that a linear transformation T is orthogonal exactly when T respect dot products. Geometrically, this means T preserves the length of every vector, as well as the angles between vectors.

Theorem 33. *A composition of orthogonal transformations is orthogonal.*
Orthogonal transformations are invertible; the inverse of an orthogonal transformation is also orthogonal.

Definition 34. A matrix A is **orthogonal** if it is **square** $n \times n$ and $A^T A = I_n$.

Theorem 35. *The following are equivalent for any $n \times n$ (square!) matrix A:*

(1) *A is orthogonal*
(2) *A is the (standard) matrix of an orthogonal transformation.*[5]
(3) $AA^T = I_n$.
(4) *The columns of A are orthonormal.*
(5) *The rows of A are orthonormal.*

Theorem 36. *Rank A = Rank A^T. Also $(AB)^T = B^T A^T$ and $(A + B)^T = A^T + B^T$.*

Definition 37. A matrix is symmetric if $A = A^T$.

Least Squares. Recall that a system $A\vec{x} = \vec{b}$ is consistent if and only if $\vec{b}$ is in the image of A.

Definition 38. The **least squares** solutions to the system $A\vec{x} = \vec{b}$ are the actual solutions to the consistent system $A\vec{x} = proj_{im\,A}(\vec{b})$. We think of these as the solutions to the "closest consistent system."

CAUTION: The least squares solutions to the system $A\vec{x} = \vec{b}$ are not actually solutions at all! They are only approximate solutions. They are the closest you can get to solutions.

Theorem 39. *The least squares solutions to the system $A\vec{x} = \vec{b}$ are the same as the actual solutions to $A^T A\vec{x} = A^T\vec{b}$.*

Moving matrices across Dot product:

Theorem 40. *Let A be any $m \times n$ matrix. Then $A\vec{x} \cdot \vec{y} = \vec{x} \cdot A^T\vec{y}$.*

Theorem 41. *Let A be any $m \times n$ matrix. Then $(\ker A)^\perp = im\, A^T$.*

[5]Caution! This is NOT the same as an orthogonal projection, which is never an orthogonal transformation unless it is the identity map.

Inner Product. An inner product combines two vectors in a vector space V to get a **scalar,** just like the dot product in $\mathbb{R}^n$. Indeed, dot product is the most familiar example of an inner product.

Definition 42. An **inner product** on a vector space V is a function

$$V \times V \longrightarrow \mathbb{R}$$

which assigns to each pair of vectors f, g some scalar $\langle f, g\rangle$, called their **inner product.** An inner product must satisfy the following axioms:

(1) Symmetry: $\langle f, g\rangle = \langle g, f\rangle$ for all vectors $f, g \in V$;
(2) Linearity in each factor: $\langle (f+g), h\rangle = \langle f, h\rangle + \langle g, h\rangle$ for all vectors $f, g, h \in V$ and $\langle \lambda f, g\rangle = \lambda\langle f, g\rangle$ for all scalars λ and all $f, g \in V$.
(3) Positive Definiteness: $\langle f, f\rangle \geq 0$ for all $f \in \vec{V}$ with $\langle f, f\rangle = 0$ only if $f = 0$.

A vector space V together with a choice of an inner product is called an **inner product space.**

Definition 43. The **magnitude** of a vector $\vec{v}$ in an inner product space, denoted $||\vec{v}||$, is the scalar $\langle \vec{v}, \vec{v}\rangle^{1/2}$. We say that $\vec{w}$ is **perpendicular** (or **orthogonal**) to $\vec{v}$ if $\langle \vec{v}, \vec{w}\rangle = 0$.

Definition 44. The orthogonal complement of a subspace W of an inner product space V is the set

$$W^{\perp} =: \{\vec{v} \in V \mid \langle v, w\rangle = 0 \,\text{for all}\, \vec{w} \in W\}.$$

Orthonormality.

Definition 45. A set of vectors $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ in an inner product space is **orthonormal** if $\langle u_i, u_j\rangle = 0$ for $i \neq j$, and $\langle u_i, u_i\rangle = 1$ for all i.

The Gram-Schmidt orthonormalization technique works in any vector space! Just replace dot product by the inner product and go! Similarly, we can compute the projection to a subspace of an inner product space just as we did in $\mathbb{R}^n$. That is, if $\vec{u}_1, \ldots, \vec{u}_d$ form an orthonormal basis for a subspace W of an inner product space V, then the projection of any vector $\vec{x} \in V$ onto W is:

$$\langle \vec{x}, \vec{u}_1\rangle\vec{u}_1 + \langle \vec{x}, \vec{u}_2\rangle\vec{u}_2 + \cdots + \langle \vec{x}, \vec{u}_d\rangle\vec{u}_d.$$

The projection on $\vec{x}$ to the subspace W is the **closest vector in** W to the vector $\vec{x} \in V$, just like the case of $\mathbb{R}^n$. The case of $\mathbb{R}^n$ is the special case where the inner product space is just coordinate space $\mathbb{R}^n$ with standard dot product for the inner product. The standard dot product of course induces the standard notion of distance, so the "closest vector to W" agrees with our usual intuition.

