Linear Algebra I: The Good Stuff

Abraham D. Smith

February 3, 2015

Preface

This packet works through the last topics for Linear Algebra I. Along with row-reduction (which we've seen plenty of times before), these are precisely the most-used tools in applications of linear algebra.

- There are two versions of this document, one called `goodstuffT.pdf` formatted for reading on tablets or printing 2-up on paper and `goodstuffP.pdf` formatted for reading normally on paper.
- Be sure that you can do all the Ex 0.0: exercises ! These are the most significant statements in this subject.
- I have stopped putting arrows over vectors, so $\vec{v}$ is now just v. You should be able to figure out what is a vector and what is a scalar in context.
- We are only concerned with finite-dimensional vector spaces, because we want to be able to write finite bases. A great deal of this can be extended to infinite-dimensional spaces, but that is a topic for another semester.
- This document will be updated daily as typos are found. Check the compile date. Ex 0.1: The first person to report any particular typo gets credit. The values are +1 point for spelling and grammar, +2 for math typos, and +4 for logical errors.

Contents

1 Linear Transformations

1.1 Basic Definitions

Given vector spaces V and W, a **linear transformation** is a function $F : V \to W$ with the property that, for any $r, s \in \mathbb{R}$ and $u, v \in V$,

$$F(ru + sv) = rF(u) + sF(v). \tag{1.1}$$

The **kernel** of F is the set $\ker F = \{v \in V \ : \ F(v) = 0\}$. The **image** of F is the set $\operatorname{im} F = \{w \in W \ : \ \exists v \in V, F(v) = w\} = F(V)$.

Linear transformations are also called **linear maps** or **homomorphisms of vector spaces**. If F is one-to-one, then we call F **injective** and write $F : V \hookrightarrow W$. If F is **onto**, then we call F **surjective** and write $F : V \twoheadrightarrow W$. If F is both injective and surjective, we call F **bijective** or say F is an **isomorphism**, writing $F : V \xrightarrow{\sim} W$. In that case, we would say V and W are **isomorphic** via F, writing $V \cong W$.

Ex 1.1: Prove that $\ker F$ and $\operatorname{im} F$ are subspaces of V and W, respectively.

Ex 1.2: Prove that $\ker F = \{0\}$ if and only if F is injective.

Ex 1.3: Prove that, if F is an isomorphism, then V and W have the same dimension.

Lemma 1.1 (Transformations to Matrices)**.** *Given bases for the domain and range, a linear transformation can be represented by an $s \times n$ matrix. This matrix representation is unique once the bases are chosen.*

Proof. Suppose that $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a basis for V and that $(\beta_1, \beta_2, \ldots, \beta_s)$ is a basis for W. Then there are numbers f_{ik} for $1 \le i \le n$ and $1 \le k \le s$ such that

$$F(\alpha_i) = \sum_{k=1}^{s} f_{ik}\beta_k, \ \forall i = 1, 2, \ldots, n. \tag{1.2}$$

These numbers f_{ik} have no ambiguity, since any vector in W has a unique decomposition in the basis β. Let A be the $s \times n$ matrix whose (k, i) entry $A_{k,i}$ is the number f_{ik}. (That is, if we think about (f_{ik}) as an $n \times s$ matrix, then $A = f^T$.) For any $v \in V$, we can decompose it in terms of the basis α in a unique way:

$$v = \sum_i v_i\alpha_i = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}_\alpha$$

Therefore, applying the linear transformation we have

$$F(v) = F\left(\sum_i v_i\alpha_i\right) = \sum_i v_i F\left(\alpha_i\right) = \sum_i v_i \sum_k f_{ik}\beta_k = \sum_{i,k} A_{k,i} v_i \beta_k = \begin{bmatrix} \sum_i A_{1,i}v_i \\ \sum_i A_{2,i}v_i \\ \vdots \\ \sum_i A_{s,i}v_i \end{bmatrix}_\beta . \tag{1.3}$$

In other words, $F(v) = w$ corresponds to the matrix equation

$$\begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s,1} & A_{s,2} & \cdots & A_{s,n} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}_\alpha = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_s \end{bmatrix}_\beta \tag{1.4}$$

□

Corollary 1.2. *If a basis (β_k) for W is understood, then* $\operatorname{Col} A = \operatorname{im} F$.

Proof. Ex 1.4: Combine your earlier proof that $\operatorname{im} F$ is a subspace of W with the previous lemma. □

1.2 Example: Linear Transformations to Matrices

Let $V = \mathcal{P}_3$ and $W = \mathcal{P}_5$. Let F be the function "integrate from 0." That is, if $p \in \mathcal{P}_3$, then $F(p) = \int_0^x p(\tau)\mathrm{d}\tau$. In calculus, you proved that $F(c_1p_1 + c_2p_2) = c_1F(p_1) + c_2F(p_2)$ for any scalars c_1 and c_2 and any "vectors" p_1 and p_2 in $\mathcal{P}_3$, so this is a linear transformation.

Let $\alpha_1 = 1$, $\alpha_2 = x - 1$, $\alpha_3 = (x-1)^2$, and $\alpha_4 = (x-1)^3$ be a basis for V. Let $\beta_1 = 25(x-1)^5$, $\beta_2 = 16(x-1)^4$, $\beta_3 = 9(x-1)^3$, $\beta_4 = (x-1)^2$, $\beta_5 = x+1$, and $\beta_6 = -1$ be a basis for W. Ex 1.5: Verify that these are bases of their respective spaces.

