MATH 2152 Notes

Kevin L

Winter Term 2 2022

Contents

1 Review of 1152

Theorem 1.1. The Non-Singularity Theorem:
Let A be a $n \times n$ matrix with entries in F, then the following are equivalent:

- A is non-singular
- $N(A) = \{0\}$
- A is invertible
- $\operatorname{rank} A = n$
- RREF $A = I_n$
- nullity $A = 0$
- rows and cols are a basis of F^n (linear independent & span)
- $A\mathbf{x} = 0 \implies \mathbf{x} = 0$
- $\exists!$ solution to $A\mathbf{x} = \mathbf{b}$ for all $\mathbf{b}$

2 Linear Maps

2.1 Matrix Mapping

Theorem 2.1. Let $T : F^n \to F^m$ be linear. Then $\exists!$ A $m \times n$ matrix such that $T(\mathbf{v}) = A\mathbf{v}$ for all $\mathbf{v} \in F$. Thus,

$$A = \Big(T(e_1), \ldots, T(e_n)\Big)$$

i.e. the columns of A are $T(e_i)$.

Example. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $T(x, y, z) = (x + y + z, \frac{1}{2}x)$.
Then,

$$\begin{aligned} A &= \Big(T(e_1) \quad T(e_2) \quad T(e_3)\Big) \\ &= \Big(T(1,0,0) \quad T(0,1,0) \quad T(0,0,1)\Big) \\ &= \begin{pmatrix} 1 & 1 & 1 \\ \frac{1}{2} & 0 & 0 \end{pmatrix} \end{aligned}$$

Theorem 2.2. Let V, W be vector spaces over F with $\dim W = m$ and $\dim V = n$. Then $\exists!$ an $m \times n$ matrix A. Let $T : V \to W$ be a linear transformation. Given basis' $\mathcal{B}_V$ and $\mathcal{B}_W$, we have

$$[T(\mathbf{v})]_{\mathcal{B}_W} = A[\mathbf{v}]_{\mathcal{B}_V}$$

for all $\mathbf{v} \in V$.
The formula to compute A is given by:

$$A = \Big([T(v_1)]_{\mathcal{B}_W} \quad \ldots \quad [T(v_n)]_{\mathcal{B}_W}\Big)$$

Example. Let $T : P_3(\mathbb{R}) \to P_2(\mathbb{R})$ be defined by $T(p(x)) = p'(x)$. Let $\mathcal{B}_{P_3(\mathbb{R})} = \{x^3, x^2, x, 1\}$ and $\mathcal{B}_{P_2(\mathbb{R})} = \{x^2, x, 1\}$.

Then,

$$\begin{aligned}
A &= \begin{pmatrix} [T(x^3)]_{\mathcal{B}_{P_2(\mathbb{R})}} & [T(x^2)]_{\mathcal{B}_{P_2(\mathbb{R})}} & [T(x)]_{\mathcal{B}_{P_2(\mathbb{R})}} & [T(1)]_{\mathcal{B}_{P_2(\mathbb{R})}} \end{pmatrix} \\
&= \begin{pmatrix} [3x^2]_{\mathcal{B}_{P_2(\mathbb{R})}} & [2x]_{\mathcal{B}_{P_2(\mathbb{R})}} & [1]_{\mathcal{B}_{P_2(\mathbb{R})}} & [0]_{\mathcal{B}_{P_2(\mathbb{R})}} \end{pmatrix} \\
&= \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}
\end{aligned}$$

Let $V = ax^3 + bx^2 + cx + d$, then

$$[T(\mathbf{v})]_{\mathcal{B}_W} = [3ax^2 + 2bx + c]_{\mathcal{B}_W} = \begin{pmatrix} 3a \\ 2b \\ c \end{pmatrix}$$

and

$$A[\mathbf{v}]_{\mathcal{B}_V} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 3a \\ 2b \\ c \end{pmatrix}$$

Proof. We must check that $[T(\mathbf{v}) = A[\mathbf{v}]_{\mathcal{B}_V}$. Let $\mathcal{B}_V = \{v_1, \dots, v_n\}$ and $\mathbf{v} = c_1v_1 + \cdots + c_nv_n$ Then on the LHS:

$$\begin{aligned}
[T(v)]_{\mathcal{B}_V} &= [T(c_1v_1 + \cdots + c_nv_n)]_{\mathcal{B}_W} \\
&= [c_1T(v_1) + \cdots + c_nT(v_n)]_{\mathcal{B}_W} \\
&= [c_1T(v_1)]_{\mathcal{B}_W} + \cdots + [c_nT(v_n)]_{\mathcal{B}_W}
\end{aligned}$$

On the RHS:

$$\begin{pmatrix} T(v_1)]_{\mathcal{B}_W} & \dots & [T(v_n)]_{\mathcal{B}_W} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = LHS$$

□

Theorem 2.3. Let V, W be finite dimensional vector spaces over the same field F, and let $S : V \to W$ and $T : U \to V$ be linear maps. Fix bases $\mathcal{B}_V$ of V and $\mathcal{B}_W$ of W, and let A_S and A_T be the matrices for S and T with respect to those bases. Let $\alpha \in F$.
Then $S + T = A_S + A_T$ and $\alpha T = \alpha A_T$.

Theorem 2.4. Let U, V, W be finite dimensional vector spaces of dimension p, n, and m respectively. Let $\mathcal{B}_U, \mathcal{B}_V, \mathcal{B}_W$ be bases of U, V, W. Let $S : V \to W$ and $T : U \to V$ be linear maps with corresponding matrices $A_S \in M_{mn}(F)$ and $A_T \in M_{np}(F)$ with respect to their corresponding bases. Then the linear map $S \circ T : U \to W$ has matrix A_SA_T with respect to $\mathcal{B}_U$ and $\mathcal{B}_W$.

Proof. Let $\mathbf{u} \in U$ and $\mathbf{v} \in V$. Then by the matrix mapping theorem, we have

$$[S\mathbf{v}]_{\mathcal{B}_W} = A_S[\mathbf{v}]_{\mathcal{B}_V}$$

and

$$[T\mathbf{u}]_{\mathcal{B}_V} = A_T[\mathbf{u}]_{\mathcal{B}_U}$$

Thus,

$$\begin{aligned}[(ST)(\mathbf{u})]_{\mathcal{B}_W} &= [S(T(\mathbf{u}))]_{\mathcal{B}_W} \\ &= A_S[T(\mathbf{u})]_{\mathcal{B}_V} \\ &= A_S A_T[\mathbf{u}]_{\mathcal{B}_U}\end{aligned}$$

□

Example. Define a linear map $T : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$T(x, y) = (-8x + 5y, -10x + 7y)$$

First, we find a matrix A_1 such that

$$T\begin{bmatrix} x \\ y \end{bmatrix} = A_1 \begin{bmatrix} x \\ y \end{bmatrix}$$

Then, let

$$\mathcal{B}_2 = \{(1, 1), (1, 2)\}$$

and find A_2, the matrix for T with respect to $\mathcal{B}_2$. Can either A_1 or A_2 be used to find a general formula for the individual components of $T^m(x, y)$.

To find A_1, we use the usual trick or the special case of the Matrix Mapping Theorem. Since $T(1, 0) = (-8, -10)$ and $T(0, 1) = (5, 7)$, the latter yields

$$A_1 = \begin{bmatrix} -8 & 5 \\ -10 & 7 \end{bmatrix}$$

To find A_2, we use the Theorem. We have $T(1, 1) = (-3, -3)$. We write this as a linear combination of the vectors in $\mathcal{B}_2 = \{(1, 1), (1, 2)\}$. Since $T(1, 1) = -3(1, 1) + 0(1, 2)$, we have $[T(1, 1)]_{\mathcal{B}_2} = (-3, 0)$. This is the first column of A_2. Similarly, $T(1, 2) = (2, 4) = 0(1, 1) + 2(1, 2)$, and so $[T(1, 2)]_{\mathcal{B}_2}$. We have

$$A_2 = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$$

Now we find a general formula for the individual components of $T^m(\mathbf{v})$. Since $T(\mathbf{v}) = A_1\mathbf{v}$, we have $T^m\mathbf{v} = A_1^m\mathbf{v}$ for

$$A_1 = \begin{bmatrix} -8 & 5 \\ -10 & 7 \end{bmatrix}$$

This is very tricky.

Since $[T(\mathbf{v})] = A_2[\mathbf{v}]_{\mathcal{B}_2}$ for

$$A_2 = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$$

. since A_2 is a diagonal matrix, it is easy to find the mth power.

$$A_2^m = \begin{bmatrix} (-3)^m & 0 \\ 0 & 2^m \end{bmatrix}$$

Now we find $[(x, y)]_{\mathcal{B}_2}$. We write (x, y) as a linear combination of the vectors of $\mathcal{B}_2$.

$$\begin{bmatrix} 1 & 1 & x \\ 1 & 2 & y \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2x - y \\ 0 & 1 & y - x \end{bmatrix}$$

Thus

$$(x, y) = (2x - y)(1, 1) + (y - x)(1, 2)$$

and so

$$[(x, y)]_{\mathcal{B}_2} = (2x - y, y - x)$$

Since $[T^m(\mathbf{v})]_{\mathcal{B}_2} = A_2^m[\mathbf{v}]_{\mathcal{B}_2}$, we have

$$\begin{aligned}\left[T^m \begin{bmatrix} x \\ y \end{bmatrix}\right]_{\mathcal{B}_2} &= A_2^m \left[\begin{bmatrix} x \\ y \end{bmatrix}\right]_{\mathcal{B}_2} \\ &= \begin{bmatrix} (-3)^m & 0 \\ 0 & 2^m \end{bmatrix} \begin{bmatrix} 2x - y \\ y - x \end{bmatrix} \\ &= \begin{bmatrix} (-3)^m(2x-y) \\ 2^m(y-x) \end{bmatrix}\end{aligned}$$

This is the coordinate vector for $T^m(x, y)$ with respect to $\mathcal{B}_2$. To find $T^m(x, y)$, we use the entries of the coordinate vector as coefficients in a linear combination of the elements of $\mathcal{B}_2$. Thus

$$\begin{aligned}T^m \begin{bmatrix} x \\ y \end{bmatrix} &= (-3)^m(2x-y) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2^m(y-x) \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} (-3)^m(2x-y) + 2^m(y-x) \\ (-3)^m(2x-y) + 2^{m+1}(y-x) \end{bmatrix}\end{aligned}$$

To compute $T^{100}(5, 2)$ for example, we use the formula to find

$$T^{100}(5, 2) = (8(-3)^{100} + 3 \cdot 2^{100}, 8(-3)^{100} + 3 \cdot 2^{101})$$

2.2 Null Space and Range

Definition 2.5. Let $T : V \to W$ be a linear map. The **null space** of T is

$$N(T) = \{v \in V \mid T(v) = 0\}$$

In other words, it is all vectors in v that map to the 0 vector in W. It is a subspace of V.

Definition 2.6. Let $T : V \to W$ be a linear map. The **range** of T is

$$\text{range}(T) = \{T(v) \mid v \in V\}$$

In other words, it is all vectors in W that get mapped to by all vectors in V. It is a subspace of W.

Example. Define $T : \mathbb{R}^3 \to \mathbb{R}^2$ by

$$T(x, y, z) = (2x - 3y + z, 5x + y - 2z)$$

Given that this is a linear map, find bases for $N(T)$ and range T.

The null space contains the vectors in $\mathbb{R}^3$ that T sends to 0. We have $T(x, y, z) = (0, 0)$ iff (x, y, z) is a solution to the system

$$\begin{aligned}2x - 3y + z &= 0 \\ 5x + y - 2z &= 0\end{aligned}$$

We solve by row reducing:

$$\begin{bmatrix} 2 & -3 & 1 \\ 5 & 1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & \frac{-5}{17} \\ 0 & 1 & \frac{-9}{17} \end{bmatrix}$$

Since z is free, set $z = t$. Then

$$\begin{aligned} N(T) &= \left\{ t \begin{bmatrix} \frac{5}{17} \\ \frac{9}{17} \\ 1 \end{bmatrix} \mid t \in \mathbb{R} \right\} \\ &= \operatorname{span} \left\{ \begin{bmatrix} \frac{5}{17} \\ \frac{9}{17} \\ 1 \end{bmatrix} \right\} \\ &= \operatorname{span} \left\{ \begin{bmatrix} 5 \\ 9 \\ 17 \end{bmatrix} \right\} \end{aligned}$$

Thus a basis for $N(T)$ is the above. We can check our work by verifying that $T(5, 9, 17) = 0$.
Let's find a basis for the range of T. We have

$$\begin{aligned} \operatorname{range} T &= \left\{ \begin{bmatrix} 2x - 3y + z \\ 5x + y - 2z \end{bmatrix} \mid x, y, z \in \mathbb{R} \right\} \\ &= \left\{ x \begin{bmatrix} 2 \\ 5 \end{bmatrix} + y \begin{bmatrix} -3 \\ 1 \end{bmatrix} + z \begin{bmatrix} 1 \\ -2 \end{bmatrix} \mid x, y, z \in \mathbb{R} \right\} \\ &= \operatorname{span} \left\{ \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\} \end{aligned}$$

Note that the range of T is equal to the column space of the matrix. We now row reduce to find

$$\begin{bmatrix} 1 & 0 & -\frac{5}{17} \\ 0 & 1 & -\frac{9}{17} \end{bmatrix}$$

that the last column is linearly dependent and so a basis of the range/col space of the matrix is $\{(2, 5), (-3, 1)\}$.