The idea of "finding the closest vector to W" is important in many branches of math and applications. For example, it can be used to approximate a crazy function g by a polynomial function: simply projecting onto the space of polynomials to find the "closest" polynomial.

To do this, we need that g is a function in some inner product space which contains the space of polynomials. For example, this will work if g is a conitnuous (or even piecewise continuous) function on some interval $[a, b]$. Then we can define an inner product using in integration, eg $\langle f, g\rangle = \int_a^b fg\,dx$.

EXAMPLE: The space $\mathbb{R}^{2\times 2}$ is an inner product space with the inner product $\langle A, B\rangle trace(A^T B)$. The standard basis $(E_{11}, E_{12}, E_{21}, E_{22})$ is an orthonormal basis, as you can check. The closest diagonal matrix to the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is the projection of this matrix onto the space of diagonal matrices. Since (E_{11}, E_{22}) is an orthonormal basis for the subspace of diagonal matrices, this projection is

$$\langle A, E_{11}\rangle E_{11} + \langle A, E_{22}\rangle E_{22} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}.$$

Fortunately, this agrees pretty well with our intuition of what ought to be the closest diagonal matrix.

Pretty much everything one can do with dot product can be done in any inner product space. This includes computing lengths of vectors, finding orthonornal bases, computing the orthogonal complement of a space, computing the projection onto a subspace. All the techniques are the same, just replacing dot product with the inner product at each step.

CHAPTER 6

Computing the Determinant. Determinants are defined only for square matrices. The determinant of a 1×1 matrix is simply its unique scalar entry. The determinant for larger size matrices is defined inductively using a technique called **Laplace expansion.**

Definition 46. Let A be an $n \times n$ matrix. The **Laplace expansion along row i of A** is the expression

$$\sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij},$$

where A_{ij} is the submatrix of A obtained by deleting row i and column j.

Likewise the **Laplace expansion along column j of A** is the expression

$$\sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}.$$

Theorem 47. *Let A be an $n \times n$ matrix. The Laplace expansion along any row or column gives the same scalar.*

Definition 48. Let A be a square matrix. The **determinant** of A is the scalar obtained by Laplace expansion along any row or column of A.

This definition is the most useful for computations, but perhaps not ideal (since, for example, it depends on Theorem 47, which is quite a beast to prove). In Math 217, we focus on how to compute and use the determinant, rather than the definition.

Properties of the Determinant.

Proposition 49. *The determinant of A and A^T are equal for any $n \times n$ matrix A.*

Theorem 50. *The determinant is* **multiplicative:** *That is,*

$$det(AB) = det A \, det B$$

for any $n \times n$ matrices A and B.

Corollary 51. *The determinant of an orthogonal matrix is ± 1.*

Theorem 52. *A square matrix is invertible if and only if its determinant is non-zero.*

Corollary 53. *The following are equivalent statements about a $n \times n$ matrix A:*

(1) $\det A = 0$
(2) *The rank of A is less than n*
(3) *There is a non-trivial relation on the columns of A.*
(4) *There is a non-trivial relation on the row of A.*
(5) *A is not invertible.*

Theorem 54. *Similar matrices have the same determinant.*

Definition 55. The **determinant** of a *linear transformation* $V \xrightarrow{T} V$ (where V is finite dimensional) is the determinant of any matrix $[T]_{\mathcal{B}}$ representing T.

This definition makes sense because even though the matrices $T_{\mathcal{B}}$ and $T_{\mathcal{A}}$ are different for different choices of bases $\mathcal{B}$ and $\mathcal{A}$, we know that they are similar, and so they have the same determinant.

Proposition 56. *Let V be a finite dimensional vector space. A linear transformation $V \xrightarrow{T} V$ is an isomorphism if and only if $det\, T$ is not zero.*

Theorem 57. *Let $\mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation given by multiplication by the $n \times n$ matrix A. Then the n-volume of the image of the unit n-cube is $|det\, A|$.*

Multilinearity. The determinant has an important property called **multilinearity** in the rows and columns. This means that fixing all but one column, the determinant is a linear function in the remaining column. Precisely:

Proposition 58. *Fix $n-1$ column vectors $\vec{v}_1, \dots, \vec{v}_{n-1}$ in $\mathbb{R}^n$. Then the function*

$$\mathbb{R}^n \to \mathbb{R}$$

$$\vec{x} \mapsto det \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_{n-1} & \vec{x} \end{bmatrix}$$

is a linear transformation. The same is true if $\vec{x}$ is inserted as the j-th column (instead of the n-th) for any j. A similar statement holds for the rows.