To compute the matrix representation of F in these bases, consider first

$$\begin{aligned} F(a_1) &= \int_0^x 1\mathrm{d}\tau = x \\ &= f_{11}25(x-1)^5 + f_{12}16(x-1)^4 + f_{13}9(x-1)^3 + f_{14}(x-1)^2 + f_{15}(x+1) + f_{16}(-1), \end{aligned} \tag{1.5}$$

examining which we can see that

$$(f_{11}, f_{12}, f_{13}, f_{14}, f_{15}, f_{16}) = (0, 0, 0, 0, 1, 1). \tag{1.6}$$

Similarly,

$$\begin{aligned} F(a_4) &= \int_0^x (\tau-1)^3\mathrm{d}\tau = \frac{1}{4}(x-1)^4 - \frac{1}{4} \\ &= f_{41}25(x-1)^5 + f_{42}16(x-1)^4 + f_{43}9(x-1)^3 + f_{44}(x-1)^2 + f_{45}(x+1) + f_{46}(-1), \end{aligned} \tag{1.7}$$

examining which we can see that

$$(f_{41}, f_{42}, f_{43}, f_{44}, f_{45}, f_{46}) = \left(0, \frac{1}{64}, 0, 0, 0, -\frac{1}{4}\right). \tag{1.8}$$

Therefore, completing the other cases as well, we find that F is represented by the matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{64} \\ 0 & 0 & \frac{1}{27} & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & \frac{1}{2} & -\frac{1}{3} & \frac{1}{4} \end{bmatrix} \tag{1.9}$$

On the other hand, you might prefer to use a different basis. Suppose instead that we use the "standard" basis for both $\mathcal{P}_3$ and $\mathcal{P}_5$. You can check that F is represented by the matrix:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{1.10}$$

These look totally different, yet they represent the same transformation!

1.3 Matrices to Linear Transformations

Suppose you have vector spaces V and W with bases (α_i) and (β_k), respectively. Then any matrix A specifies a unique linear transformation, F_A, defined by $F_A(\alpha_i) = A_{k,i}\beta_k$ and extending by linearity.

Proof. First, we need to make sure that F_A actually defines a function. Then, we need to make sure that the function is a linear transformation. Suppose that $v \in V$. Because (α_i) is a basis of V, there is a unique decomposition $v = \sum_i u_i\alpha_i$. Therefore, "extending by linearity" means that $F_A(v)$ is defined as

$$F_A(v) = \sum_i v_i F(\alpha_i) = \sum_{i,k} A_{k,i} v_i \beta_k, \tag{1.11}$$

which is a well-defined vector in W. Therefore, every $v \in V$ has a well-defined output value in W. Now to linearity: suppose that $r, s \in \mathbb{R}$ and $u, v \in V$ with $u = \sum_i u_i\alpha_i$ and $v = \sum_i v_i\alpha_i$. Ex 1.6: Now expand and use the previous formula to see that $F_A(ru+sv) = rF_A(u)+sF_A(v)$. □

1.4 Example: Matrices alone are meaningless

If you have two matrices that look the same, do they represent the same linear transformation? Consider your favorite matrix,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

You usually think of this as the "identity transformation on $\mathbb{R}^2$" but that interpretation relies on the context of $\mathbb{R}^2$ with a specific basis. Here are two other ways that the matrix A could be interpreted in other contexts:

- Let $V = \mathbb{R}^2$ with the standard basis. Suppose that $W = \{(x, y, z) \in \mathbb{R}^3 : 2x + 3y - z = 0\}$ with the basis $\beta_1 = (1, 0, -2)$ and $\beta_2 = (-3, 2, 0)$. In this basis the transformation F_A takes $\mathbb{R}^2$ to some sort of oblique, stretched plane in $\mathbb{R}^3$. This transformation cannot be the identity, since $V \neq W$.

- Here's an even crazier example: suppose that $V = \mathcal{P}_1$ and $W = \mathbb{C}$ (a.k.a., the complex plane) and $\alpha_1 = 1$ and $\alpha_0 = x$ and $\beta_1 = -i$ and $\beta_2 = \frac{3}{2}$. Then the corresponding transformation is
$$F_A(p) = \left(\frac{\partial p}{\partial x} - \int_0^x p \right) \Big|_{x=\sqrt{-1}}$$
So $F_A(a + bx) = b - ai - \frac{1}{2}bi^2 = \frac{3}{2}b - ai \in \mathbb{C}$.

1.5 Why do we dwell on Linear Algebra?

A matrix is meaningless until you know which spaces and which bases are in use. For most of your mathematical life, everyone was using the standard basis in $\mathbb{R}^n$, but they didn't have to! An amazing consequence is that everything you learned about $\mathbb{R}^n$ and systems of linear equations and matrices for $\mathbb{R}^n$ applies equally well to all these crazy spaces and their homomorphisms. **That is the point of linear algebra!** Some of the most important linear transformations and vector spaces appear as

1. calculus on smooth functions (including polynomials)
2. construction of solutions of differential equations
3. frequency analysis on wave-like functions
4. compression, error-detection, and encoding of digital data
5. interactions of elementary particles
6. gravitational energy and force in space-time

In at least one way, these are all the same family of problems, and this is why Linear Algebra is the gateway to all higher mathematics and a huge amount of applied science.

1.6 Isomorphisms as Basis-Changers

Lemma 1.3. *If $G : V \to W$ is an isomorphism, then it takes any basis of V to a corresponding basis of W.*

Proof. This is probably how you proved Exercise 1.1. □

That means that we can interpret isomorphisms as basis-changers.