Example. Let $D : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ be the differentiation map, which is linear. Find bases for $N(D)$ and $\operatorname{range} D$.
If $p(x) = ax^2 + bx + c$, then $D(p(x)) = p'(x) = 2ax + b$. The null space of D consists of all polynomials that D sends to 0. This is only $a = b = 0$ with c free. Thus

$$\begin{aligned} N(D) &= \{c \mid c \in \mathbb{R}\} \\ &= \operatorname{span}\{1\} \end{aligned}$$

The null space therefore consists of all constant polynominals.
We have

$$\begin{aligned} \operatorname{range}(D) &= \{2ax + b \mid a, b \in \mathbb{R}\} \\ &= \operatorname{span}\{2x, 1\} \\ &= \operatorname{span}\{x, 1\} \end{aligned}$$

Thus a basis of $\operatorname{range} T$ is $\{x, 1\}$.

2.3 Fundamental Theorem of Linear Maps

Theorem 2.7. Let V, W be finite dimensional vector spaces over F. Let $\mathcal{B}_V, \mathcal{B}_W$ be bases. Let $T : V \to W$ be a linear map. Let A be the matrix for T, as given by the matrix mapping theorem. Then,

$$v \in N(T) \Leftrightarrow [v]_{\mathcal{B}_V} \in N(A)$$

and

$$w \in \operatorname{range} T \Leftrightarrow [w]_{\mathcal{B}_W} \in \operatorname{colspace} A$$

Proof. By the matrix mapping theorem, we have

$$[T(v)]_{\mathcal{B}_W} = A[v]_{\mathcal{B}_V}$$

Thus

$$\begin{aligned} v \in N(T) &\Leftrightarrow T(v) = 0 \\ &\Leftrightarrow [T(v)]_{\mathcal{B}_W} = 0 \\ &\Leftrightarrow A[v]_{\mathcal{B}_V} = 0 \\ &\Leftrightarrow [v]_{\mathcal{B}_V} \in N(A) \end{aligned}$$

We also have

$$\begin{aligned} w \in \operatorname{range}(T) &\Leftrightarrow w = T(v) \text{ for } v \in V \\ &\Leftrightarrow [w]_{\mathcal{B}_W} = [T(v)]_{\mathcal{B}_W} \text{ for } v \in V \\ &\Leftrightarrow [w]_{\mathcal{B}_W} = A[v]_{\mathcal{B}_V} \text{ for } v \in V \\ &\Leftrightarrow [w]_{\mathcal{B}_W} \in \operatorname{colspace} A \end{aligned}$$

Since $A[v]_{\mathcal{B}_V}$ is a linear combination of the columns of A. □

The Fundamental Theorem of Linear Maps. Let V, W be finite dimensional vector spaces and let $T : V \to W$ be a linear map. Then,

$$\dim V = \dim(N(T)) + \dim(\operatorname{range}(T))$$

Proof. Fix bases of V, W. Let $\dim v = n$ and $\dim W = m$. Then the matrix A for T with respect to bases is $m \times n$. By previous theorem, $N(A)$ consists of the co-ordinate vectors for vectors in $N(T)$ Thus

$$\dim(\operatorname{range}(T)) = \dim(\operatorname{colspace} A) = \operatorname{rank} A$$

By the rank-nullity theorem, we have $\operatorname{rank} A + \operatorname{nullity} A = n$. Thus we have the result. □

Corollary 2.8. We have

$$\dim(N(T)) \geq 0$$

and

$$\dim(\operatorname{range}(T)) \leq \dim V$$

2.4 Injective and Surjective Linear Maps

Definition 2.9. Let V, W be vector spaces over the same field, and let $T : V \to W$ be a linear map. T is **injective** if for all $u, v \in V$, $T(u) = T(v) \implies u = v$.

Theorem 2.10. Let V, W be vector spaces over the same field, and let $T : V \to W$ be a linear map. Then T is injective if and only if $N(T) = \{0\}$.

Proof. We prove the forwards direction:
Suppose T is injective. Since $T(0) = 0$, we have $0 \in N(T)$. Now suppose $v \in N(T)$. Then

$$T(v) = 0 = T(0)$$

Since T is injective, this implies $v = 0$. Thus $N(T) = \{0\}$.
We prove the reverse direction:
Suppose $N(T) = \{0\}$. Let $u, v \in V$, and suppose $T(u) = T(v)$. Then since T is linear, we have

$$0 = T(u) - T(v) = T(u - v)$$

Thus $u - v \in N(T)$ and hence $u - v = 0 \implies u = v$. $\square$

Example. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear map defined by

$$T(x, y, z) = (x + y + 2z, y + z, 2x - y + z)$$

Determine if T is injective.
We use the matrix mapping theorem. The matrix A is

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

which row reduces to

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

Since there is a free variable, there are non-trivial solutions to the homogeneous matrix, and so T is not injective.

Definition 2.11. Let V, W be vector spaces over the same field, and let $T : V \to W$ be a linear map. Then T is surjective if $\text{range}(T) = W$.

Example. Let $D_1 : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ and $D_2 : P_2(\mathbb{R}) \to P_1(\mathbb{R})$ both be differentiation maps. Determine if D_i is injective or surjective.
Consider the null spaces of the maps. It will consist of polynomials that have the 0 polynominal as their derivative. This is satisfied by all constant polynomials, and so it is not injective.
The range of the map consists of all polynomials of degree 1 or less, so D_2 is surjective but not D_1.

Example. Let $T_1 : \mathbb{R}^3 \to \mathbb{R}^2$ and $T_2 : R^2 \to R^3$ be linear maps. Can T_1 be injective and an T_2 be surjective?
The answer is no. We have

$$\dim(N(T)) + \dim(\text{range}(T)) = 3$$

Since $\text{range}(T_1)$ is a subspace of $\mathbb{R}^2$, we have

$$\dim(\text{range}(T_1)) \leq \dim(\mathbb{R}^2) = 2$$

Thus

$$\dim(N(T_1)) = 3 - \dim(\text{range}(T_1)) \geq 3 - 2 = 1$$

and so $N(T_1) \neq \{0\}$ (not injective).
Since $T_2 : \mathbb{R}^2 \to \mathbb{R}^3$, we have

$$\dim(N(T_2)) + \dim(\text{range}(T_2)) = 2$$

Since $\dim(N(T_2)) \geq 0$, we have

$$\dim(\text{range}(T_2)) = 2 - \dim(N(T_2)) \leq 2$$

Since $\mathbb{R}^3$ has dimension 3, we have $\text{range}(T_2) \neq \mathbb{R}^3$ and so it is not surjective.

2.5 Invertible Linear Maps

Definition 2.12. Let V, W be vector spaces over the same field. A linear map $T : V \to W$ is invertible if there is a linear map $S : W \to V$ with $ST = I_V$ and $TS = I_W$. We call S the **inverse** of T, and write $S = T^{-1}$.

Theorem 2.13. An invertible linear map has a unique inverse.

Proof. Suppose $T : V \to W$ is invertible and S_1, S_2 are inverses of T. Then,

$$\begin{aligned} S_1 &= S_1 I_W \\ &= S_1(TS_2) \\ &= (S_1T)S_2 \\ &= I_v S_2 \\ &= S_2 \end{aligned}$$

□

Theorem 2.14. Let V, W be vector spaces over the same field, and let $T : V \to W$ be linear. Then T is invertible iff T is injective and surjective.

Theorem 2.15. Let V and W be finite dimensional vector spaces over the same field, and let $T : V \to W$ be a linear map. If T is invertible, then $\dim V = \dim W$.

Proof. Since T is invertible, the previous theorem implies T is injective and surjective. Since T is injective, we have $N(T) = \{0\}$ and so $\dim(N(T)) = 0$. Since T is surjective, we have $\operatorname{range} T = W$, and so $\dim(\operatorname{range}(T)) = \dim W$. The FTLM yields

$$\dim V = \dim(N(T)) + \dim(\operatorname{range}(T)) = \dim W$$

□

Theorem 2.16. Let V, W be finite dimensional vector spaces, and let $T : V \to W$ be a linear map. Fix bases of V, W and let A be the matrix for T. Then T is invertible iff A is invertible. Furthermore, the matrix for T^{-1} is A^{-1}.

2.6 Isomorphisms

Definition 2.17. Let V, W be vector spaces over the same field. If there is an invertible linear map $T : V \to W$, then we say V is **isomorphic** to W, and T is called an **isomorphism**.

Remark. An invertible map between vector spaces V and W pairs up the vectors between the vector spaces. Each vector in V is matched with exactly one vector in W , and vice versa. Furthermore, since the map is linear, this pairing respects the vector space operations. If there is an invertible linear map from one vector space to another, it essentially shows that other than possibly the names we are using for the vectors, they are the same vector space at their core, and the invertible linear map is telling us how to translate from one vector space to the other. The one vector space is really the same as the other, it's just "in disguise." When this happens, we saw that V is isomorphic to W , a word that literally means "same shape."

Corollary 2.18. Let V be a vector space over F and let $\dim V = n$. Then $V \simeq F^n$.

Theorem 2.19. Let V be a finite dimensional vector space with $\dim V = n$. Fix a basis $\mathcal{B}$ of V. Let $T : V \to F^n$ be defined by $T(v) = [v]_{\mathcal{B}}$. Then T is an isomorphism and V is isomorphic to F^n.

Remark. It turns out that any vector space of dimension n is isomorphic to F^n. Every n-dimensional vector space is really F^n "in disguise!"

Theorem 2.20. Let V, W be finite dimensional vector spaces over F. Then V is isomorphic to W iff $\dim V = \dim W$.

2.7 Change of Basis

Theorem 2.21. Let V be a finite dimensional vector space, and let $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ and $\mathcal{B}_2$ be two bases of V. Let P be the $n \times n$ matrix with ith column equal to $[v_i]_{\mathcal{B}_2}$. Then for any $v \in V$,

$$[v]_{\mathcal{B}_2} = P[v]_{\mathcal{B}_1}$$

P is called the **change of basis matrix** from $\mathcal{B}_1$ to $\mathcal{B}_2$.

Theorem 2.22. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two bases of a finite dimensional space V, and let P be the change of basis matrix from $\mathcal{B}_1$ to $\mathcal{B}_2$. Then P is non-singular and P^{-1} is the change of basis matrix from $\mathcal{B}_2$ to $\mathcal{B}_1$.

Example. Consider $\mathbb{R}^3$ over $\mathbb{R}$ with bases

$$\mathcal{B}_1 = \{(1,0,2), (-2,-1,-1), (8,3,8)\}$$

and

$$\mathcal{B}_2 = \{(1,1,1), (0,1,1), (0,0,1)\}$$

Find P, the change of basis matrix from $\mathcal{B}_1$ to $\mathcal{B}_2$.
Let v_i be the ith vector of $\mathcal{B}_1$. By the Change of Basis Theorem, the ith column of P is $[v_i]_{\mathcal{B}_2}$. Thus the first column of P is (a, b, c) where $a, b, c \in \mathbb{R}$ satisfies

$$a \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

and the second column is

$$a \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ -1 \end{bmatrix}$$

and the third column is

$$a \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \\ 8 \end{bmatrix}$$

We can solve all three systems at once by including a constant column:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 8 \\ 1 & 1 & 0 & 0 & -1 & 3 \\ 1 & 1 & 1 & -2 & -1 & 8 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 8 \\ 0 & 1 & 0 & -1 & 1 & -5 \\ 0 & 0 & 1 & 2 & 0 & 5 \end{array}\right]$$

Thus P is the right side of the above.
Let's test it. Let v be the vector where $[v]_{\mathcal{B}_1} = (1,2,3)$. We have

$$\begin{aligned} [v]_{\mathcal{B}_2} &= P[v]_{\mathcal{B}_1} \\ &= \begin{bmatrix} 1 & -2 & 8 \\ -1 & 1 & -5 \\ 2 & 0 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 21 \\ -14 \\ 17 \end{bmatrix} \end{aligned}$$

On one hand, since $[v]_{\mathcal{B}_1} = (1, 2, 3)$, we have

$$v = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + 2 \begin{bmatrix} -2 \\ -1 \\ -1 \end{bmatrix} + 3 \begin{bmatrix} 8 \\ 3 \\ 8 \end{bmatrix} = \begin{bmatrix} 21 \\ 7 \\ 24 \end{bmatrix}$$

and on the other, we have $[v]_{\mathcal{B}_2} = (21, -14, 17)$

$$v = 21 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 14 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + 17 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 21 \\ 7 \\ 24 \end{bmatrix}$$

Theorem 2.23. Let $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ and $\mathcal{B}_2 = \{w_1, \ldots, w_n\}$ be two bases for F^n. Then the RREF of the matrix

$$\left[\begin{array}{cccc|cccc} w_1 & w_2 & \ldots & w_n & v_1 & v_2 & \ldots & v_n \end{array}\right]$$

is

$$\left[\begin{array}{c|c} I & P \end{array}\right]$$

where P is the change of basis matrix from $\mathcal{B}_1$ to $\mathcal{B}_2$.