Example 59. Because the determinant is linear in the first column, we have

$$det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} + det \begin{bmatrix} 0 & 2 & 3 \\ 0 & 5 & 6 \\ 100 & 8 & 9 \end{bmatrix} = det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 107 & 8 & 9 \end{bmatrix}.$$

Note that here, the second two columns are the same, only the first column has elements being added. Also

$$det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} + 2det \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 7 & 8 & 9 \end{bmatrix} = det \begin{bmatrix} 1 & 2 & 3 \\ 6 & 7 & 8 \\ 7 & 8 & 9 \end{bmatrix}$$

by linearity in the middle row.

Another important property is the **alternating property** of the determinant:

Proposition 60. *Let A be a square matrix. Let A' be the matrix obtained by swapping two columns. Then $\det A = -\det A'$. The same holds if we swap two rows.*

Note that this implies that if we make k swaps of rows/columns, the determinant is multiplied by $(-1)^k$.

Chapter 7

Definition 61. An **eigenvector** of a linear transformation $V \xrightarrow{T} V$ is any non-zero vector $\vec{v} \in V$ such that $T(\vec{v}) = \lambda\vec{v}$ for some scalar λ. The scalar λ is called the **eigenvalue** of the eigenvector $\vec{v}$.

Caution: Not every transformation has an eigenvector!

Definition 62. Let $V \xrightarrow{T} V$ be a linear transformation. An **eigenbasis** is a basis for V consisting of eigenvectors for T.

Caution: Not every transformation has an eigenbasis!

Theorem 63. *Let $V \xrightarrow{T} V$ be a linear transformation, where V is finite dimensional. Then a basis $\mathfrak{B}$ is an eigenbasis if and only if the $\mathfrak{B}$-matrix $[T]_{\mathfrak{B}}$ is diagonal.*

To check your understanding, make sure you see why Theorem 63 is true.

Example The map $\mathbb{R}^{2\times 2} \xrightarrow{T} \mathbb{R}^{2\times 2}$ sending A to $A + A^T$ is a linear map. Note that if A is symmetric, then $T(A) = 2A$, so the set of symmetric matrices is contained in the 2-eigenspace. The symmetric matrices have basis $(E_{11}, E_{22}, E_{12} + E_{21})$. Are there any other eigenvalues and vectors? Well, also $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ can be easily seen to be in the kernel, so it is a zero eigenvector, and must span the zero-eigenspace. So $\mathfrak{B} = (E_{11}, E_{22}, E_{12} + E_{21}, E_{12} - E_{21})$ is an eigenbasis for T. The corresponding diagonal $\mathfrak{B}$-matrix is $[T]_{\mathfrak{B}} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Definition 64. A linear transformation $V \xrightarrow{T} V$ is **diagonalizable** if V admits an eigenbasis for T. A matrix A is diagonalizable if the linear transformation defined by left multiplication by A is diagonalizable.

The following is an alternative (but equivalent) way to define a diagonalizable matrix:

Theorem 65. *The matrix is A is diagonalizable if and only if there exists an invertible matrix S such that $S^{-1}AS$ is diagonal. In this case, the columns of S form an eigenbasis for A.*

Be sure you understand **why** the above characterization of a diagonalizable matrix is equivalent to having an eigenbasis. Note that using the change of basis formulation, if $\mathfrak{B}$ is an eigenbasis for A and $\mathcal{E}$ denotes the standard basis, then

$$[A]_{\mathfrak{B}} = S_{\mathcal{E}\to\mathfrak{B}}[A]_{\mathcal{E}}S_{\mathfrak{B}\to\mathcal{E}} = S^{-1}AS$$

where S is the matrix made from the columns of $\mathfrak{B}$.

Definition 66. Let λ be an eigenvalue of a linear transformation T. The subset of V

$$V_\lambda = \{\vec{v} \in V \mid T(\vec{v}) = \lambda\vec{v}\}$$

is called the λ-eigenspace of T. Its dimension is the **geometric multiplicity** of λ.

The λ-eigenspace is a *subspace* of V. It consists of all the λ-eigenvalues, together with the zero vector.

Remark 67. The book discusses *eigenvectors, eigenvalues, eigenspaces and eigenbasis* of a **matrix** A. In all cases, this just means the eigen-whatever of the transformation $\mathbb{R}^n \to \mathbb{R}^n$ given by left multiplication by A. This abuse of terminology is consistent with the book's general convention of thinking of a matrix as a linear transformation. Be sure you see why these definitions agree with the definitions in the early sections of Chapter 7. (The book considers vector spaces that are not just $\mathbb{R}^n$ only in 7.4).