Suppose that you are solving some interesting problem in $\mathbb{R}^3$ with the standard basis, (e_i). You have discovered that a particular vector is important for your work, say

$$v = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}_e = 4e_1 + e_2 + 2e_3 \tag{1.12}$$

On the other hand, your friend Igor hath a neck injury (and a lithp) and theeth everything thidewayth and thlightly thtretched. Igor preferth a bathith $(\beta_1, \beta_2, \beta_3)$. But, you would write these as

$$\beta_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}_e = 2e_1 + e_2, \text{ and } \beta_2 = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}_e = -e_1 + 3e_2, \text{ and } \beta_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}_e = e_3 \tag{1.13}$$

How would Igor write thith vector? Igor hath thome numberth u_1, u_2, u_3 where $v = u_1\beta_1 + u_2\beta_2 + u_3\beta_3$, so

$$\begin{aligned} 0 &= v - v \\ &= (4e_1 + e_2 + 2e_3) - (u_1\beta_1 + u_2\beta_2 + u_3\beta_3) \\ &= (4e_1 + e_2 + 2e_3) - (u_1(2e_1 + e_2) + u_2(-e_1 + 3e_2) + u_3e_3) \\ &= (4 - 2u_1 + u_2)e_1 + (1 - u_1 - 3u_2)e_2 + (2 - u_3)e_3. \end{aligned} \tag{1.14}$$

Therefore, you can find Igor'th coordinateth by solving the system

$$\begin{bmatrix} 2 & -1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}. \tag{1.15}$$

Ex 1.7: Tholve thith thythtem to thee thingth from Igor'th perthpective.

Implicitly here, we have written an isomorphism $G : V \to W{=}V$ that has the following properties:

1. $G(e_1) = \beta_1$ and $G(e_2) = \beta_2$ and $G(e_3) = \beta_3$;
2. Using the basis (e_i) for the domain and (β_i) for the range, G is represented by the matrix I, but remember the lesson of Example 1.4;
3. Using the basis (e_i) for both the domain and the range, G is represented by the non-singular matrix
$$P = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1.16}$$
4. To change from your coordinates to Igor'th coordinateth, apply P^{-1}.

Ex 1.8: Go back to Example 1.2 about integrals and find the basis transformations.

1.7 What can you and Igor agree on?

Once we choose a basis $(\alpha_1, \ldots, \alpha_n)$ for V, we know that $F : V \to V$ is represented by an $n \times n$ matrix, A. Question: What properties of F are "the same" no matter what basis you choose?

Lemma 1.4. *Suppose that $F : V \to V$ is represented by the matrix A when the basis (α_i) is used for both the domain and range. Similarly, suppose that F is represented by the matrix B when the basis (β_i) is used for both the domain and range. Suppose also that $G : V \to V$ is the isomorphism defined by $G(\alpha_1) = \beta_1$, $G(\alpha_2) = \beta_2$, $\ldots$, $G(\alpha_n) = \beta_n$. Suppose that, using the basis (α_i) for both the domain and the range, the transformation G is represented by the matrix P. Then*

$$B = P^{-1}AP \text{ and } A = PBP^{-1}. \tag{1.17}$$

In other words, if you say "F is represented by A", then Igor will say "F ith reprethented by B." So, if $w = F(v)$, then we have the following diagrams:

$$\begin{array}{ccc} V_\alpha & \xrightarrow{F} & W_\alpha \\ G \big\downarrow & & \big\downarrow G \\ V_\beta & \xrightarrow[F]{} & W_\beta \end{array}$$

$$\begin{array}{ccc} v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}_\alpha & \xrightarrow{A\cdot} & w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_\alpha \\ P^{-1}\cdot \big\downarrow & & \big\downarrow P^{-1}\cdot \\ v = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_\beta & \xrightarrow[B\cdot]{} & w = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}_\beta \end{array}$$

Suppose that there is a property "\$" of matrices such that $\$(A) = \$(P^{-1}AP) = \$(A)$ for any non-singular P. No matter how twisted Igor's perspective is, you and Igor will agree about these properties. A change of the form $A \mapsto P^{-1}AP$ is called **conjugation** or a **basis-change** of A.

So far as we can tell, physical laws do not have preferred coordinate systems.[1] Therefore, if mathematics is to be applied to science, we care primarily about those properties that are invariant under conjugation. The most significant of these properties are the final topics in this course.

[1] Actually, the most recent scans of the cosmic microwave background radiation suggest that there is giant asymmetry, with one area of the sky being significantly colder than the rest, so I guess you could call that "up." The meaning of this remains to be seen! Another counterexample is the 2nd law of thermodynamics, meaning that closed systems have a forward-time arrow. The only other counterexample I know of is that all biologically produced organic compounds are right-handed, which suggests that all current life on Earth originated at a single chemical event. If such an event occurred multiple times and if its consequences were still here, we'd expect to see a random mix of right- and left-handed compounds. When we make organic compounds in lab environments, they generally come out evenly distributed between right- and left-handed.

2 Determinants

What does a linear transformation $F : V \to V$ look like?