Example. Consider $P_1(\mathbb{R})$ with bases $\mathcal{B}_1 = \{x+1, x-1\}$ and $\mathcal{B}_2 = \{x+2, 2x+1\}$. Find P, the change of basis matrix from $\mathcal{B}_1 \to \mathcal{B}_2$.
Let $p_1(x) = x + 1$ and $p_2(x) = x - 1$. Then by the Change of Basis Theorem, the ith column of P is $[p_i(x)]_{\mathcal{B}_2}$. Thus the first column of P will be given by $a, b \in \mathbb{R}$ with

$$a(x + 2) + b(2x - 1) = x + 1$$

so we must solve the system

$$\begin{aligned} a + 2b &= 1 \\ 2a - b &= 1 \end{aligned}$$

The second column of P will be given by $a, b \in \mathbb{R}$ with

$$a(x + 2) + b(2x - 1) = x - 1$$

so we must solve the system

$$\begin{aligned} a + 2b* &= 1 \\ 2a + b &= -1 \end{aligned}$$

We have

$$\begin{array}{cc|cc} 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & -1 \end{array} \to \begin{array}{cc|cc} 1 & 0 & \frac{1}{3} & -1 \\ 0 & 1 & \frac{1}{3} & 1 \end{array}$$

Thus

$$P = \begin{bmatrix} \frac{1}{3} & -1 \\ \frac{1}{3} & 1 \end{bmatrix}$$

Theorem 2.24. Let $\mathcal{B}_1, \mathcal{B}_2$ be bases of a finite dimensional vector space V. Let P be the change of basis matrix from

$\mathcal{B}_1 \to \mathcal{B}_2$. Let $T : V \to V$ be a linear map. Let $A_{\mathcal{B}_1}, A_{\mathcal{B}_2}$ be the matrices for T with respect to A_1, A_2. Then

$$A_{\mathcal{B}_1} = P^{-1} A_{\mathcal{B}_2} P$$

Proof. By change of basis theorems, we have for any $v \in V$

$$[v]_{\mathcal{B}_2} = P[v]_{\mathcal{B}_1}$$

and

$$[v]_{\mathcal{B}_1} = P^{-1}[v]_{\mathcal{B}_2}$$

The matrix mapping theorem yields

$$[T(v)]_{\mathcal{B}_1} = A_{\mathcal{B}_1}[v]_{\mathcal{B}_1}$$

and

$$[T(v)_{\mathcal{B}_2}] = A_{\mathcal{B}_2}[v]_{\mathcal{B}_2}$$

Thus,

$$\begin{aligned} P^{-1} A_{\mathcal{B}_2} P[v]_{\mathcal{B}_1} &= P^{-1} A_{\mathcal{B}_2}[v]_{\mathcal{B}_2} \\ &= P^{-1}[T(v)]_{\mathcal{B}_2} \\ &= [T(v)]_{\mathcal{B}_1} \\ &= A_{\mathcal{B}_1}[v]_{\mathcal{B}_1} \end{aligned}$$

and since this holds for any vector in V, we have

$$P^{-1} A_{\mathcal{B}_2} P = A_{\mathcal{B}_1}$$

$\square$

Example. Consider $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T(x, y) = (3x + y, x + 3y)$$

Let $\mathcal{B}_1 = \{(1,1), (1,-1)\}$ and $\mathcal{B}_2 = \{(1,0), (0,1)\}$. Find $A_{\mathcal{B}_1}$ in 2 ways: first directly by using the matrix mapping theorem and then by using the $\mathcal{B}_2$ and the change of basis theorem. Finally, find a general formula for the individual components of $T^m(x, y)$.

Let's use the matrix mapping theorem. We know that $A_{\mathcal{B}_1}$ has columns equal to $[T(v_i)]_{\mathcal{B}_2}$. Since $T(1,1) = (4,4)$, we find $a, b \in \mathbb{R}$ with

$$a(1,1) + b(1,-1) = (4,4)$$

and similarly

$$a(1,1) + b(1,-1) = (2,-2)$$

We have

$$\left[\begin{array}{cc|cc} 1 & 1 & 4 & 2 \\ 1 & -1 & 4 & -2 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 4 & 0 \\ 0 & 1 & 0 & 2 \end{array}\right]$$

Thus we have $A_{\mathcal{B}_1} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$

We now use the new method. Since $\mathcal{B}_2$ is the standard basis, we have $A_{\mathcal{B}_2} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ Now we form the matrix $\left[\begin{array}{c|c} \mathcal{B}_2 & \mathcal{B}_1 \end{array}\right]$ which is

$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{array}\right]$$

which is already in RREF. Thus $P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Now we find P^{-1} using the usual RREF method, which is

$$\frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Then our new theorem yields

$$\begin{aligned} A_{\mathcal{B}_1} &= P^{-1}A_{\mathcal{B}_2}P \\ &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \end{aligned}$$

Now we find a formula for $T^m(x, y)$. We have

$$T^m \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}^m \begin{bmatrix} x \\ y \end{bmatrix}$$

However, we can use $A_{\mathcal{B}_1}$ and rearrange the change of basis formula.

$$\begin{aligned} A_{\mathcal{B}_2}^m &= (PA_{\mathcal{B}_1}P^{-1})^m \\ &= PA_{\mathcal{B}_1}P^{-1}PA_{\mathcal{B}_1}P^{-1}\dots && (m \text{ copies}) \\ &= PA_{\mathcal{B}_1}A_{\mathcal{B}_1}\dots A_{\mathcal{B}_1}P^{-1} && (m \text{ copies of } A_{\mathcal{B}_1}) \\ &= PA_{\mathcal{B}_1}^m P^{-1} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}^m \left(\frac{1}{2}\right)\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 4^m + 2^m & 4^m - 2^m \\ 4^m - 2^m & 4^m + 2^m \end{bmatrix} \end{aligned}$$

and so $T^m \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (4^m + 2^m)x & (4^m - 2^m)y \\ (4^m - 2^m)x & (4^m + 2^m)y \end{bmatrix}$

2.8 Similar Matrices

Definition 2.25. Let A, B be $n \times n$ matrices. Then A is **similar** to B if there exists a non-singular matrix P with $B = P^{-1}AP$. Matrix similarity is an equivalence relation.

Theorem 2.26. Two $n \times n$ matrices are similar iff they are matrices for the same linear map $T : F^n \to F^n$ with respect to two bases.

Theorem 2.27. Let A, B be similar matrices. Then $\operatorname{rank} A = \operatorname{rank} B$.

Proof. Let A, B be similar matrices. Then A, B are both matrices for the same linear map T wrt 2 bases. Since both $\operatorname{colspace} A$ and $\operatorname{colspace} B$ contain the co-ordinate vectors for $\operatorname{range} T$, we have $\dim \operatorname{colspace} A = \dim \operatorname{range} T$ and $\dim \operatorname{colspace} B = \dim \operatorname{range} T$. Thus

$$\begin{aligned} \operatorname{rank} A &= \dim \operatorname{colspace} A \\ &= \dim \operatorname{range} T \\ &= \dim \operatorname{colspace} B \\ &= \operatorname{rank} B \end{aligned}$$

3 Determinants

Definition 3.1. Given an $n \times n$ matrix A, let M_{ij} be the $n-1 \times n-1$ matrix obtained by removing the ith row and jth column of A. M_{ij} is called a **minor** of A.
We define the **determinant** of A, denoted $\det A$ or $|A|$, as follows.
If A is 1×1, then $\det A = [A]_{11}$. If A is $n \times n$ for $n \geq 2$, then

$$\det A = [A]_{11} \det M_{11} - [A]_{12} \det M_{12} + \cdots + (-1)^{n+1}[A]_{1n} \det M_{1n}$$

or

$$\sum_{i=1}^{n} (-1)^{i+1}[A]_{1i} \det M_{1i}$$

Definition 3.2. Let A be $n \times n$. The **cofactor** of $[A]_{ij}$, denoted C_{ij} is

$$C_{ij} = (-1)^{i+j} \det M_{ij}$$

Theorem 3.3. Let A be $n \times n$. Fix i with $1 \leq i \leq n$. Then

$$\det A = [A]_{i1}C_{11} + [A]_{i2}C_{i2} + \cdots + [A]_{in}C_{in}$$

or

$$\sum_{j=1}^{n} [A]_{ij}C_{ij}$$

3.1 Row Operations and Determinants

Theorem 3.4. Let A be a square matrix.

1. If B‘is obtained from A by interchanging two rows or columns, then $\det B = -\det A$.
2. If B is obtained from A by multiplying a row or column by a scalar c, then $\det B = c \det A$.
3. If B is obtained from A by adding a multiple of a row or column to another, then $\det B = \det A$.

Example. Compute the determinant of

$$\begin{bmatrix} 3 & 7 & -1 \\ -6 & -16 & 5 \\ 3 & 5 & 7 \end{bmatrix}$$

We have

$$\begin{aligned} \begin{bmatrix} 3 & 7 & -1 \\ -6 & -16 & 5 \\ 3 & 5 & 7 \end{bmatrix} &\to \begin{bmatrix} 3 & 7 & -1 \\ 0 & -2 & 3 \\ 0 & 2 & 8 \end{bmatrix} \begin{matrix} \\ R2 + 2R1 \\ R3 - R1 \end{matrix} \\ &\to \begin{bmatrix} 3 & 7 & -1 \\ 0 & -2 & 3 \\ 0 & 0 & 5 \end{bmatrix} \begin{matrix} \\ \\ R3 - R2 \end{matrix} \end{aligned}$$

This is an upper triangular matrix, and so its determinant is the product of the diagonal entries $(3)(-2)(5) = -30$. The row operations leave the determinant unchanged and so the determinant of the original matrix is also -30.

3.2 Determinants and Singularity

Theorem 3.5. The determinants of elementary matrices are as follows:

1. If E_{ij} is the elementary matrix for the ERO that interchanges row i with j, then $\det E_{ij} = -1$.
2. If $E_i(\alpha)$ is the elementary matrix for the ERO that multiplies row i by α, then $\det E_i(\alpha) = \alpha$.
3. If $E_{ij}(\alpha)$ is the elementary matrix for the ERO that adds α times row i to row j, then $\det E_{ij}(\alpha) = 1$.

Theorem 3.6. Let A be a square matrix, and E an elementary matrix of the same size. Then $\det EA = \det E \cdot \det A$.

Theorem 3.7. A square matrix A is singular if and only if $\det A = 0$.

Proof. Let B be the RREF of A. Then $\exists$ elementary matrices E_i with

$$E_k \dots E_i A = B$$

By previous theorem, we have

$$\begin{aligned}
\det B &= \det(E_k E_{k-1} \dots E_1 A) \\
&= \det E_k (\det E_{k-1} \dots E_1 A) \\
&= \dots \\
&= (\det E_k) \dots (\det E_1)(\det A)
\end{aligned}$$

Since the determinant of an elementary matrix is either $1, -1$ or α for $\alpha \neq 0$, it follows that

$$\det A = 0 \Leftrightarrow \det B = 0$$

Now we prove that A is singular if and only if $\det A = 0$.
Suppose $n \times n$ matrix A is singular, then $\operatorname{rank} A < n$, and so B has at least one zero row, so that means $\det B = 0$ by expanding on this row. Thus it follows that $\det A = 0$.
Suppose A is non-singular. Then $B = I_n$ and so $\det B = 1 \neq 0$. It follows that $\det A \neq 0$. □

Example. Is

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

singular?
We have

$$\begin{aligned}
\det A &= 1 \det \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} - 2 \det \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} + 3 \det \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix} \\
&= 45 - 48 = 2(36 - 42) + 3(32 - 35) \\
&= 0
\end{aligned}$$

Since the determinant is 0, we have that A is singular.
Alternatively, we could have seen that the columns of A are linearly dependent, or we could have row reduced.

Example. Find all values of $x \in \mathbb{R}$ for which

$$A = \begin{bmatrix} -2 & x & 1 \\ x & 1 & 1 \\ 2 & 3 & -1 \end{bmatrix}$$

is invertible.
We have

$$\begin{aligned} \det A &== 2 \det \begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix} - x \det \begin{bmatrix} x & 1 \\ 2 & -1 \end{bmatrix} + 1 \det \begin{bmatrix} x & 1 \\ 2 & 3 \end{bmatrix} \\ &= -2(1-3) - x(-x-2) + 3x - 2 \\ &= x^2 + 5x + 6 \\ &= (x+2)(x+3) \end{aligned}$$

Thus $\det A = 0 \Leftrightarrow x = -3, -2$. So A is invertible for all x except $-2, -3$.

3.3 Additional Results on Determinants

Theorem 3.8. Let A and B be square matrices of the same size, then

$$\det AB = \det A \cdot \det B$$

Theorem 3.9. Let A be a non-singular square matrix. Then $\det A^{-1} = (\det A)^{-1}$.

Theorem 3.10. Let A, B be similar. Then $\det A = \det B$.

Theorem 3.11. Let A be square matrix. Then $\det A^t = \det A$.

Theorem 3.12. Let A be upper/lower triangular matrix. Then

$$\det A = \prod_{i=1}^{n} a_{ii}$$

4 Eigenvalues and Eigenvectors

Definition 4.1. Let V be a vector space over a field of scalars F. Let $T : V \to V$ be a linear map. If there is a scalar λ such that

$$T(v) = \lambda v$$

for some $v \in V$ with $v \neq 0$, then λ is called an **eigenvalue** of T with corresponding **eigenvector** v.

Example. Consider the linear map $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T(x, y) = (-8x + ty, -10x + 7y)$$

We were given basis $\mathcal{B} = \{(1,1), (1,2)\}$. We have

$$\begin{aligned} T(1,1) &= -3(1,1) \\ T(1,2) &= 2(1,2) \end{aligned}$$

thus $-3, 2$ are eigenvalues with $(1,1), (1,2)$ their corresponding eigenvectors.
Notice that

$$\begin{aligned}T(c(1,1)) &= cT(1,1)\\ &= c(-3,-3)\\ &= -3c(1,1)\end{aligned}$$

and so $c(1,1)$ is also an eigenvector of T corresponding to eigenvalue -3.

Definition 4.2. Let A be an $n \times n$ matrix with entries in F. If $\exists \lambda \in F$ such that

$$Av = \lambda v$$

for some $v \in F^n$ with $v \neq 0$, then λ is called an **eigenvalue** of A with corresponding **eigenvector** v.