In practice, we find eigenspaces for a linear transformation of a *finite dimensional space* using the following practical technique. Be sure you understand *why* this works:

Proposition 68. *Fix an eigenvalue λ of the $n \times n$ matrix A. The λ-eigenspace of A is the space*

$$V_\lambda = \ker(A - \lambda I_n).$$

Rank-nullity is super useful for computing geometric multiplicity. For example, you should immediately verify that $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ has eigenvalue 0 with geometric multiplicity 1.

Consider a transformation $V \xrightarrow{T} V$ where V is finite dimensional. We can fix a basis $\mathcal{B}$ so as to identity T with matrix multiplication $\mathbb{R}^n \xrightarrow{B} \mathbb{R}^n$ between the space of $\mathcal{B}$-coordinates of V. Here of course B is the $\mathcal{B}$-matrix of T, namely $[T]_\mathcal{B}$. The eigenvalues, eigenvectors, and eigenspaces can be computed using the matrix B. Don't forget to reinterpret the elements back in V again using the basis $\mathcal{B}$!

Theorem 69. *Let $V \xrightarrow{T} V$ be a linear transformationm where V is n dimensional. Then T has an eigenbasis if and only if the sum of the geometric multiplicities of all its eigenvalues adds up to n.*

Restating the theorem with matrices we have:

Theorem 70. *Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if the sum of the geometric multiplicities of all its eigenvalues adds up to n.*

You should make sure you understand why these two theorems are equivalent.

Definition 71. Consider an $n \times n$ matrix A. The **characteristic polynomial** of A is the degree n polynomial

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

Definition 72. Let $V \xrightarrow{T} V$ be a linear transformation on a vector space V of finite dimension n. The **characteristic polynomial** of T is the degree n polynomial

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

where A is the matrix of T in *any* basis for V.

NOTE: The book's formulation of the characteristic polynomial is somewhat clumsy. It is perhaps helpful to write it as a polyomial in $(-x)$:

$$(-x)^n + a_1(-x)^{n-1} + a_2(-x)^{n-2} + \cdots + a_n$$

. With this formulation, we have that $a_n = \det A$ and $a_1 = \text{trace } A$.

Example 73. Suppose that $V \xrightarrow{T} V$ is diagonalizable. This means there exists a basis $\mathcal{B}$ (an eigenbasis) such that

$$[T]_{\mathcal{B}} = \begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ 0 & \vdots & \dots & \vdots & 0 \\ 0 & 0 & \dots & 0 & a_{nn} \end{bmatrix}.$$

The diagonal elements a_{ii} are the eigenvalues of T. Computing the characteristic polynomial, we see that

$$\chi_T(x) = (x - a_{11})(x - a_{22}) \cdots (x - a_{nn}),$$

where the a_{ii} are the eigenvalues of T (possibly repeated multiple times).

Remark 74. The characteristic polynomial may or may not have any real roots at all! Remember, however, that over the complex numbers, every polynomial factors completely into linear factors

$$(x - \lambda_1)^{a_1}(x - \lambda_2)^{a_2} \cdots (x - \lambda_t)^{a_t}$$

where the roots $\lambda_i \in \mathbb{C}$. The roots of the characteristic polynomial will be called eigenvalues *even when they are complex.*[6]

Definition 75. Let $V \xrightarrow{T} V$ be a linear transformation of a finite dimensional vector space V. The **algebraic multiplicity** of an eigenvalue λ is the largest power r such that $(x - \lambda)^r$ divides the characteristic polynomial.

Theorem 76. *For each eigenvalue of a linear transformation, the geometric multiplicity is at most the algebraic multiplicity: gemu*$(\lambda) \leq$ *almu* (λ).

Corollary 77. *If a linear transformation of an n-dimensional space has n distinct eigenvalues, then it is diagonalizable.*

The **Spectral Theorem** is really remarkable:

Theorem 78. *A symmetric $n \times n$ matrix has an orthonormal eigenbasis.*

Stated in terms of matrices, this becomes

[6] If λ is an eigenvalue of a matrix A, then it turns out that there is a column vector with *complex entries* $\vec{v}$ such that $A\vec{v} = \lambda\vec{v}$.

Theorem 79. *Let A be a symmetric $n \times n$ matrix. There exists an orthogonal matrix S such that $S^T AS$ is diagonal.*

To check your understanding: be sure you see why the columns of S in the theorem above form an eigenbasis for A, and why the elements on the diagonal of the diagonal matrix are all eigenvalues.