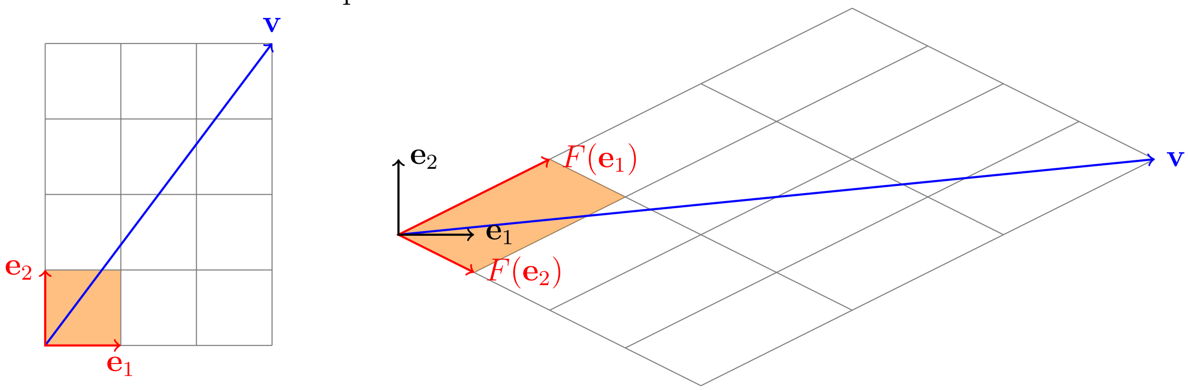
Even for $F : \mathbb{R}^2 \to \mathbb{R}^2$, cannot draw a complete graph in any reasonable way! However, if we choose basis vectors, we can see where they go. Consider $V = W = \mathbb{R}^2$ with the standard basis, and suppose F is the transformation represented by the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix} \tag{2.1}$$

We can see that $v = 3e_1 + 4e_2$ goes to

$$F(v) = 3F(e_1) + 4F(e_2) = 3\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 4\begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix} = 10e_1 + 1e_2 \tag{2.2}$$

We can see all of this in one picture:



Notice that all of the geometry here is described by the orange boxes. In the domain, the orange box is the rectangle given by the basis vectors. In the range, the orange box is the quadrilateral described as the image of the basis vectors. Certainly, its area has changed, but also the orientation has changed: you rotate anticlockwise from e_1 to e_2, but you rotate clockwise from $F(e_1)$ to $F(e_1)$

2.1 Motivating Idea

If we use the same basis for the domain and range, how do the volume and orientation of an n-prism change through a linear transformation $F : V \to V$?

To answer this question, let's choose a basis and represent F by an $n \times n$ matrix A. Then the columns of A are the images of the domain's basis, so we are considering the prism described by the columns of A, and we want to study this function, which we call "the determinant of A."

$|A|$ = "the volume of the image of the basis n-prism, signed by orientation."

2.2 Properties and Uniqueness

Here are some observations Ex 2.1: (which you should confirm)

1. If E is an elementary row operation that swaps two rows, then $|EA| = -|A|$.
2. If E is an elementary row operation that re-scales a single row by c, then $|EA| = c|A|$.
3. If E is an elementary row operation that adds a replaces ρ_i with $\rho_i + c\rho_j$, then $|EA| = |A|$.
4. $|I| = 1$.
5. If A has a row of 0s, then $|A| = 0$.

Lemma 2.1. *Let $\mathcal{M}_n$ denote the space of all $n \times n$ matrices. Suppose that* $\det : \mathcal{M}_n \to \mathbb{R}$ *is a function with the properties listed in part 2.2. Then* $\det(A) = |A|$ *for all $A \in \mathcal{M}_n$. In other words, these properties can be taken as the definition of a function called* $\det(A)$*, and in fact* $\det A$ *and* $|A|$ *are the same function.*[2]

Proof. Suppose that $\det$ and $\widehat{\det}$ are two functions that satisfy the properties above. Suppose that A is some matrix. Our goal is to show that $\det(A) = \widehat{\det}(A)$ for all A.

Let U denote the reduced row-echelon form of A. Recall that this is unique, and there is some sequence of elementary row operations $E_1, E_2, \ldots, E_r$ such that

$$E_r E_{r-1} \cdots E_2 E_2 A = U \text{ and } A = E_1^{-1} E_2^{-1} \cdots E_{r-1}^{-1} E_r^{-1} U. \tag{2.3}$$

(The *sequence* of row-operations is not unique, but this will not matter.) If A is degenerate, then U has a row of zeros and the properties mean $\det(U) = 0 = \widehat{\det}(U)$. If A is nondegenerate, then $U = I$ and the properties mean $\det(U) = 1 = \widehat{\det}(U)$. Now, working from the right, we apply the inverse elementary row-operations to U one-by-one. Each of these either changes the sign or scales or does nothing, but it does the same thing to both $\det$ and $\widehat{\det}$ until we finish. □

We will write either $\det A$ or $|A|$ interchangeably based purely on aesthetics.[3] While our definition and properties make sense geometrically, we still have a fairly nasty problem of *actually computing* $|A|$. The next three corollaries allow actual computation:

Corollary 2.2. *You can compute determinants by doing row-reduction on the fully augmented matrix $[A|I]$ and keeping track of the row operations. See the Example 2.3 below.*

Corollary 2.3. $|AB| = |A||B|$. *In particular,* $\det A \neq 0$ *if and only if A is nondegenerate, and* $|A^{-1}| = \frac{1}{|A|}$.

Ex 2.2: Prove this.

Corollary 2.4. *If $B = P^{-1}AP$, then* $\det B = \det A$.

Therefore, following section 1.7, we see that $\det$ is a property of $F : V \to V$ that is invariant under conjugation, so it does not depend on the representation. Hence, you and Igor agree on $\det F$.

[2] This is like the famous calculus argument "There is exactly one function, call it exp, with the properties $\exp' = \exp$ and $\exp(0) = 1$. Oh look, $\exp(x) = e^x$."

[3] ...much like you'd write $\exp(x)$ or e^x depending on what looks better on paper.

Corollary 2.5. $\det A = \det A^T$.

Proof. Ex 2.3: First, check this under the additional assumption that A is itself an elementary row operation, E.