Remark. How can we find the eigenvectors and values? We have

$$Av = \lambda v \Leftrightarrow Av - \lambda v = 0$$

We add an identity so we can factor and get

$$Av = \lambda v \Leftrightarrow (A - \lambda I)v = 0$$

and so

$$Av = \lambda v \Leftrightarrow v \in N(A - \lambda I)$$

Theorem 4.3. Suppose λ is an eigenvalue of $n \times n$ matrix A. Then eigenvectors corresponding to λ are non-zero vectors in $N(A - \lambda I)$, subspace of F^n. Denote this subspace by E_λ and call it the **eigenspace** of A corresponding to λ.

Remark. If $N(A - \lambda I) = \{0\}$, that is if it is non singular, then λ has no corresponding eigenvectors, that is λ is not an eigenvalue. Thus we must have $A - \lambda I$ singular, which means the determinant must be 0.

Definition 4.4. Let A be square matrix. Then $\det(A - \lambda I)$ is called the **characteristic polynominal of** A, denoted by $p_A(\lambda)$.

Theorem 4.5. Let A be square matrix. Then λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$.

Example. Find the eigenvalues and vectors for

$$A = \begin{bmatrix} -8 & 5 \\ -10 & 7 \end{bmatrix}$$

We have

$$\begin{aligned}\det(A - \lambda I) &= \det\left(\begin{bmatrix} -8 & 5 \\ -10 & 7 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\\ &= \det \begin{bmatrix} -8 - \lambda & 5 \\ -10 & 7 - \lambda \end{bmatrix}\\ &= \lambda^2 + \lambda - 6\\ &= (\lambda - 2)(\lambda + 3)\end{aligned}$$

and so $\det(A - \lambda I) = 0$ for $\lambda = -3, 2$. These are the eigenvalues of A.
Now we find E_{-3}. We have

$$E_{-3} = N(A - (-3)I) = N\left(\begin{bmatrix} -5 & 5 \\ -10 & 10 \end{bmatrix}\right)$$

which row reduces to

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

Thus

$$E_{-3} = \left\{ t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mid t \in \mathbb{R} \right\} = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Similarly we have

$$E_2 = N(A - 2I) = N\left(\begin{bmatrix} -10 & 5 \\ -10 & 5 \end{bmatrix}\right)$$

which row reduces to

$$\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}$$

Thus

$$E_2 = \left\{ t \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix} \mid t \in \mathbb{R} \right\} = \text{span}\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$

Example. Find the eigenvalues and corresponding eigenspaces for

$$A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{bmatrix}$$

We have

$$\begin{aligned}
\det(A - \lambda I) &= \det \begin{bmatrix} 1-\lambda & -3 & 3 \\ 3 & -5-\lambda & 3 \\ 6 & -6 & 4-\lambda \end{bmatrix} \\
&= \det \begin{bmatrix} 1-\lambda & -2-\lambda & 3 \\ 2+\lambda & 0 & 0 \\ 6 & 0 & 4-\lambda \end{bmatrix} \\
&= (-1)^{1+2}(-2-\lambda) \det \begin{bmatrix} 2+\lambda & 0 \\ 6 & 4-\lambda \end{bmatrix} \\
&= (\lambda+2)((2+\lambda)(4-\lambda) - 0) \\
&= -(\lambda+2)^2(\lambda-4)
\end{aligned}$$

and so $\lambda = -2, 4$ which are the eigenvalues.
We have

$$E_4 = N(A - 4I) = N\left(\begin{bmatrix} -3 & 3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{bmatrix}\right)$$

which row reduces to

$$\begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

Thus

$$E_4 = \left\{ t \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} \mid t \in \mathbb{R} \right\} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \right\}$$

We have

$$E_{-2} = N(A - (-2)I) = N\left(\begin{bmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{bmatrix} \right)$$

which row reduces to

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus

$$E_4 = \left\{ s \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \mid s, t \in \mathbb{R} \right\} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

4.1 Solving Higher Degree Polynomial Equations

Fundamental Theorem of Algebra. Every degree n polynominal with coefficients in $\mathbb{C}$ factors uniquely as

$$a(x - c_1) \dots (x - c_n)$$

for $a, c_i \in \mathbb{C}$ with c_i not necessarily distinct. When factoring over $\mathbb{R}$, there may be irreducible quadratic factors. Every degree n polynomial with coefficients in $\mathbb{R}$ factors uniquely as

$$a(x - r_1) \dots (x - r_m)(x^2 + b_1 x + c_1) \dots (x^2 + b_p x + c_p)$$

for $a, r_i, b_i, c_i \in \mathbb{R}$ where $m + 2p = n$ and each $x^2 + b_i x + c_i$ is irreducible over $\mathbb{R}$. Note it may be that $m = 0$ or $p = 0$.

Theorem 4.6. Factoring in $\mathbb{Z}_p$ is more complicated. For example, consider

$$x^2 + 1 = 0$$

Over $\mathbb{Z}_3$, it is best to substitute $1, 2, 3$ into the equation and reducing. Over $\mathbb{Z}_2$, since $2 \equiv 0$, we have

$$x^2 + 1 \equiv (x + 1)^2 \pmod 2$$

We see that $x \equiv 1$ is a solution.

Remark. Remember factor theorem and long division from high school.

4.2 Basic Theorems on Eigenvalues and Eigenvectors

Theorem 4.7. Let A be square. 0 is an eigenvalue of A if and only if A is singular.

Proof. We have

$$\begin{aligned} 0 \text{ is an eigenvalue of } A &\Leftrightarrow \det(A - 0I) = 0 \\ &\Leftrightarrow \det A = 0 \\ &\Leftrightarrow A \text{ is singular.} \end{aligned}$$

□

Theorem 4.8. Let A be square matrix with eigenvalue λ.

1. λ is an eigenvalue of A^t.
2. If $\lambda \neq 0$, then λ^{-1} is an eigenvalue of A^{-1}.
3. Let c be a scalar, then $c\lambda$ is an eigenvalue of cA.
4. Let $m \in \mathbb{N}$. Then λ^m is an eigenvalue of A^m.

Theorem 4.9. Let A, B be similar matrices. Then A, B have the same characteristic polynominal, and hence the same eigenvalues.
Note that the eigenvectors may be different, with $w = Pv$ as an eigenvalue.

Proof. We have $B = P^{-1}AP$ for some non-singular matrix P. Thus

$$\begin{aligned}
p_B(\lambda) &= \det(B - \lambda I) \\
&= \det\big(P^{-1}AP - \lambda P^{-1}IP\big) \\
&= \det\big(P^{-1}AP - P^{-1}(\lambda I)P\big) \\
&= \det\big(P^{-1}(A - \lambda I)P\big) \\
&= (\det P^{-1})(\det(A - \lambda I))(\det P) \\
&= (\det P)^{-1}(\det P)(\det(A - \lambda I)) \\
&= \det(A - \lambda I) \\
&= p_A(\lambda)
\end{aligned}$$

□

4.3 Complex Eigenvalues for Real Matrices

Example. Find the eigenvalues and associated eigenspaces for

$$A = \begin{bmatrix} 2 & 2 \\ -4 & -2 \end{bmatrix}$$

We have

$$\begin{aligned}
\det(A - \lambda I) &= \det \begin{bmatrix} 2-\lambda & 2 \\ -4 & -2-\lambda \end{bmatrix} \\
&= (2-\lambda)(-2-\lambda) - (2)(-4) \\
&= \lambda^2 + 4
\end{aligned}$$

This has no real eigenvalues, but if we are working over $\mathbb{C}$, then we have $\lambda^2 = -4$ and so $\lambda = \pm 2i$. We have $\lambda_1 = 2i$ and $\lambda_2 = -2i$.
Let's find E_{2i}. We have

$$E_{2i} = N(A - 2iI) = N\left(\begin{bmatrix} 2-2i & 2 \\ -3 & -2-2i \end{bmatrix}\right)$$

We row reduce:

$$\begin{bmatrix} 2-2i & 2 \\ -4 & -2-2i \end{bmatrix} \to \begin{bmatrix} 1 & \frac{1+i}{2} \\ 0 & 0 \end{bmatrix}$$

Thus $x_2 = t$ is free and $x_1 = \left(-\frac{1+i}{2}\right) t$. We have

$$E_{2i} = \left\{ t \begin{bmatrix} -\frac{1+i}{2} \\ 1 \end{bmatrix} \mid t \in \mathbb{C} \right\}$$
$$= \operatorname{span} \left\{ \begin{bmatrix} 1+i \\ -2 \end{bmatrix} \right\}$$

Theorem 4.10. Let $A \in M_{nn}(\mathbb{R})$. Let λ be an eigenvalue of A with corresponding eigenvector v. Let $\overline{v}$ be the vector with entries that are the complex conjugates of the entries of v. Then $\overline{\lambda}$ is an eigenvalue of A with corresponding eigenvector $\overline{v}$.

Example. For the example above,

$$E_{-2i} = \operatorname{span} \left\{ \begin{bmatrix} 1-i \\ -2 \end{bmatrix} \right\}$$

4.4 The Characteristic Polynomial

Theorem 4.11. Let F be a field and let A be $n \times n$ with entries in $P_1(F)$. Then $\det A \in P_n(F)$.

Theorem 4.12. Let F be a field, $n \geq 2$, and $A \in M_{nn}(F)$. Let $[A]_{ij} = a_{ij}$. Then

$$p_A(\lambda) = (a_{11} - \lambda) \dots (a_{nn} - \lambda) + g(\lambda)$$

where $g(\lambda) \in P_{n-2}(\lambda)$.

Theorem 4.13. Let F be a field and let $A_{nn}(F)$. Then $p_A(\lambda)$ is a degree n polynominal in λ with

$$p_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \operatorname{Tr} A \lambda^{n-1} + c_{n-2}\lambda^{n-2} + \cdots + c_1 \lambda + \det A$$

Example. If

$$A = \begin{bmatrix} 5 & 2 \\ -3 & 7 \end{bmatrix}$$

Then $\operatorname{Tr} A = 12$ and $\det A = 41$. Thus

$$p_A(\lambda) = \lambda^2 - 12\lambda + 41$$

Definition 4.14. If $f(x) \in P_n(F)$ factors into a product of degree one factors over F, that is if

$$f(x) = a \prod_{i=1}^{n} (x - c_i)$$

for $a, c_i \in F$, then we say that f **splits** over F.

Theorem 4.15. If

$$p_A(\lambda) = (-1)^n \prod_{i=1}^{n} (\lambda - \lambda_i)$$

then

$$\det A = \prod_{i=1}^{n} \lambda_i$$

and

$$\operatorname{Tr} A = \sum_{i=1}^{n} \lambda_i$$

That is, if $p_A(\lambda)$ splits over F, then $\det A$ is equal to the product of eigenvalues with multiplicity and $\operatorname{Tr} A$ is equal to the sum of eigenvalues with multiplicity.

4.5 Algebraic and Geometric Multiplicity

Definition 4.16. Suppose $A \in M_{nn}(F)$ has eigenvalue λ_1 and

$$p_A(\lambda) = (\lambda - \lambda_1)^{\alpha_A(\lambda_1} q_A(\lambda)$$

where $\alpha_A(\lambda_1) \in \mathbb{N}$ and $q_A(\lambda_1) \neq 0$. Then $\alpha_A(\lambda_1)$ is the **algebraic multiplicity** of λ_1. The **geometric multiplicity** $\gamma_A(\lambda_1)$ of λ_1 is equal to $\dim(E_{\lambda_1}$ or $\dim(A - \lambda_1 I)$. If the matrix A is clear, can write $\alpha(\lambda_1)$ and $\gamma(\lambda_1)$ instead.

Theorem 4.17. Suppose $A \in M_{nn}(F)$ has eigenvalue λ_1. Then,

$$1 \leq \gamma(\lambda_1) \leq \alpha(\lambda_1) \leq n$$

Proof. Since λ_1 is an eigenvalue with algebraic multiplicity $\alpha(\lambda_1)$, $(\lambda - \lambda_1)^{\alpha(\lambda_1)}$ is a factor of the characteristic polynominal $p_A(\lambda)$, a polynominal of degree n. Thus, $\alpha(\lambda_1) \leq n$. Since λ_1 is an eigenvalue, there is some $v \neq 0$ in E_{λ_1}. Thus $E_{\lambda_1} \neq \{0\}$ and so $\gamma(\lambda_1) = \dim E_{\lambda_1} \geq 1$.
Let $g = \gamma(\lambda_1)$, and let $v_1, \ldots, v_g$ be a basis of E_{λ_1}. By the Linear Independence to Basis Theorem, we can find $u_1, \ldots, u_{n-g} \in F^n$ so that $v_1, \ldots, v_g, u_1, \ldots, u_{n-g}$ is a basis of F^n. Let $P = \begin{bmatrix} v_1 & \ldots & v_g & u_1 & \ldots & u_{n-g} \end{bmatrix}$. Then P is $n \times n$. Since the columns of P are linearly independent, P is non-singular and is invertible. Note that

$$\begin{bmatrix} e_1 & \ldots & e_n \end{bmatrix} = I = P^{-1}P = P^{-1} \begin{bmatrix} v_1 & \ldots & v_g & u_1 & \ldots & u_{n-g} \end{bmatrix}$$

and so for $1 \leq i \leq g$, we have

$$P^{-1} v_i = e_i$$

Let $B = P^{-1}AP$. Since A, B are similar, we have $p_A(\lambda) = p_B(\lambda)$. Since

$$B = P^{-1}AP = P^{-1}A \begin{bmatrix} v_1 & \ldots & v_g & u_1 & \ldots & u_{n-g} \end{bmatrix}$$

for $1 \leq i \leq g$, the ith column of B is equal to $P^{-1}Av_i$.
Since $v_i \in E_{\lambda_1}$, for $1 \leq i \leq g$, the ith column of B is equal to

$$\begin{aligned} P^{-1}Av_i &= P^{-1}(\lambda_1 v_i) \\ &= \lambda_1 P^{-1} v_i \\ &= \lambda_1 e_i \end{aligned}$$

