8 Week Report

Shambhavi Singh
MATS638

1 Some basic Definitions

In this section we shall define a few of the basic structures used in the report

1.1 Field

A Field is a set along with 2 binary operations, + and *

1. **Closed under + and ***
   $\forall u, v \in\ F,\ (u+v)\ \in\ F\ and\ u*v\ \in\ F$

2. **Commutative**
   $\forall u, v\ \in\ F,\ u+v=v+u$

3. **Associative**
   $\forall u, v, w\ \in\ F,\ (u+v)+w=u+(v+w)$

4. **Additive and Multiplicative Identity**
   $\exists\ 0\ s.t\ u+0=u\ \forall\ u1 \in F$
   $\exists\ 1\ s.t\ u*1=u\ \forall\ u \in F$

5. **Additive and Multiplicative Inverse**
   $\forall u \in F\ \exists -u+\ \ and\ u^{-1} \in F\ s.t\ u+(-u)=0\ \&\ u*u^{-1}=1$

6. **Distributivity of multiplication over addition**
   $\forall u, v, w \in F\ u*(v+w)=u*v+u*w$

For the rest of the report, we shall take F to be either $\mathbb{R}$ or $\mathbb{C}$

1.2 Vector Space

A vector space is a set V of vectors over a field F along with vector addition $\oplus$ and scalar multiplication which satisfies the following properties

1. **Closed under $\oplus$ and scalar multiplication**
   $\forall u, v\ \in\ V\ \&\ a\ \in\ F(u \oplus av) \in V$

2. **Commutative**
   $\forall u, v\ \in\ V,\ u \oplus v=v \oplus u$

3. **Associative**
   $\forall u, v, w \ \in \ V, \ (u \oplus v) \oplus w = u \oplus (v \oplus w)$

4. **Additive Identity**
   $\exists \ 0 \ s.t \ u \oplus 0 = u \ \forall \ u \in V$
   We shall call this the **zero** of the vector space

5. **Additive Inverse**
   $\forall u \in V \ \exists -u$
   $in \ F \ s.t \ u \oplus (-u) = 0$

1.3 Subspace

A subspace is a subset of a vector space such that it's closed under vector addition and scalar multiplication, i.e. for $S \subseteq V$

$$(u + av) \ \in \ S \ \forall \ u, v \in S \ \& \ a \in F$$

1.4 Span

We say that a set of vectors $\{v_1, v_2, \ldots, v_n\}$ spans the vector space V when every element in V can be represented as the linear combination of the vectors in the given set, that is

$$\forall \ v \ \in V, \exists \ a_1, a_2, \ldots, a_n \ s.t$$
$$v = \sum_{i=1}^{n} a_i v_i$$
$$\text{For atleast one } a_i \neq 0$$

1

1.5 Linearly Independent Vectors

A set of vectors $\{v_1, v_2, \ldots, v_n\}$ is said to be linearly independent if

$$\sum_{i=1}^{n+1} a_i v_i = 0 \Leftrightarrow \ a_i = 0 \ \forall \ i = 1, \ldots, n$$

2 Linear Transformation Map

A function $T : V \to W$ is called a linear transformation map if

$$T(\alpha x + y) = \alpha T(x) + T(y) \forall x, y \in V \ and \ \alpha \in F$$

where V *and* W are the vector spaces over field F
Example:

$$V = \mathbb{R}^n,\ W = \mathbb{R}^m,\ \left[a_{ij}\right]_{mxn}$$
$$\text{Define} S : V \to W as$$
$$S(x) = S\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = Ax$$

S is a linear map as it clearly satisfies all the properties of a linear map
A linear map $T : \mathbb{R}^n \to \mathbb{R}$ is called a *linear functional*

2.1 Operators

An operator is a linear map from a vector space to itself

$$T : V \to V$$

The set of all operators on a vector space V shall be denoted as $\mathcal{L}(V)$

3 Inner Product Space

Definition: We define the inner product space as a vector space along with binary operator $\langle .,. \rangle$ such that it satisfies the following properties

$$\langle .,. \rangle : V \to F$$

1. **Positivity**
   $\langle v, v \rangle \geq 0 \ \forall \ v \ \in \ V$

2. **Definiteness**
   $\langle v, v \rangle = 0 \ \Leftrightarrow \ v = 0$

3. **Conjugate Symmetry**
   $\langle v, u \rangle = \overline{\langle u, v \rangle} \ \forall \ u, v \ \in \ V$