For generic A, decompose it into row operations. This is the same as decomposing A^T into column operations. Since $(E_1E_2)^T = E_2^T E_1^T$ and $\det(E_1) = \det(E_1^T)$, and so on, and using the multiplication property above. ((detail)) □

2.3 Example: 2×2 determinants.

Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then we row-reduce the augmented matrix. While doing so, we will need to divide by some things that better not be zero, but we'll proceed and account for all of that at the end:

$$\begin{aligned}
&\begin{bmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{bmatrix} \\
&\xrightarrow{\frac{1}{a}\rho_1} \begin{bmatrix} 1 & b/a & 1/a & 0 \\ c & d & 0 & 1 \end{bmatrix} \\
&\xrightarrow{\rho_2 - c\rho_1} \begin{bmatrix} 1 & b/a & 1/a & 0 \\ 0 & (ad-bc)/a & -c/a & 1 \end{bmatrix} \\
&\xrightarrow{\frac{a}{ad-bc}\rho_2} \begin{bmatrix} 1 & b/a & 1/a & 0 \\ 0 & 1 & -c/(ad-bc) & a/(ad-bc) \end{bmatrix} \\
&\xrightarrow{\rho_1 - \frac{b}{a}\rho_2} \begin{bmatrix} 1 & 0 & 1/a + \frac{bc}{a(ad-bc)} & -b/(ad-bc) \\ 0 & 1 & -c/(ad-bc) & a/(ad-bc) \end{bmatrix} = \begin{bmatrix} 1 & 0 & d/(ad-bc) & -b/(ad-bc) \\ 0 & 1 & -c/(ad-bc) & a/(ad-bc) \end{bmatrix}
\end{aligned} \tag{2.4}$$

There were four row operations. Working backwards from $\det I = 1$, we get

$$\det A = a\left(1\left(\frac{ad-bc}{a}\left(1\left(\det I\right)\right)\right)\right) = ad - bc \tag{2.5}$$

Moreover, we now know that $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

In order to do this computation, we needed to divide by a and $ad - bc$. If $a = 0$ and $c \neq 0$, then we could have swapped rows and multiply $\det(A)$ by (-1). If $a = c = 0$, then A was degenerate and we know $\det A = 0$ anyway. If $ad - bc = 0$, then this reduction does not work; however, reexamining the context of the problem, this just says that "a degenerate matrix cannot be reduced to the identity," which we already knew!

In short, without using any trigonometry, we now know that the quadrilateral with sides (a, c) and (b, d) has (signed) area $ad - bc$.

Corollary 2.6. *Suppose that A is an $n \times n$ matrix of the form*

$$A = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & (B) & \\ 0 & & & \end{bmatrix} \tag{2.6}$$

where B is an $(n-1)\times(n-1)$ matrix. Then $\det A = \det B$.

Proof. This is a n-prism with "$(n-1)$ base" $\det B$ and height 1. □

Corollary 2.7. *Using linear combinations, we can compute the* $\det A$ *in terms of its* $(n-1)\times(n-1)$ *submatrices, and each of those in terms of its* $(n-2)\times(n-2)$ *submatrices, and so on. Eventually you get to* 2×2 *matrices, which we understand. This is called "cofactor expansion."*

The formula is $\det A = \sum_j (-1)^{1+j} A_{1,j} m_{i,j}$ *where* $m_{i,j} = \det(M_{i,j})$ *is the determinant of the* $(n-1)\times(n-1)$ *sub-matrix opposite* $A_{1,j}$*. Either* $m_{i,j}$ *or* $M_{i,j}$ *is sometimes called a "minor."*

Some people (with no taste) take this as the definition. From our perspective, the proof is simply that this formula satisfies the properties of the determinant, and the determinant is the only function with those properties.

Ex 2.4: For 3×3 matrices, verify that this definition satisfies all of the properties of a determinant, so it must be the determinant.

Instead of using the first row, you can use a similar expansion for any row or any column.

2.4 Example: Computing Higher Determinants

Consider the matrix $A = \begin{bmatrix} 3 & 4 & 6 \\ 2 & 1 & 2 \\ 0 & -1 & 5 \end{bmatrix}$ and write the co-factor formula for the top row, $(3,4,5)$. The submatrix opposite $A_{1,1} = 3$ is $M_{1,1} = \begin{bmatrix} 1 & 2 \\ -1 & 5 \end{bmatrix}$, so $m_{1,1} = 7$. The submatrix opposite $A_{1,2} = 4$ is $M_{1,2} = \begin{bmatrix} 2 & 2 \\ 0 & 5 \end{bmatrix}$, so $m_{1,2} = 10$. The submatrix opposite $A_{1,3} = 6$ is $M_{1,3} = \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}$, so $m_{1,3} = -2$. Therefore the determinant is

$$\det A = 3(7) - 4(10) + 6(-2) = 21 - 40 - 12 = -31.$$

Alternatively, you could do this using row operations:

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 6 \\ 2 & 1 & 2 \\ 0 & -1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 4/3 & 2 \\ 0 & 5/6 & -1 \\ 0 & -1 & 5 \end{bmatrix} \tag{2.7}$$

This is of the form $E_3E_2E_1A = B$, so $\det B = \det E_3 \det E_2 \det E_1 \det A = 1\frac{1}{2}\frac{1}{3}\det A$. Then, apply column operations to clear out the first row:

$$\begin{bmatrix} 1 & 4/3 & 2 \\ 0 & 5/6 & -1 \\ 0 & -1 & 5 \end{bmatrix} \begin{bmatrix} 1 & -4/3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5/6 & -2 \\ 0 & -1 & 5 \end{bmatrix} \tag{2.8}$$

This is of the form $C = BF_1F_2$, so $\det C = \det B \det F_1 \det F_2 = \det B$. Now, C is of the form in the corollary, so $\det C = -\frac{31}{6}$, and $\det A = -31$.