Thus

$$\begin{aligned} p_A(\lambda) &= p_B(\lambda) \\ &= \det \begin{bmatrix} \lambda_1 - \lambda & 0 & \ldots & 0 & c_{11} & \ldots & c_{1(n-g)} \\ 0 & \lambda_1 - \lambda & \ldots & 0 & c_{21} & \ldots & c_{2(n-g)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_1 - \lambda & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 0 & \vdots & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & c_{n1} & \ldots & c_{n(n-g)} \end{bmatrix} & = (\lambda_1 - \lambda)^g \det M \end{aligned}$$

by expanding the determinant along the first column, repeated g times for some matrix M (where the c_{ij} may depend on λ)

We have shown that

$$p_A(\lambda) = (\lambda_1 - \lambda)^g \det M = (-1)^g(\lambda - \lambda_1)^g \det M$$

for some matrix M. Since $\det M$ is some polynominal in λ, we have shown that $(\lambda - \lambda_1)^g$ is a factor of $p_A(\lambda)$. Since $\alpha(\lambda_1)$ is the greatest power of $(\lambda - \lambda_1)$ that divides $p_A(\lambda)$, we have $\gamma(\lambda_1) = g \leq \alpha(\lambda_1)$. □

Example. Find a basis for each eigenspace of $A = \begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

We have

$$\det(A - \lambda I) = -(\lambda - 8)(\lambda + 1)^2$$

Thus the eigenvalues are $-1, 8$.
Since the algebraic multiplicity of $\lambda = 8$ is $\alpha(8) = 1$, and by previous inequality, have $\dim E_8 = \gamma(8) = 1$.
Since the algebraic multiplicity of $\lambda = -1$ is $\alpha(-1) = 2$, and $1 \leq \gamma(-1) \leq \alpha(-1) = 2$, have $\dim E_{-1} = \gamma(-1)$ either equal to 1 or 2.
We have

$$E_{-1} = N(A + I) = N\left(\begin{bmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{bmatrix}\right)$$

If we can find linear combinations of the columns that yield 0, then the vector containing the coefficients is in E_{-1}, which we know is either 1-dimensional or 2-dimensional.
Note that column 1 minus 3 is 0, thus $(1, 0, -1) \in E_{-1}$ and that column 1 minus twice column 2 is 0, thus $(1, -2, 0) \in E_{-1}$. A basis of E_{-1} are those two vectors.
If we weren't able to find the above 2 by inspection, we would have to row reduce.
Have

$$E_8 = N(A - 8I) = N\left(\begin{bmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{bmatrix}\right)$$

We know it is 1-dimensional. Can notice that the sum of the first and last columns is $-\frac{1}{2}$ times the second, or $2v_1 + v_2 + 2v_3 = 0$ and so a basis is

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \right\}$$

4.6 Combining Spaces of Eigenspaces

Theorem 4.18. Let A be square matrix, and for $i = 1, \ldots, m$, let v_i be an eigenvector corresponding to eigenvalue λ_i where each λ_i is distinct ($\lambda_i \neq \lambda_j$ for $i \neq j$). Then $v_i, \ldots, v_m$ are linearly independent.

Proof. If $m = 1$, then v_1 is linearly independent since eigenvectors are non-zero.
Suppose $m \geq 2$, and use a proof by contradiction. Suppose $v_1, \ldots, v_m$ are linearly dependent. Let k be an integer such that $v_1, \ldots, v_k$ is linearly dependent but $v_1, \ldots, v_k$ is linearly independent. Note that the vector v_1 by itself is linearly independent. Note that if the collection $v_1, \ldots, v_k$ is linearly dependent, then $v_1, \ldots, v_\ell$ where $\ell \geq k$ is as well.
Since $v_1, \ldots, v_k$ are dependent, there are c_i not all 0 with

$$c_1 v_1 + \cdots + c_k v_k = 0$$

Using the fact that v_i is an eigenvector of A corresponding to eigenvalue λ_i, we have

$$\begin{aligned}
0 &= (A - \lambda_k I)0 \\
&= (A - \lambda_k I)(c_1 v_1 + \cdots + c_k v_k) \\
&= c_1(A - \lambda_k I)v_1 + \cdots + c_{k-1}(A - \lambda_k I)v_{k-1} + c_k(A - \lambda_k I)v_k \\
&= c_1(Av_1 - \lambda_k v_1) + \cdots + c_{k-1}(Av_{k-1} - \lambda_k v_{k-1}) + c_k(Av_k - \lambda_k v_k) \\
&= c_1(\lambda_1 v_1 - \lambda_k v_1) + \cdots + c_{k-1}(\lambda_{k-1} v_{k-1} - \lambda_k v_{k-1}) + c_k(\lambda_k v_k - \lambda_k v_k) \\
&= c_1(\lambda_1 - \lambda_k)v_1 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} + 0 \\
&= c_1(\lambda_1 - \lambda_k)v_1 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1}
\end{aligned}$$

Since $v_1, \ldots, v_{k-1}$ are independent, have $c_i(\lambda_i - \lambda_k) = 0$ for all $1 \le i \le k-1$. Since $\lambda_i \neq \lambda_k$, have $c_i = 0$ for all i.
Recall have

$$c_1 v_1 + \cdots + c_k v_k = 0$$

for not all c_i zero. Since $c_i = 0$ for $1 \le i \le k-1$, this yields

$$c_k v_k = 0$$

with $c_k \neq 0$. Thus $v_k = 0$. This is a contradiction as v_k is an eigenvector and cannot be 0. □

Theorem 4.19. Let A be square. The union of bases of distinct eigenspaces of A is linearly independent.

Proof. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of A. Let

$$\mathcal{B}_i = \{v_{1i}, \ldots, v_{in_i}\}$$

be a basis for E_{λ_i} so $n_i = \dim E_{\lambda_i}$. Suppose

$$\sum_{i=1}^{m} \sum_{j=1}^{n_i} c_{ij} v_{ij} = 0$$

Let

$$w_i = \sum_{j=1}^{n_i} c_{ij} v_{ij}$$

Since w_i is a linear combination of eigenvectors corresponding to λ_i, we have $w_i \in E_{\lambda_i}$. Thus either w_i is an eigenvector corresponding to λ_i or $w_i = 0$. Then we have

$$\sum_{i=1}^{m} w_i = 0$$

If some $w_i = 0$, then they can be dropped from the sum without changing it. This is a sum of eigenvectors from distinct eigenspaces. By the previous theorem, the vectors appearing in this sum are linearly independent, and so this sum must be empty (since if not, then these vectors would be linearly dependent). That is, for each i we have $w_i = 0$.
Using the definition of w_i, for each i we have

$$0 = w_i = \sum_{j=1}^{n_i} c_{ij} v_{ij}$$

Since $v_{i1}, \ldots, v_{in}$ form a basis of E_{λ_i}, they are independent. Thus for each i we have $c_{ij} = 0$ for all j. That is, $c_{ij} = 0$ for all i, j, and so v_{ij} are independent. □

4.7 Diagonalization

Definition 4.20. Let A be square. A is **diagonalizable** if A is similar to a diagonal matrix. That is, if there is a diagonal matrix D and a non-singular matrix P such that $D = P^{-1}AP$.

Theorem 4.21. A $n \times n$ matrix A is diagonalizable iff there is a basis of F^n consisting of eigenvectors of A.

Proof. We prove the reverse direction. Let

$$\mathcal{B} = \{v_1, \dots, v_n\}$$

be a basis of F^n consisting of eigenvectors of A, so that for each i we have

$$Av_i = \lambda_i v_i$$

for some eigenvalue λ_i (not necessarily distinct). We must show that A is similar to a diagonal matrix. Let P be the change of basis matrix from $\mathcal{B}$ to the standard basis $\mathcal{S}$ of F^n. Then by the change of basis theorem, we have

$$P = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$$

Note that $[v_i]_{\mathcal{B}} = e_i$. Since P is the change of basis matrix from $\mathcal{B}$ to $\mathcal{S}$, it follows that P^{-1} is the change of basis matrix from $\mathcal{S}$ to $\mathcal{B}$. Thus for any v, we have $P[v]_{\mathcal{B}} = [v]_{\mathcal{S}}$ and $P^{-1}[v]_{\mathcal{S}} = [v]_{\mathcal{B}}$. Also since $\mathcal{S}$ is the standard basis, for any v we have $[v]_{\mathcal{S}} = v$.
The ith column of $P^{-1}AP$ is

$$\begin{aligned} P^{-1}APe_i &= P^{-1}AP[v_i]_{\mathcal{B}} \\ &= P^{-1}A[v_i]_{\mathcal{S}} \\ &= P^{-1}Av_i \\ &= P^{-1}\lambda_i v_i \\ &= \lambda_i P^{-1} v_i \\ &= \lambda_i P^{-1}[v_i]_{\mathcal{S}} \\ &= \lambda_i [v_i]_{\mathcal{B}} \\ &= \lambda_i e_i \end{aligned}$$

Since the ith column of $P^{-1}AP$ is $\lambda_i e_i$, $P^{-1}AP = D$ where D is a diagonal matrix with $[D]_{ii} = \lambda_i$. Thus A is similar.
We now prove the forward direction. Suppose A is similar to a diagonal matrix D. Then $D = P^{-1}AP$ for some non-singular matrix P. We must show there is a basis of F^n consisting of eigenvectors of A. Let $\mathcal{B} = \{v_1, \dots, v_n\}$ consist of the columns of P. Since P is non-singular, $\mathcal{B}$ is a basis of F^n. We must now show that each v_i is an eigenvector of A. Note that since $\mathcal{B}$ is a basis, we must have $v_i \neq 0$ for each i.
Let $[D]_{ii} = \lambda_i$. We have $AP = PD$. Since the ith column of PD is a linear combination of the columns of P using the entries of the ith column of D as coefficients, it is equal to $\lambda_i v_i$. On the other hand, the ith column of AP is Av_i. Thus we have

$$Av_i = \lambda_i v_i$$

and so each v_i is an eigenvector of A. $\square$

Example. We saw that the eigenvalues of $A = \begin{bmatrix} 1 & -3 & 3 \\ 3 & -5 & 4 \\ 6 & -6 & 4 \end{bmatrix}$ are 4 and -2 with $E_4 = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\right\}$ and $E_{-2} = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$

Thus $A = PDP^{-1}$ for $P = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

. Let A be square matrix. Suppose $p_A(\lambda)$ splits over F. Then

$$\gamma(\lambda_i) = \alpha(\lambda_i)$$

for each i iff A is diagonalizable.

Proof. Let A be $n \times n$ matrix. Then $p_A(\lambda)$ is a degree n polynominal. Let λ_i be the distinct eigenvalues of A. We show the forwards direction. We have

$$n = \sum_i \alpha(\lambda_i) = \sum_i \gamma(\lambda_i) = \sum_i \dim E_{\lambda_i}$$

Thus the union of bases of distinct eigenspaces contains n vectors. By a previous theorem, the vectors in this union are linearly independent. Since $\dim F^n = n$, this implies that these vectors form a basis of F^n, and hence A is diagonalizable.
We show the reverse direction. Suppose we have $\gamma(\lambda_k) \neq \alpha(\lambda_k)$ for some k. Since we have previously shown that for all i we have $\gamma(\lambda_i) \leq \alpha(\lambda_i)$, it follows that $\gamma(\lambda_k) < \alpha(\lambda_k)$. Then

$$n = \sum_i \alpha(\lambda_i) > \sum_i \gamma(\lambda_i) = \sum_i \dim E_{\lambda_i}$$

Thus the union of bases of distinct eigenspaces contains strictly less than n vectors. It follows that there is no basis of F^n consisting of eigenvectors of A since if there were, we would have $\sum_i \dim E_{\lambda_i} = n$ and so A is not diagonalizable. □

Theorem 4.22. Let A be a diagonalizable matrix with $A = PDP^{-1}$ for D a diagonal matrix and P non-singular. Then for all $k \in \mathbb{N}$, we have

$$A^k = PD^kP^{-1}$$

for a non-singular matrix P.

Proof. We have

$$\begin{aligned} A^k &= (PDP^{-1})^k \\ &= (PDP^{-1})(PDP^{-1})\ldots(PDP^{-1}) \\ &= PD(P^{-1}P)D(P^{-1}P)\ldots(P^{-1}P)DP^{-1} \\ &= PD\ldots DP^{-1} \\ &= PD^kP^{-1} \end{aligned}$$

□

Example. Let $A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$ and let $k \in \mathbb{N}$. Find A^k.
We have

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 2-\lambda & -4 \\ -1 & -1-\lambda \end{bmatrix} \\ &= (2-\lambda)(-1-\lambda) - (-4)(-1) \\ &= \lambda^2 - \lambda + 6 \\ &= (\lambda - 3)(\lambda + 2) \end{aligned}$$

Thus the eigenvalues of A are $-2, 3$ and are both $1D$.