4. **Homogeneity and Additivity in the Second Slot**
   $\langle u, av + w \rangle = a\langle u, v \rangle + \langle u, w \rangle \ \forall \ u, v, w \ \in \ V$ and $a \ \in \ F$

5. **Conjugate Homogeneity in the First Slot**
   $\langle av, u \rangle = \overline{a}\langle u, v \rangle \ \forall \ u, v \ \in \ V$ and $\in F$

Note that several texts state the homogeneity and additivity to hold in the first slot, but we'll prefer this notation as it'll ease computation as we move on

A standard inner product space on F^n is the Standard Product Space or the **Euclidean Product Space** defined as

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

3.1 Norm

The norm of a vector is defined as follows

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

For example, for F^n, the norm with the standard inner product is given by

$$\|(x_1, x_2, \ldots, x_n)\| = \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2}$$

So, by the definition of inner product the norm of a non zero vector will always have a positive norm

3.2 Orthonormal Vectors and Projections

Definition: Two vectors are said to be orthonormal when $\langle u, v\rangle = 0$

Definition: A Projection is a linear operator such that $P^2 = P$. It has the following properties, where U and V are the range and kernel subspaces respectively:

- $P^2 = P$
- P is the identity operator on U, i.e. $\forall\ u\ \in\ U$
- $W = U \oplus V$

A projection is said to be orthogonal when $V = U^{\perp}$
Orthogonal Projections have special properties

- $P^* = P$
- $Q = I - P$ is also an orthonormal projection, where the range and kernel space are V and U respectively
  Here Q will have kernel and range as U and V respectively
- $W = U \oplus V$

Let the vector space have an orthonormal basis $(e_1, \ldots, e_n)$. Then the projection P such that $Px = \sum_{i=1}^{k} \langle e_i, x\rangle e_i$ is an orthonormal projection.
Here the range U will be the subspace formed by the orthonormal basis of $(e_1, \ldots, e_k)$ and the kernel V will be the subspace formed by $(e_{k+1}, \ldots, e_n)$
We'll show that the properties hold for it to be an orthonormal projection, we'll prove 2 of the properties, the others being true by definition

- $P^2 = P$

$$\begin{aligned}
P^2x = P(Px) =& P(\sum_{i=1}^{k}\langle e_i, x\rangle e_i) \\
=& \sum_{j=1}^{k}\left\langle e_j, \sum_{i=1}^{k}\langle e_i, x\rangle e_i \right\rangle e_j \\
=& \sum_{j=1}^{k}\sum_{i=1}^{k}\langle e_i, x\rangle\langle e_j, e_i\rangle e_j \\
=& \sum_{j=1}^{k}\sum_{i=1}^{k}\langle e_i, x\rangle\delta_{ij}e_j \\
=& \sum_{j=1}^{k}\sum_{i=1}^{k}\langle e_i, x\rangle\delta_{ij}e_i \\
=& \sum_{i=1}^{k}\langle e_i, x\rangle e_i \\
=& Px
\end{aligned}$$

- P*=P

$$\begin{aligned}
\langle y, Px\rangle =& \sum_{i=1}^{k}\langle e_i, x\rangle\langle y, e_i\rangle \\
\langle y, P^*x\rangle =& \langle Py, x\rangle \\
=& \overline{x, Py} \\
=& \overline{\sum_{i=1}^{k}\langle x, e_i\rangle\langle e_i, y\rangle} \\
=& \sum_{i=1}^{k}\overline{\langle x, e_i\rangle}\ \overline{\langle e_i, y\rangle} \\
=& \sum_{i=1}^{k}\langle e_i, x\rangle\langle y, e_i\rangle
\end{aligned}$$

  Since $\langle y, P^*x\rangle = \langle y, Px\rangle\ \forall\ x, y$ in the vector space, $P^* = P$

- $Q = I - P$ is also an orthogonal projection Clearly, $Qx = x - \sum_{i=1}^{k}\langle e_i, x\rangle e_i$
  $\Rightarrow\ Q = \sum_{i=k+1}^{n}\langle e_i, x\rangle e_i$
  We can see that the range and kernel space of Q are the kernel and range space of P, and since Q has a structure similar to P, the above claims hold for it as well

3.3 Some Important Results

In this section, we shall state and prove some of the important inequalities that involve inner product spaces

3.3.1 Pythagorean Theorem

If u, v are orthogonal vectors in V, then

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

Proof.