Notice that neither F_1 nor F_2 required any scaling, since that was already handled by E_2 above. Therefore, the column operations are not actually necessary, as expected from the co-factor formula.

Corollary 2.8. *The following matrices have an easy-to-compute determinant: diagonal matrices; upper-triangular matrices, lower-triangular matrices; matrices one of whose rows or columns is mostly zeroes.*

Ex 2.5: Write down a (nonsingular) 4×4 matrix of each kind, and compute its determinant.

Ex 2.6: Compute several random 3×3 and 4×4 determinants. Do each of them three times: once by hand using Gaussian elimination, once by hand using the co-factor formula, and once using a computer algebra system. Most people prefer the latter method.

3 Eigenvalues and Eigenvectors

What are the most useful numbers? Probably 0 and 1. What are the most useful functions? Probably $f(x) = x$ and $f(x) = e^x$. Each of these is an example of a *fixed point* for some operation. For example, we like e^x primarily because $\frac{\partial}{\partial x}e^x = e^x$. It is *fixed* or *preserved* by the derivative transformation. That also means that $\frac{\partial}{\partial x}f(x) = \lambda f(x)$ for any $f(x) = e^{\lambda x}$ with constant λ, which is nearly as nice.

If we want to understand a certain transformation, a lot can be learned by studying the objects it preserves, $F(v) = v$. There are a few caveats to keep in mind: We can only say "$F(v) = v$" if the domain and range of F are the same space, so $F : V \to W{=}V$. We already know that $F(0) = 0$ for any linear transformation, so we don't care about $v = 0$. Also, just like the example $f(x) \mapsto \lambda f(x)$ above, we want to allow pure re-scaling of vectors, so we'll examine $F(v) = \lambda v$. This means that we really care about "directions that are preserved" not "vectors that are preserved."

Definition 3.1. *Suppose that $F : V \to V$. Suppose that λ is a number and v is a non-zero vector such that $F(v) = \lambda v$. We call v an* ***eigenvector***[4] *The set of all eigenvalues is called the* ***spectrum*** *of F.*

This definition does *not* rely on bases and matrices. It is purely a property of the linear transformation F. We will explore this soon.

It turns out that it is easier to first find λ, then find v afterwards.

3.1 Finding Eigenvalues

Suppose that $F : V \to V$ is represented by A when using a basis (α_i) for both the domain and range. Now, suppose that $Av = \lambda v$ has a non-zero solution, v. Then $0 = Av - \lambda v = Av - \lambda Iv = (A - \lambda I)v$, so the matrix $(A - \lambda I)$ must be degenerate. Therefore, $\det(A - \lambda I) = 0$. This is very useful because we can treat λ as a variable, and solving $\det(A - \lambda I) = 0$ for λ is "just" the classic problem of finding roots of polynomials.

Once you know an eigenvalue, λ, you can find the corresponding eigenvector by fixing that λ and solving the homogeneous system $(A - \lambda I)v = 0$ in the normal way.

Lemma 3.2. *If $\lambda = 0$ is an eigenvalue of F, then F is not an isomorphism.*

Proof. Read the previous paragraph, setting $\lambda = 0$. □

Here is a complete example. Suppose that (in some basis) F is given by the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}, \text{ so } (A - \lambda I) = \begin{bmatrix} 1-\lambda & 1 & 1 \\ 2 & 2-\lambda & 2 \\ 3 & 3 & 3-\lambda \end{bmatrix}. \tag{3.1}$$

So, we find the determinant:

$$\det(A-\lambda I) = (1-\lambda)(2-\lambda)(3-\lambda)+6+6-3(2-\lambda)-6(1-\lambda)-2(3-\lambda) = -\lambda^3+6\lambda^2 = \lambda^2(6-\lambda)$$

[4] Eigen means "own" in German, as in "This is my own favorite sandwich." The transformation F might say "this is my own favorite vector."

Therefore, the solutions are the eigenvalues $\lambda = 6$, $\lambda = 0$, and $\lambda = 0$. We might say that the spectrum of F is $\{0, 0, 6\}$. Now, let us find the corresponding eigenvectors.

- Consider $\lambda = 6$, so we need to solve $(A - 6I)v = 0$.

$$\begin{bmatrix} -5 & 1 & 1 \\ 2 & -4 & 2 \\ 3 & 3 & -3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1/3 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.2}$$

The null-space is the 1-dimensional space:

$$\left\{ \begin{bmatrix} 1/3 \\ 2/3 \\ 1 \end{bmatrix} t : \ t \in \mathbb{R} \right\} \tag{3.3}$$

- Suppose that $\lambda = 0$, so we need to solve $(A - 0I)v = Av = 0$.

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.4}$$

The null-space is the two-dimensional space

$$\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} t + \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} s : \ s, t \in \mathbb{R} \right\} \tag{3.5}$$

Note! There is a pile of additional examples in Sections 3.4 and 3.5.