Since $A - 3I = \begin{bmatrix} -1 & -4 \\ -1 & -4 \end{bmatrix}$, by inspection a basis for E_3 is span $\left\{ \begin{bmatrix} 4 \\ -1 \end{bmatrix} \right\}$. Since $A + 2I = \begin{bmatrix} 4 & -4 \\ -1 & 1 \end{bmatrix}$, by inspection a basis for E_{-2} is span $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$. Thus A is diagonalizable with $P = \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}$. Note that $P^{-1} = \frac{1}{4(1)-(1)(-1)} \begin{bmatrix} 1 & -1 \\ 1 & 4 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 1 & -1 \\ 1 & 4 \end{bmatrix}$ Thus

$$\begin{aligned}
A^k &= PD^kP^{-1} \\
&= \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}^k \frac{1}{5} \begin{bmatrix} 1 & -1 \\ 1 & 4 \end{bmatrix} \\
&= \frac{1}{5} \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3^k & 0 \\ 0(-2)^k & \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 4 \end{bmatrix} \\
&= \frac{1}{5} \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 3^k & -3^k \\ (-2)^k & 4(-2)^k \end{bmatrix} \\
&= \frac{1}{5} \begin{bmatrix} 4 \cdot 3^k + (-2)^k & -4 \cdot 3^k + 4(-2)^k \\ -3^k + (-2)^k & 3^k + 4(-2)^k \end{bmatrix}
\end{aligned}$$

4.8 Fibonacci Numbers

Definition 4.23. The **Fibonacci numbers** satisfy $F_1 = 1, F_2 = 1$ and

$$F_n = F_{n-1} + F_{n-2}$$

for $n \geq 3$.

Example. Notice that

$$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{n-1} \\ F_{n-2} \end{bmatrix}$$

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

$$\begin{aligned}
\begin{bmatrix} F_3 \\ F_2 \end{bmatrix} &= A \begin{bmatrix} F_2 \\ F_1 \end{bmatrix} = A \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
\begin{bmatrix} F_4 \\ F_3 \end{bmatrix} &= A \begin{bmatrix} F_3 \\ F_2 \end{bmatrix} = A^2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
\begin{bmatrix} F_5 \\ F_4 \end{bmatrix} &= A \begin{bmatrix} F_4 \\ F_3 \end{bmatrix} = A^3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\end{aligned}$$

and in general,

$$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = A^{n-2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

We find the eigenvalues and eigenvectors of A.

Have

$$\det(A-\lambda I) = \det\begin{bmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{bmatrix}$$
$$= (1-\lambda)(-\lambda) - 1$$
$$= \lambda^2 - \lambda - 1$$

and so

$$\lambda = \frac{-(-1) \pm \sqrt{(-1)^2 - 4(-1)}}{2} = \frac{1 \pm \sqrt{5}}{2}$$

and the eigenvalues are $\lambda_1 = \frac{1+\sqrt{5}}{2}$ and $\lambda_2 = \frac{1-\sqrt{5}}{2}$.
Now, find a basis for E_{λ_1}.
Have

$$A - \lambda_1 I = \begin{bmatrix} \frac{1-\sqrt{5}}{2} & 1 \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}$$

Since column 1 minus $\frac{1-\sqrt{5}}{2}$ times column 2 is zero, see that

$$\begin{bmatrix} 1 \\ -\frac{1-\sqrt{5}}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ -\lambda_2 \end{bmatrix}$$

is a basis for E_{λ_1}. Similarly, a basis for E_{λ_2} is

$$\begin{bmatrix} 1 \\ -\lambda_1 \end{bmatrix}$$

Now, write $(1,1)$ as a linear combination of eigenvectors.

$$c_1 \begin{bmatrix} 1 \\ -\lambda_2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -\lambda_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

So, have

$$\begin{bmatrix} 1 & 1 & 1 \\ -\lambda_2 & -\lambda_1 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -\frac{1+\lambda_1}{\lambda_2-\lambda_1} \\ 0 & 1 & \frac{1+\lambda_2}{\lambda_2-\lambda_1} \end{bmatrix}$$

Since λ_i satisfies $\lambda^2 - \lambda - 1 = 0$, have $\lambda + 1 = \lambda^2$. Also,

$$\lambda_2 - \lambda_1 = \frac{1-\sqrt{5}}{2} - \frac{1+\sqrt{5}}{2} = -\sqrt{5}$$

Thus $c_1 = \frac{1}{\sqrt{5}}\lambda_1^2$ and $c_2 = -\frac{1}{\sqrt{5}}\lambda_2^2$.
Therefore,

$$\begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} = A^{n-2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$= A^{n-2}\left(\frac{1}{\sqrt{5}}\lambda_1^2 \begin{bmatrix} 1 \\ -\lambda_2 \end{bmatrix} - \frac{1}{\sqrt{5}}\lambda_2^2 \begin{bmatrix} 1 \\ -\lambda_1 \end{bmatrix}\right)$$
$$= \frac{1}{\sqrt{5}}\left(\lambda_1^2 A^{n-2}\begin{bmatrix} 1 \\ -\lambda_2 \end{bmatrix} - \lambda_2^2 A^{n-2}\begin{bmatrix} 1 \\ -\lambda_1 \end{bmatrix}\right)$$
$$= \frac{1}{\sqrt{5}}\left(\lambda_1^2 \lambda_1^{n-2}\begin{bmatrix} 1 \\ -\lambda_2 \end{bmatrix} - \lambda_2^2 \lambda_2^{n-2}\begin{bmatrix} 1 \\ -\lambda_1 \end{bmatrix}\right)$$
$$= \frac{1}{\sqrt{5}}\left(\lambda_1^n \begin{bmatrix} 1 \\ -\lambda_2 \end{bmatrix} - \lambda_2^n \begin{bmatrix} 1 \\ -\lambda_1 \end{bmatrix}\right)$$

Thus
$$F_n = \frac{1}{\sqrt{5}}(\lambda_1^n - \lambda_2^n) = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$$
is an explicit formula for F_n.

4.9 Population Models

Example. Suppose in year n there are x_n children and y_n adult goats. Each year, the following happens

- Each adult has 1.5 offspring. The offspring will be children in the next year.
- 30% of children survive to be an adult in the next year.
- 40% of adults survive in the next year.

What is the long term behaviour?

Have
$$\begin{aligned} x_{n+1} &= 1.5y_n \\ y_{n+1} &= 0.3x_n + 0.4y_n \end{aligned}$$
and so
$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1.5 \\ 0.3 & 0.4 \end{bmatrix} \begin{bmatrix} x_n \\ y_n \end{bmatrix}$$
Letting $A = \begin{bmatrix} 0 & 1.5 \\ 0.3 & 0.4 \end{bmatrix}$ we have
$$\begin{aligned} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} &= A \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \\ \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} &= A \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = A^2 \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \end{aligned}$$
and in general,
$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = A^n \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$
We have
$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} -\lambda & 1.5 \\ 0.3 & 0.4-\lambda \end{bmatrix} \\ &= \lambda^2 - 0.4\lambda - 0.45 \\ &= (\lambda - 0.9)(\lambda + 0.5) \end{aligned}$$
Thus the eigenvalues of A are $\lambda_1 = 0.9$ and $\lambda_2 = -0.5$. By inspection, we have
$$E_{0.9} = N(A - 0.9I) = \operatorname{span}\left\{ \begin{bmatrix} 5 \\ 3 \end{bmatrix} \right\}$$
and
$$E_{-0.5} = N(A + 0.5I) = \operatorname{span}\left\{ \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right\}$$
Let
$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = c_1 \begin{bmatrix} 5 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

and so

$$\begin{aligned}
\begin{bmatrix} x_n \\ y_n \end{bmatrix} &= A^n \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \\
&= A^n \left(c_1 \begin{bmatrix} 5 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right) \\
&= c_1 A^n \begin{bmatrix} 5 \\ 3 \end{bmatrix} + c_2 A^n \begin{bmatrix} 3 \\ -1 \end{bmatrix} \\
&= c_1 (0.9)^n \begin{bmatrix} 5 \\ 3 \end{bmatrix} + c_2 (-0.5)^n \begin{bmatrix} 3 \\ -1 \end{bmatrix}
\end{aligned}$$

For large n, $0.9^n >> 0.5^n$ and so

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} \approx c_1 (0.9)^n \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

0.9 is called the dominant eigenvalue.
The asymptotic growth rate of the population is 0.9, which means that it is shrinking. The asymptotic proportion of children to adults is $5:3$.
We can change the birth rate from 1.5 to b to find where the population stabilizes, i.e., the asymptotic growth rate is 1.
Have $A = \begin{bmatrix} 0 & b \\ 0.3 & 0.4 \end{bmatrix}$ so

$$\begin{aligned}
\det(A - \lambda I) &= \det \begin{bmatrix} -\lambda & b \\ 0.3 & 0.4 - \lambda \end{bmatrix} \\
&= \lambda^2 - 0.4\lambda - 0.3b
\end{aligned}$$

We want $\lambda = 1$ to be an eigenvalue and therefore a root. Letting $\lambda = 1$ and setting equal to 0 yields

$$0 = 1 - 0.4 - 0.3b = 0.6 - 0.3b$$

and so $b = 2$. Setting $b = 2$, the characteristic polynominal is

$$\lambda^2 - 0.4\lambda - 0.6 = (\lambda - 1)(\lambda + 0.6)$$

and so the other polynominal is -0.6 so 1 is the dominant. Thus

$$E_1 = N(A - I) = \text{span} \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$$

and

$$E_{-0.6} = N(A + 0.6I) = \text{span} \left\{ \begin{bmatrix} 10 \\ -3 \end{bmatrix} \right\}$$

Thus

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = c_1 (1)^n \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 (-0.6)^n \begin{bmatrix} 10 \\ -3 \end{bmatrix}$$

and so as $n \to \infty$, the proportion of children to adults approaches $2:1$.

5 Geometry

Definition 5.1. The **norm** of $w = (x_1, \dots, x_n) \in \mathbb{R}^n$ is defined as

$$\|w\| = \sqrt{x_1^2 + \dots + x_n^2}$$

Corollary 5.2. The norm maps from $\mathbb{R}^n \to \mathbb{R}$, but it is not a linear map. In particular,

$$\begin{aligned} \|cw\| &= \sqrt{(cx_1)^2 + \dots + (cx_n)^2} \\ &= \sqrt{c^2(x_1^2 + \dots + x_n^2)} \\ &= |c|\sqrt{x_1^2 + \dots + x_n^2} \\ &= |c|\|w\| \end{aligned}$$

Definition 5.3. The **inner product** or **dot product** of $u = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $v = (y_1, \dots, y_n) \in \mathbb{R}^n$ is

$$u \cdot v = x_1y_1 + \dots + x_ny_n$$

This maps from $\mathbb{R}^n \to \mathbb{R}$.

Corollary 5.4.

$$u \cdot u = x_1^2 + \dots + x_n^2 = \|u\|^2$$

Theorem 5.5. Fix $v \in \mathbb{R}^n$. Then the map from $\mathbb{R}^n \to \mathbb{R}$ that sends u to $u \cdot v$ is linear.

Remark. Let $u = (a_1, a_2), v = (b_1, b_2)$ and $w = u - v = (a_1 - b_1, a_2 - b_2)$. Then

$$\begin{aligned} w \cdot w &= (a_1 - b_1)^2 + (a_2 - b_2)^2 \\ &= (a_1^2 + a_2^2) + (b_1^2 + b_2^2) - 2(a_1b_1 + a_2b_2) \\ &= u \cdot u + v \cdot v - 2(u \cdot v) \end{aligned}$$

Thus

$$\|w\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos\theta$$

Equating these yields

$$u \cdot v = \|u\|\|v\| \cos\theta$$

Definition 5.6. The angle θ between 2 vectors u, v in $\mathbb{R}^n$ is the value of θ between 0 and π that satisfies

$$\cos\theta = \frac{u \cdot v}{\|u\|\|v\|}$$

Definition 5.7. Vectors $u, v \in \mathbb{R}^n$ are **orthogonal** or perpendicular iff $u \cdot v = 0$.

Example. Find all vectors $(x, y, z) \in \mathbb{R}^3$ that are orthogonal to $(1, 2, 3)$.
We need $(x, y, z) \cdot (1, 2, 3) = 0$ and so $x + 2y + 3z = 0$. This is a plane in $\mathbb{R}^3$, let $y = s$ and $z = t$ and so $x = -2s - 3t$ and so the solution is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2s - 3t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

That is, the set of all vectors in $\mathbb{R}^3$ orthogonal to $(1, 2, 3)$ is the plane

$$\operatorname{span}\left\{\begin{bmatrix}-2\\1\\0\end{bmatrix}, \begin{bmatrix}-3\\0\\1\end{bmatrix}\right\}$$

Corollary 5.8. We can use matrix multiplication to represent the dot product. Let $u = \begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$ and $v = \begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}$. Then

$$u \cdot v = \begin{bmatrix}x_1 & \dots & x_n\end{bmatrix}\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix} = u^t v$$

We can also use the dot product to describe matrix multiplication. Let $A \in M_{mn}(\mathbb{R})$ and $B \in M_{nr}(\mathbb{R})$. Let $u_1, \dots, u_m$ be the rows of A and $v_1, \dots, v_r$ be the columns of B. Then

$$[AB]_{ij} = \sum_{k=1}^{n} [A]_{ik}[B]_{kj} = u_i \cdot v_k$$

Definition 5.9. Let $v = (z_1, \dots, z_n) \in \mathbb{C}^n$. Then

$$\|v\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}$$

where $|z_i|$ is the modulus defined by

$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2} = \sqrt{z\overline{z}}$$

Definition 5.10. Let $u = (w_1, \dots, w_n)$ and $v = (z_1, \dots, z_n)$ in $\mathbb{C}^n$. The inner product in $\mathbb{C}$ is defined as

$$\langle u, v\rangle = w_1\overline{z_1} + \cdots + w_n\overline{z_n}$$

Corollary 5.11. Not commutative.

$$\begin{aligned}\langle v, u\rangle &= z_1\overline{w_1} + \cdots + z_n\overline{w_n}\\ &= \overline{w_1\overline{z_1} + \cdots + w_n\overline{z_n}}\\ &= \overline{\langle u, v\rangle}\end{aligned}$$

5.1 Inner Product Spaces

Remark. We have looked at the Euclidean Inner Product on $\mathbb{R}^n$ and $\mathbb{C}^n$. Now, we generalize for any vector space.