3.2 Existence and Uniqueness of Eigenvalues

Definition 3.3. *Given a matrix A, the polynomial $p(x) = \det(A - xI)$ is called the **characteristic polynomial** of A. Its roots are the eigenvalues of A.*

Lemma 3.4. *Any $n \times n$ matrix with real or complex entries must have n complex eigenvalues (counted with multiplicity). If n is odd, then any $n \times n$ matrix A with real entries must have at least one real eigenvalue. If n is even, then an $n \times n$ matrix A may or may not have any real eigenvalues.*

Proof. Eigenvalues are found by finding the roots of the characteristic polynomial, whose coefficients are products and sums of the coefficients of A. Examination of the definition of the determinant shows that this polynomial has degree n. So, this is a direct application of the Fundamental Theorem of Algebra. □

Theorem 3.5. *If F is an isomorphism with an eigenvalue λ, then $\frac{1}{\lambda}$ is an eigenvalue of F^{-1}.*

Ex 3.1: Prove this. Hint: notice that an eigenvector of F is also an eigenvector of F^{-1}.

Theorem 3.6. *If $B = P^{-1}AP$, then A and B share the same eigenvalues (but the eigenvectors may look completely different, due to the change-of-basis P).*

Ex 3.2: Prove this. Hint: use the characteristic polynomial as in the example above.

This means that eigenvalues are a property of a formal linear transformation F, and they do not depend on whatever particular basis you may use to write F as a matrix. Hence, you and Igor will agree on eigenvalues, though you will write the corresponding eigenvector differently.

Ex 3.3: Look back at matrix A in Equation 3.1, rewrite A for Igor using the change-of-basis P from Equation 1.16. Verify that Igor'th eigenvalueth are identical to your eigenvalues.

3.3 Eigenspaces

Now that we understand how to understand eigenvalues as the roots of the characteristic polynomial, we need to turn our attention to the corresponding eigenvectors.

Definition 3.7. *Suppose that λ is an eigenvalue of F. Let $K_\lambda = K_\lambda = \{v \in V : F(v) = \lambda v\}$. This is called the* ***eigenspace*** *associated to the eigenvalue λ. It is the set of vectors with that particular eigenvalue.*

Lemma 3.8. *For any λ, the eigenspace K_λ is a vector subspace of V. Moreover, if $\lambda \neq \mu$, then $K_\lambda \cap K_\mu = 0$.*

Ex 3.4: Prove this.

Corollary 3.9. *Suppose that λ is a* non-zero *eigenvalue of F. Show that $F(K_\lambda) = K_\lambda$. In other words, show that $F|_{K_\lambda}$ is an isomorphism, and if K_λ is dimension r, then K_λ is represented in any basis by the $r \times r$ matrix*

$$\lambda I_r = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix}.$$

Ex 3.5: Prove this. Hint: The subspace K_λ has a basis. Where does the basis go?

So, V has a collection of subspaces on which the transformation F is "nice." It would be awfully convenient if we could decompose the *entire* space V into spaces on which F is this nice. This is not always possible, but it is *nearly* possible. Understanding this is the topic of the Jordan Canonical Form of a matrix in Linear Algebra II. First, let's see an example.

Ex 3.6: Consider the matrix

$$A = \begin{bmatrix} 5 & 6 & 4 \\ 4 & 7 & 4 \\ -11 & -13 & -9 \end{bmatrix}.$$

Show that A has three distinct eigenvalues, namely $-2, 1, 3$, and show that A is similar to the matrix:

$$B = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

In this exercise, $\mathbb{R}^3$ decomposes into three subspaces, K_{-1}, K_2, and K_3, each of which is a one-dimensional subspace. By applying a change-of-basis $B = P^{-1}AP$, we can change these subspaces to become the three standard lines $\{te_1 : t \in \mathbb{R}\}$, $\{te_2 : t \in \mathbb{R}\}$, and $\{te_3 : t \in \mathbb{R}\}$.

On the other hand, you could have a 3×3 matrix with exactly one real eigenvalue, as in the next exercise.

Ex 3.7: Suppose that A is a 3×3 matrix of the following form where $a, b, c, d \neq 0$.

$$A = \begin{bmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{bmatrix}.$$

1. Prove that A is non-singular.
2. Consider the matrix $B = A - aI_3$. Prove that the matrix B has rank 2 and nullity 1.
3. Prove that the set $E = \{\vec{v} \in \mathbb{R}^3 : A\vec{v} = a\vec{v}\}$ is a 1-dimensional subspace of $\mathbb{R}^3$, and find a basis.
4. Prove that the matrix B^2 has rank 1 and nullity 2.
5. Prove that the set $E' = \{\vec{v} \in \mathbb{R}^3 : A^2\vec{v} - 2aA\vec{v} + a^2\vec{v} = \vec{0}\}$ is a 2-dimensional subspace of $\mathbb{R}^3$ that contains E, and find a basis.
6. Prove that the matrix B^3 has rank 0 and nullity 3.
7. Prove that the set $E'' = \{\vec{v} \in \mathbb{R}^3 : A^3\vec{v} - 3aA^2\vec{v} + 3a^2A\vec{v} - a^3\vec{v} = \vec{0}\}$ is all of $\mathbb{R}^3$.

Notice that we have a sequence of subspaces of dimensions 0,1,2,3

$$0 \subset E \subset E' \subset E''$$

described as the null-spaces of the matrices A, $(A - aI)$, $(A - aI)^2$ and $(A - aI)^3$. The second of these, E, is really the eigenspace K_a. The others, E' and E'', are not really eigenspaces, but you can see that they are related somehow. This phenomenon is the core of the most important theorem in Linear Algebra, the "Jordan Canonical Form" which will be a major focus in Linear Algebra II and is the key to solving differential equations.