Definition 5.12. Let V be a vector space over F with $F = \mathbb{R}$ or $F = \mathbb{C}$. An **inner product** on V is a function that maps an ordered pair of vectors (u, v) to a number $\langle u, v\rangle$ in F that satisfies the following:
For any $u, v, w \in V$ and $c \in F$, have

- $\langle u, v\rangle \geq 0$
- $\langle u, u\rangle = 0 \Leftrightarrow u = 0$
- $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$
- $\langle cu, v\rangle = c\langle u, v\rangle$
- $\langle u, v\rangle = \overline{\langle v, u\rangle}$

V and its inner product are called an inner product space.

Example. The **Euclidean inner product** on F^n is defined by

$$\langle (u_1, \dots, u_n), (v_1, \dots, v_n) \rangle = u_1\overline{v_1} + \cdots + u_n\overline{v_n} = \sum_{i=1}^{n} u_i\overline{v_i}$$

Let's verify this is an inner product. Let $u = (u_1, \dots, u_n)$ and $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots, w_n)$ be in F^n and $c \in F$.

1. We have

$$\langle u, u \rangle = \sum_{i=1}^{n} v_i\overline{v_i} = \sum_{i=1}^{n} |v_i|^2$$

Since $|v_i| \in \mathbb{R}$, $\langle v, v \rangle$ is the sums of squares of real numbers, which are non-negative.

2. We have

$$\begin{aligned} \langle v, v \rangle = 0 &\Leftrightarrow \sum_{i=1}^{n} |v_i|^2 = 0 \\ &\Leftrightarrow v_i = 0 && \text{(for each } i\text{)} \\ &\Leftrightarrow v = 0 \end{aligned}$$

3. We have

$$\begin{aligned} \langle u + v, w \rangle &= \sum_{i=1}^{n} (u_i + v_i)\overline{w_i} \\ &= \sum_{i=1}^{n} (u_i\overline{w_i} + v_i\overline{w_i}) \\ &= \sum_{i=1}^{n} (u_i\overline{w_i}) + \sum_{i=1}^{n} (v_i\overline{w_i}) \\ &= \langle u, w \rangle + \langle v, w \rangle \end{aligned}$$

4. We have

$$\begin{aligned} \langle cu, v \rangle &= \sum_{i=1}^{n} (cu_i)\overline{v_i} \\ &= c\sum_{i=1}^{n} u_i\overline{v_i} \\ &= c\langle u, v \rangle \end{aligned}$$

5. We have

$$\begin{aligned} \overline{\langle v, u \rangle} &= \overline{\sum_{i=1}^{n} v_i\overline{u_i}} \\ &= \sum_{i=1}^{n} \overline{v_i}u_i \\ &= \sum_{i=1}^{n} u_i\overline{v_i} \\ &= \langle u, v \rangle \end{aligned}$$

Definition 5.13. Let $c_1, \dots, c_n$ be fixed positive numbers. The **weighted Euclidean inner product** on F^n is defined by

$$\langle (u_1, \dots, u_n), (v_1, \dots, v_n) \rangle = c_1u_1\overline{v_1} + \cdots + c_nu_n\overline{v_n} = \sum_{i=1}^{n} c_iu_i\overline{v_i}$$

Example. Let V be the vector space of continuous real valued functions defined on $[-1, 1]$ over field of scalars $\mathbb{R}$. We define

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\ dx$$

We show that is is an inner product.

1. We have

$$\langle f, f\rangle = \int_{-1}^{1} (f(x))^2\ dx \geq 0$$

2. We have

$$\begin{aligned}\langle f, f\rangle = 0 &\Leftrightarrow \int_{-1}^{1} (f(x))^2\ dx = 0\\ &\Leftrightarrow f(x) = 0\end{aligned}$$

3. We have

$$\begin{aligned}\langle f + g, h\rangle &= \int_{-1}^{1} (f(x) + g(x))h(x)\ dx\\ &= \int_{-1}^{1} f(x)h(x) + g(x)h(x)\ dx\\ &= \int_{-1}^{1} f(x)h(x)\ dx + \int_{-1}^{1} g(x)h(x)\ dx\\ &= \langle f, h\rangle + \langle g, h\rangle\end{aligned}$$

4. We have

$$\begin{aligned}\langle cf, g\rangle &= \int_{-1}^{1} cf(x)g(x)\ dx\\ &= c\int_{-1}^{1} f(x)g(x)\ dx\\ &= c\langle f, g\rangle\end{aligned}$$

5. We have

$$\begin{aligned}\langle f, g\rangle &= \int_{-1}^{1} f(x)g(x)\ dx\\ &= \int_{-1}^{1} g(x)f(x)\ dx\\ &= \langle g, f\rangle\end{aligned}$$

Theorem 5.14. Let V be an inner product space. Then the following hold:

- Fix $v \in V$. the map that sends $u \in V$ to $\langle u, v\rangle \in F$ is linear.
- For any $u \in V$, $\langle u, 0\rangle = \langle 0, u\rangle = 0$.
- $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ for all $u, v, w \in V$.
- $\langle u, cv\rangle = \overline{c}\langle u, v\rangle$ for all $u, v \in V$ and $c \in F$.

Proof. 3. We have

$$\begin{aligned}\langle u, v + w\rangle &= \overline{\langle v + w, u\rangle}\\ &= \overline{\langle v, u\rangle + \langle w, u\rangle}\\ &= \langle u, v\rangle + \langle u, w\rangle\end{aligned}$$

4. We have

$$\begin{aligned}\langle u, cv\rangle &= \overline{\langle cv, u\rangle}\\ &= \bar{c}\overline{\langle v, u\rangle}\\ &= \bar{c}\langle u, v\rangle\end{aligned}$$

□

Definition 5.15. Let V be an inner product space. The **norm**$\|v\|$ of $v \in V$ is defined by

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

Note that since $\langle v, v\rangle \in \mathbb{R}$ with $\langle v, v\rangle \geq 0$ we have $\|v\| \in \mathbb{R}$ with $\|v\| \geq 0$.

Example. Consider F^n with the Euclidean inner product. Let $v = (v_1, \dots, v_n)$. Then

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

and so

$$\|v\| = \sqrt{|v_1|^2 + \cdots + |v_n|^2}$$

as before. If $F = \mathbb{R}$, we have

$$\|v\| = \sqrt{v_1^2 + \cdots + v_n^2}$$

For example, if $f(x) = x$, then

$$\|f\| = \sqrt{\int_{-1}^{1} x^2\, dx} = \sqrt{\frac{1}{3}x^3\Big|_{-1}^{1}} = \sqrt{\frac{2}{3}}$$

Theorem 5.16. Let V be an inner product space, and let $v \in V$. Then

- $\|v\| \geq 0$
- $\|v\| = 0 \Leftrightarrow v = 0$
- $\|cv\| = |c|\|v\|$

Theorem 5.17. Let V be an inner product space. If $v \neq 0$, then $\frac{v}{\|v\|}$ is a unit vector and a positive multiple of v.

Example. Find a unit vector scalar multiple of $(1, 2, 3)$. We have

$$\|(1, 2, 3)\| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$$

Thus the unit vector is

$$\frac{1}{\sqrt{14}}(1, 2, 3)$$

5.2 Orthogonality

Definition 5.18. Vectors u, v are an inner product space V are orthogonal if $\langle u, v\rangle = 0$.

Example. Show that $\sin(n\pi x)$ and $\cos(n\pi x)$ are orthogonal.

We have

$$
\begin{aligned}
\langle \sin(n\pi x), \cos(n\pi x)\rangle &= \int_{-1}^{1} \sin(n\pi x)\cos(n\pi x)\ dx \\
&= \frac{1}{2}\int_{-1}^{1} \sin(2n\pi x)\ dx \\
&= -\frac{1}{4n\pi}\cos(2n\pi x)\Big|_{-1}^{1} \\
&= -\frac{1}{4n\pi}(\cos(2n\pi) - cos(-2\pi)) \\
&= -\frac{1}{4n\pi}(1-1) \\
&= 0
\end{aligned}
$$

Theorem 5.19. Every $v \in V$ is orthogonal to 0. Furthermore, 0 is the only vector V that is orthogonal to itself.

Theorem 5.20. If u is orthogonal to v, then for any scalar c, u is orthogonal to cv.

Theorem 5.21. Suppose u, v are orthogonal. Then

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2$$

Proof. We have

$$
\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u, u+v\rangle + \langle v, u+v\rangle \\
&= \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\
&= \|u\|^2 + \|v\|^2
\end{aligned}
$$

□

Theorem 5.22. Let $u, v \in V$ with $v \neq 0$. Let $c = \frac{\langle u,v\rangle}{\langle v,v\rangle}$ and let $w = u - cv$. Then $u = cv + w$ and $\langle w, v\rangle = 0$.

Proof. Clearly we have $cv + w = cv + (u - cv) = u$. We have

$$
\begin{aligned}
\langle w, v\rangle &= \langle u - cv, v\rangle \\
&= \langle u,v\rangle - c\langle v,v\rangle \\
&= \langle u,v\rangle - \frac{\langle u,v\rangle}{\langle v,v\rangle}\langle v,v\rangle \\
&= 0
\end{aligned}
$$

□

Definition 5.23. The **projection** of u onto v is $\operatorname{proj}_v u = \frac{\langle u,v\rangle}{\langle v,v\rangle} v$

The Cauchy-Schwartz Inequality. Let $u, v \in V$. Then

$$|\langle u,v\rangle| \leq \|u\|\|v\|$$

Furthermore, equality holds iff u, v are scalar multiples of each other.

Proof. If $v = 0$, then $|\langle u, 0\rangle| = 0$ and $\|u\|\|0\| = 0$ so equality holds. Suppose $v \neq 0$. By previous theorem, we can write

$$u = \frac{\langle u, v\rangle}{\langle v, v\rangle} v + w$$

with v orthogonal to w and hence its scalar multiple. Then

$$\begin{aligned}
\|u\|^2 &= \left\| \frac{\langle u, v\rangle}{\langle v, v\rangle} v \right\|^2 + \|w\|^2 \\
&= \frac{|\langle u, v\rangle|^2}{(\langle v, v\rangle)^2} + \|w\|^2 \\
&= \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\|^2 \\
&\geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}
\end{aligned}$$

Thus $|\langle u, v\rangle|^2 \leq \|u\|^2 \|v\|^2$ and since norm is non-negative the result holds. We have equality iff $\|w\|^2 = 0 \Leftrightarrow w = 0$. Since $u = cv + w$, this holds iff u is a scalar multiple of v. □

The Triangle Inequality. Let $u, v \in V$. Then

$$\|u + v\| \leq \|u\| + \|v\|$$

with equality iff u and v are scalar multiples of each other.

Proof. We have

$$\begin{aligned}
\|u + v\|^2 &= \langle u + v, u + v\rangle \\
&= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
&= \langle u, u\rangle + \langle u, v\rangle + \overline{\langle u, v\rangle} + \langle v, v\rangle \\
&= \|u\|^2 + 2\operatorname{Re}\langle u, v\rangle + \|v\|^2 \\
&\leq \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \\
&\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\
&= (\|u\| + \|v\|)^2
\end{aligned}$$

In the 5th line, we used

$$|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2} \geq \sqrt{(\operatorname{Re} z)^2} = |\operatorname{Re} z| \geq \operatorname{Re} z$$

we have equality in the first inequality above if and only if $\langle u, v\rangle$ is real and non-negative. We then have equality in the second if and only if $\langle u, v\rangle = \|u\|\|v\|$.
Suppose $u = cv$ with $c \geq 0$. Then

$$\langle u, v\rangle = \langle cv, v\rangle = c\langle v, v\rangle = |c|\|v\|^2 = \|cv\|\|v\| = \|u\|\|v\|$$

Conversely, suppose we have equality. Then $\langle u, v\rangle = \|u\|\|v\|$. By Cauchy-Schwarz, one of u, v must be a scalar multiple of the other. Say $u = cv$. Then

$$\|u\|\|v\| = \langle cv, v\rangle = c\langle v, v\rangle$$

Since $\langle v, v\rangle \geq 0$, we have $c \geq 0$. □

Definition 5.24. Vectors $v_1, \ldots, v_m$ are **orthonormal** if each v_i is a unit vector and they are pairwise orthogonal. That is, if

$$\langle u, v\rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Example. The standard basis e_i is orthonormal.

Example. The following 3 vectors are orthonormal:

$$\begin{aligned} v_1 &= \frac{1}{\sqrt{3}}(1,1,1) \\ v_2 &= \frac{1}{\sqrt{2}}(-1,1,0) \\ v_3 &= \frac{1}{\sqrt{6}}(1,1,-2) \end{aligned}$$

Theorem 5.25. If $v_1, \dots, v_m$ are orthonormal, then for any $c_i \in F$ we have

$$\|c_1v_1 + \cdots + c_mv_m\|^2 = |c_1|^2 + \cdots + |c_m|^2$$

Proof. Note that c_iv_i is orthogonal to $\sum_j c_jv_j$ where the j run over a subset of $\{1, \dots, m\}$ that doesn't include i since

$$\langle \sum_j c_jv_j, c_iv_i\rangle = \sum_j c_j\overline{c_i}\langle v_j, v_i\rangle = 0$$

By the Pythagorean Theorem, we have

$$\begin{aligned} \|c_1v_1 + \cdots + c_mv_m\|^2 &= \|c_iv_i\|^2 + \|c_2v_2 + \cdots + c_mv_m\|^2 \\ &= |c_1|^2\|v_1\|^2 + \|c_2v_2 + \cdots + c_mv_m\|^2 \\ &= |c_1|^2 + \|c_2v_2 + \cdots + c_mv_m\|^2 \\ &= \dots \\ &= |c_1|^2 + |c_2|^2 + \cdots + |c_m|^2 \end{aligned}$$

□

Example. Let

$$\begin{aligned} v_1 &= \frac{1}{\sqrt{3}}(1,1,1) \\ v_2 &= \frac{1}{\sqrt{2}}(-1,1,0) \\ v_3 &= \frac{1}{\sqrt{6}}(1,1,-2) \end{aligned}$$

If $v = 17v_1 - 13v_2 - 5v_3$, then

$$\|v\|^2 = 17^2 + (-13)^2 + (-5)^2 = 483$$

and so $\|v\| = \sqrt{483}$

Theorem 5.26. If $v_1, \dots, v_m$ are orthonormal, they are linearly independent.