Theorem 3.10. *Suppose V is an n-dimensional vector space and that $F : V \to V$ has n distinct eigenvalues. Then there is a basis in which F is represented by a diagonal matrix. Moreover,* $\det F$ *is the product of the eigenvalues.*

We see in Exercise 3.3 that the hypothesis of n distinct eigenvalues is not always true. There are stronger, more nuanced conditions that also imply diagonalizability, but they are a topic for another day.

Proof. Each eigenvalue is associated to an eigenvector, and these eigenvectors live in 1-dimensional subspaces, $K_{\lambda_1}, K_{\lambda_2}, \ldots, K_{\lambda_n}$ which intersect trivially. Therefore, we can choose one eigenvector for each eigenvalue and use those eigenvector to write F in a basis. By applying Lemma 3.9 to each eigenspace, this implies that F is represented by a diagonal matrix in that basis. □

3.4 Eigenspace decomposition in two dimensions

In this section, we see several examples of how a 2×2 matrix can possibly look when the space V is decomposed using eigenvectors.

In general, the characteristic polynomial of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $(a-x)(d-x) - bc = x^2 - (a+d)x + (ad - bc)$. This is always a quadratic function, so over the reals is may have two, one, or zero solutions. In other words, it may factor as $(x - \lambda_1)(x - \lambda_2)$, or it may factor as $(x - \lambda_1)^2$, or it may not factor at all over $\mathbb{R}$. The vertex is at $\frac{1}{2}(a + d)$.

3.4.1 Two real eigenvalues.

Consider the matrix

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}. \tag{3.6}$$

The characteristic polynomial is $(2-x)(4-x) - 3 = x^2 - 6x + 5 = (x-1)(x-5)$, so A has two eigenvalues, $\lambda_1 = 1$ and $\lambda_2 = 5$. Now we need to find the eigenspaces:

- $K_1 = \{v : (A - \lambda_1 I)v = 0\}$, so we need to solve

$$\begin{bmatrix} 2-1 & 1 \\ 3 & 4-1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \text{ so} \tag{3.7}$$

$$K_1 = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} t, \ t \in \mathbb{R} \right\} \subset \mathbb{R}^2.$$

- $K_2 = \{v : (A - \lambda_2 I)v = 0\}$, so we need to solve

$$\begin{bmatrix} 2-5 & 1 \\ 3 & 4-5 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \text{ so} \tag{3.8}$$

$$K_2 = \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} t, \ t \in \mathbb{R} \right\} \subset \mathbb{R}^2.$$

Now, notice that K_1 and K_2 together span all of $\mathbb{R}^2$. In other words, we can apply the change-of-basis $P = \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix}$ to rewrite A as

$$P^{-1}AP = \frac{1}{4} \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}. \tag{3.9}$$

3.4.2 One real eigenvalue.

There are two different types of 2×2 matrices that have exactly one eigenvalue. Here are two simple examples that show the difference. Let

$$B = \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}, \ C = \begin{bmatrix} 7 & 1 \\ 0 & 7 \end{bmatrix}. \tag{3.10}$$

Notice that B and C have the same characteristic polynomial $p(x) = (7-x)^2$, whose only solution is 7. To find the corresponding eigenspaces, we examine the kernels of the matrices $(B - 7I)$ and $(C - 7I)$.

Of course, $(B - 7I)$ is the zero matrix, so every vector is an eigenvector of B, and we can take e_1 and e_2 to be a basis of the eigenspace.

On the other hand, $C - 7I$ is a matrix of rank 1. Its kernel is spanned by e_1. The eigenspace is only one-dimensional.

Therefore, B and C cannot be similar, because they cannot represent the same linear transformation. Ex 3.8: Explain why they cannot be similar. Hint: what would an isomorphism P do to the eigenspaces?

This problem gets worse in higher dimensions, and it is why the Jordan Canonical Form theorem, which extends Theorem 3.10, is rather subtle.

3.4.3 Two complex eigenvalues.

Here is an example of a real matrix with no real eigenvalues:

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{3.11}$$

So,

$$0 = \det(A - \lambda I) = \det \begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 + 1. \tag{3.12}$$

The only possible solutions are $\pm\sqrt{-1}$. So, A has two complex eigenvalues but no real eigenvalues.

What does this matrix do as a transformation F_A? Well, $F_A(e_1) = e_2$, and $F_A(e_2) = -e_1$, so this is an anti-clockwise rotation of $\frac{\pi}{2}$ in the plane. This is a general phenomenon: **pairs of complex-conjugate eigenvalues correspond to rotations**.

Ex 3.9: Find the complex eigenvalues of the matrix $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ for an angle θ.

3.5 Eigenspace decomposition in three dimensions

Here are several different "nice" 3×3 matrices. Each must have at least one real eigenvalue. Ex 3.10: Find the eigenvalues and the associated eigenspaces.

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}, \ \begin{bmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}, \ \begin{bmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{bmatrix}, \ \begin{bmatrix} a & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

4 Main Theorem

Let $F : V \to W$ be a linear transformation where $\dim V = \dim W = n$. The following are equivalent:

1. F is an isomorphism.
2. $\ker F = 0$.
3. $\operatorname{im} F = W$.
4. All eigenvalues of F are non-zero.
5. There is an inverse transformation $F^{-1} : W \to V$.

Moreover, if A is an $n \times n$ matrix that represents F in some bases, these are also equivalent to those above:

6. The system of linear equations described by A is non-degenerate.
7. The nullity of A is 0.
8. The row-rank and column-rank of A are n.
9. All eigenvalues of A are non-zero.
10. The determinant of A is non-zero.
11. A can be row-reduced to the $n \times n$ identity matrix, $RA = I$.
12. A can be obtained by a product of elementary row operations.

Ex 4.1: Piece together everything you know to prove this.