Proof. Suppose $c_1v_1 + \cdots + c_mv_m = 0$ for $c_i \in F$. By previous theorem, we have $|c_1|^2 + \cdots + |c_m|^2 = 0$, which implies $|c_i|^2 = 0$ for each i and so $c_i = 0$. □

Theorem 5.27. If $\dim V = n$ and $v_i, \dots, v_n$ are orthonormal, they are a basis of V.

Proof. By previous theorem, the v_i are linearly independent. Since there are n of them, they form a basis for V. □

Definition 5.28. If $v_1, \dots, v_n$ are orthonormal and a basis of V, then they are called an **orthonormal basis** of V.

Definition 5.29. The standard basis is an orthonormal basis of F^n.

Theorem 5.30. Let $v_1, \ldots, v_n$ be an orthonormal basis of V and let $w \in V$. Then

$$w = \langle w, v_1 \rangle v_1 + \cdots + \langle w, v_n \rangle v_n$$

and

$$\|w\|^2 = |\langle w, v_1 \rangle|^2 + \cdots + |\langle w, v_n \rangle|^2$$

Proof. The second follows from the first part by theorem on norm of a linear combination of orthonormal vectors. Since $v_1, \ldots, v_n$ is a basis of V, there exist $c_i \in F$ with

$$w = c_1 v_1 + \cdots + c_n v_n$$

We must show that $c_i = \langle w, v_i \rangle$. We have

$$\begin{aligned}
\langle w, v_i \rangle &= \langle c_1 v_1 + \cdots + c_n v_n, v_i \rangle \\
&= \langle c_1 v_1, v_i \rangle + \cdots + \langle c_n v_n, v_i \rangle \\
&= c_1 \langle v_1, v_i \rangle + \cdots + c_n \langle v_n, v_i \rangle \\
&= c_i
\end{aligned}$$

□

Example. Let

$$\begin{aligned}
v_1 &= \frac{1}{\sqrt{3}}(1, 1, 1) \\
v_2 &= \frac{1}{\sqrt{2}}(-1, 1, 0) \\
v_3 &= \frac{1}{\sqrt{6}}(1, 1, -2)
\end{aligned}$$

Write $w = (1, 2, 3)$ as a linear combination of the v_i.
The v_i form an orthonormal set of vectors. We have

$$\begin{aligned}
\langle w, v_1 \rangle &= \frac{1}{\sqrt{3}} \langle (1, 2, 3), (1, 1, 1) \rangle = \frac{6}{\sqrt{3}} \\
\langle w, v_2 \rangle &= \frac{1}{\sqrt{2}} \langle (1, 2, 3), (-1, 1, 0) \rangle = \frac{1}{\sqrt{2}} \\
\langle w, v_3 \rangle &= \frac{1}{\sqrt{6}} \langle (1, 2, 3), (1, 1, -2) \rangle = -\frac{3}{\sqrt{6}}
\end{aligned}$$

Thus $(1, 2, 3) = \frac{6}{\sqrt{3}} v_1 + \frac{1}{\sqrt{2}} v_2 - \frac{3}{\sqrt{6}} v_3$.

5.3 The Gram-Schmidt Algorithm

Theorem 5.31. Let $u_1, \ldots, u_m$ be linearly independent. Set $v_1 = u_1$, and for $2 \leq k \leq m$, set

$$\begin{aligned}
v_k &= u_k - \operatorname{proj}_{v_1} u_k - \cdots - \operatorname{proj}_{v_{k-1}} u_k \\
&= u_k - \sum_{j=1}^{k-1} \frac{\langle u_k, v_j \rangle}{\langle v_j, v_j \rangle} v_j
\end{aligned}$$

Now set $w_i = \frac{v_i}{\|v_i\|}$. Then $w_1, \ldots, w_m$ are orthonormal with

$$\operatorname{span}\{u_1, \ldots, u_m\} = \operatorname{span}\{w_1, \ldots, w_m\}$$

Example. Let $u_1 = (1,0,1,0)$, $u_2 = (1,1,1,1)$, $u_3 = (1,2,3,4)$.
Use the G-S Algorithm to find an orthonormal basis for $\text{span}\{u_1, u_2, u_3\}$.
Recall that $v_k = u_k - \sum_{j=1}^{k-1} \frac{\langle u_k, v_j \rangle}{\langle v_j, v_j \rangle} v_j$.
Then,

$$\begin{aligned} v_1 &= u_1 = (1,0,1,0) \\ v_2 &= u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 = (0,1,0,1) \\ v_3 &= u_3 - \frac{\langle u_3, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_3, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 = (-1,-1,1,1) \end{aligned}$$

Thus $w_1 = \frac{1}{\sqrt{2}}(1,0,1,0)$, $w_2 = \frac{1}{\sqrt{2}}(0,1,0,1)$, $w_3 = \frac{1}{2}(-1,-1,1,1)$ is an orthonormal for the span.

5.4 Orthogonal Matrices

Definition 5.32. A matrix $A \in M_{nn}(\mathbb{R})$ is called **orthogonal** if its columns form an orthonormal basis of $\mathbb{R}^n$.

Example. The $n \times n$ identity matrix is orthogonal.

Theorem 5.33. If A is orthogonal, then $A^{-1} = A^t$.

Theorem 5.34. A is orthogonal iff $A^{-1} = A^t$.

Theorem 5.35. If A is orthogonal, then $\det A = \pm 1$.

Theorem 5.36. If A is orthogonal, then so is A^{-1}. If A, B are both orthogonal, then so is AB.

Definition 5.37. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear. If $\|T(v)\| = v$ for all $v \in \mathbb{R}$, then T is called a **length preserving** linear map or **isometry**.

Theorem 5.38. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be linear and let A be the matrix for T w.r.t. to the standard basis so $T(v) = Av$. Then A is orthogonal iff T is an isometry.

Definition 5.39. If $A \in M_{nn}(\mathbb{R})$ is diagonalizable with real eigenvalues and if there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A, then A is **orthogonally diagonalizable**.

Theorem 5.40. If A is orthogonally diagonalizable, then A is symmetric.

The Spectral Theorem. A is orthogonally diagonalizable iff A is symmetric.

Example. Let

$$A = \begin{bmatrix} 1 & -2 & 2 \\ -2 & 4 & -4 \\ 2 & -4 & 4 \end{bmatrix}$$

Since A is symmetric, it is orthogonally diagonalizable. Find D and P with $A = PDP^t$.
We have $\det(A - \lambda I) = -\lambda^2(\lambda - 9)$. Thus the e. values are 0 and 9.
We have $E_9 = \text{span}\left\{ \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \right\}$. Since $\|(1,-2,2)\| = \sqrt{1^2 + (-2)^2 + 2^2} = 3$, an orthonormal basis of E_9 consists of $\frac{1}{3}(1,-2,2)$.

Similarly, $E_0 = N(A) = \text{span}\left\{\begin{bmatrix}2\\1\\0\end{bmatrix}, \begin{bmatrix}2\\0\\1\end{bmatrix}\right\}$.
We use G-S on the basis of E_0.

$$\begin{aligned} v_1 &= u_1 = (2,1,0) \\ v_2 &= u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 \\ &= \frac{1}{5}(2,-4,-5) \end{aligned}$$

Let's instead use $v_2 = (2,-4,5)$. Then $w_1 = \frac{1}{\sqrt{5}}(2,1,0)$ and $w_2 = \frac{1}{3\sqrt{5}}(2,-4,-5)$ form an orthonormal basis of E_0. Hence,

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & -\frac{5}{3\sqrt{5}} \end{bmatrix}$$

and

$$D = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Definition 5.41. For $A \in M_{nn}(\mathbb{C})$, the **conjugate transpose** of A is $A* = \overline{A^t}$. A is **unitary** iff $A^{-1} = A*$. (analogue of an orthogonal matrix). A is **Hermitian** iff $A* = A$.

Theorem 5.42. Let A be Hermitian. Then there exists unitary P and diagonal D such that $A = PDP^*$.

5.5 Projections onto a Subspace

Definition 5.43. We can project a vector onto a subspace. Let V be a subspace of U and let $b \in U$. We want $w \in V$ so that $\forall x \in V, \langle x, b-w\rangle = 0$. Such a w is called the **orthogonal projection** of b onto V, denoted $\text{proj}_V b$.

Theorem 5.44. Let $v_1, \ldots, v_k$ be an orthogonal basis of V. Let

$$w = \text{proj}_V b = \sum_{i=1}^{k} \frac{\langle b, v_i\rangle}{\langle v_i, v_i\rangle} v_i$$

Then $w \in V$ and for any $x \in V$, we have $\langle x, b-w\rangle = 0$.

Theorem 5.45. Let $w \in \text{proj}_V b$. Let $x \in V$. Then $\|b-x\| \geq \|b-w\|$.

5.6 Least Squares Method

Remark. Let A be $m \times n$ in $\mathbb{R}$. Consider the system $Ax = b$. The system has a solution $x \in \mathbb{R}^n$ iff $b \in \text{colspace}\, A = V$. What if there is no solution? By above a vector x for which Ax is closest to b is x with $Ax = \text{proj}_V b$. Note that such an x must exist since $\text{proj}_V b \in \text{colspace}\, A$. Once we obtain such an x, the error is $\|Ax - b\|$. Writing $[A]_{ij} = a_{ij}$ and $[x]_{i1} = x_i$ and $[b]_{i1} = b_i$, we have

$$\|Ax-b\|^2 = \sum_{k=1}^{m}([Ax]_{k1} - b_k)^2 = \sum_{k=1}^{m}\left(\sum_{\ell=1}^{n} a_{k\ell}x_\ell - b_k\right)^2$$

We are minimizing a sum of squares, so this is called the **least squares method**. Given system $Ax = b$ with no solution, how do we find x with $Ax = \text{proj}_V b$? We can find an orthogonal basis of colspace A with G-S, then compute the projection, then solve the system using row reduction. However, this is cumbersome and there is a better way.

Let $A = [a_1 \dots a_n]$ and let $V = \text{colspace}\, A$. We have

$$\begin{aligned} Ax = \text{proj}_V\, b &\Leftrightarrow \langle a_i, b - Ax\rangle = 0\ \forall i && (a_i \text{ form a basis of } v) \\ &\Leftrightarrow a_i^t(b - Ax) = 0\ \forall i \\ &\Leftrightarrow A^t(b - Ax) = 0 \\ &\Leftrightarrow A^tAx = A^tb \end{aligned}$$

Theorem 5.46. Given system $Ax = b$ with no solution, we solve $A^tAx = A^tb$. Then the error $\|Ax - b\|$ is minimized.

Example. We have experimental data:

$$(-2, 4), (-1, 2), (0, 1), (2, 1), (3, 1)$$

Find LoBF $y = a + bx$ for constants $a, b \in \mathbb{R}$. The system $Ax = b$ is

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

for this data, have

$$A = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, b = \begin{bmatrix} 4 \\ 2 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

We have

$$A^tA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 2 & 18 \end{bmatrix}$$

and

$$A^tb = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ -5 \end{bmatrix}$$

so

$$A^tAx = A^tb \implies \begin{bmatrix} 5 & 2 \\ 2 & 18 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \begin{bmatrix} 9 \\ -5 \end{bmatrix}$$

which has solution $(2, -\frac{1}{2})$. Thus the line of best fit is $y = 2 - \frac{x}{2}$.
The error is

$$\|Ax - b\| = \left\| \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -\frac{1}{2} \end{bmatrix} - \begin{bmatrix} 4 \\ 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -1 \\ \frac{1}{2} \\ 1 \\ 0 \\ -\frac{1}{2} \end{bmatrix} \right\| = \sqrt{\frac{5}{2}}$$

Example. Find a quadratic LoBF instead. Our system $Ax = b$ is

$$\begin{bmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 & \\ \vdots & y_n \end{bmatrix}$$

For this data,

$$A = \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}, b = \begin{bmatrix} 4 \\ 2 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

We have

$$A^t A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 2 & 3 \\ 4 & 1 & 0 & 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 18 \\ 2 & 18 & 26 \\ 18 & 26 & 114 \end{bmatrix}$$

and

$$A^t b = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 2 & 3 \\ 4 & 1 & 0 & 4 & 9 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 9 \\ -5 \\ 31 \end{bmatrix}$$

Then

$$\begin{bmatrix} 5 & 2 & 18 \\ 2 & 18 & 26 \\ 18 & 26 & 114 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 9 \\ -5 \\ 31 \end{bmatrix}$$

has solution $\left(\frac{86}{77}, \frac{-62}{77}, \frac{43}{154}\right)$ so the equation is $y = \frac{86}{77} - \frac{62}{77}x + \frac{43}{154}x^2$

Example. If doing a plane of best fit, $Ax = b$ is

$$\begin{bmatrix} 1 & x_1 & y_1 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$$

For example, if data is $(1, 0, 2), (1, 2, 5), (-1, 0, 2), (-2, -1, 2)$, then we have

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 0 \\ 1 & -2 & -1 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 5 \\ 2 \\ 2 \end{bmatrix}$$