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

□

3.3.2 Cauchy Schwarz Inequality

If $u, v \in V$, then

$$|\langle u, v\rangle|^2 \leq \|u\|\|v\|$$

Proof. u can be decomposed into 2 vectors- one along v and the other normal to it

$$\begin{aligned}
u &= \frac{\langle u, v\rangle}{\|v\|^2} v + w \text{ where } \langle v, w\rangle = 0 \\
\Rightarrow \|u\|^2 &= \left\| \frac{\langle u, v\rangle}{\|v\|^2} v \right\|^2 + \|w\|^2 \text{ by Pythagorean Theorem} \\
&= \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\| \\
&\geq \frac{|\langle u, v\rangle|^2}{\|v\|^2} \\
&\Rightarrow \|u\|\|v\| \geq |\langle u, v\rangle|^2
\end{aligned}$$

□

3.3.3 Triangle Inequality

If $u, v \in V$, then

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

Proof.

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

□

4 Basis of a Vector Space

Definition: A set of vectors $\{v_1, v_2, ..., v_n\}$ is called a basis for a vector space if $\{v_1, v_2, , v_n\}$ are linearly independent and they span the whole vector space

Claim: Any 2 bases have the same number of elements
We'll not prove this claim here
Corollary: Max no. of LI vectors in the vector space are equal to those in a basis

Proof. Assume that this is not so
Then we can have a vector v_{n+1} s.t $\{v_1, v_2, ..., v_n + 1\}$ is a set of LI vectors, where $\{v_1, v_2, ..., v_n\}$ is the basis
Since the basis spans the whole space we have

$$v_{n+1} = \sum_{i=1}^{n} a_i v_i for\ some\ a_i \neq 0$$

Taking v_{n+1} to RHS we have

$$0 = \sum_{i=1}^{n+1} a_i v_i = 0,\ where a_{n+1} = -1$$

but the set of vectors was LI, which is contradicted by this equation
Hence we can only get n LI vectors □

Example Verify the following is a subspace of $\mathbb{R}^n$ and find a basis for it:

$$W = \{x \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i = 0\}$$

*Solution:*Clearly, $(0, \ldots, 0) \in W$
Also for all $x, y \in W$ $\sum_{i=1}^{n} x_i = 0 = \sum_{i=1}^{n} y_i$

$\Rightarrow \sum_{i=1}^{n} ax_i + y_i = 0$
$\Rightarrow ax + y \in W$
So W is a subspace of $\mathbb{R}^n$
For a basis we can have $\left(\frac{1}{\sqrt{2}}, 0, \ldots, 0, \frac{-1}{\sqrt{2}}\right)\left(0, \frac{1}{\sqrt{2}}, 0, \ldots, 0, \frac{-1}{\sqrt{2}}\right)\ldots\left(0, \ldots, 0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right)$

4.1 Dimension of a Vector Space

Using the property of all bases for a vecor space having the same number of vectors and that number being the same as the maximum number of LI vectors one can have in that space, we'll define the dimension of a vector space

Definition: The dimension of a vector space is the number of vectors in the basis of that vector space

4.2 Orthonormal Basis

We now will define a special type of basis where all the vectors are orthonormal to each other
Definition: A basis $\{v_1, v_2, ..., v_n\}$ of a vector space V is said to be orthogonal when

$$\langle v_i, v_j \rangle = 0 \ \forall \ i \neq j$$

So we can say that in an orthonormal basis There are several advantages of having an orthogonal basis for a vector space one of them being the ease of representation of a vector in the space using such a basis as for any $v \in V$
First let us fix an orthonormal basis $\{v_1, v_2, ..., v_n\}$ s.t

$$\langle v_i, v_k \rangle = \delta_{ij}$$

where

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

so for any vector in V

$$v = \sum_{i=1}^{n} a_i v_i for\ atleast\ one\ a_i \neq 0$$

Now

$$\langle v_k, v \rangle = \sum_{i=1}^{n} ai \langle v_i, v_k \rangle = a_i$$

So we can say that

$$v = \sum_{i=1}^{n} ai \langle v_i, v_k \rangle = a_i v$$

4.2.1 Gram Schmidt Orthogonalization

Since an orthonormal basis makes vector representation easier, it'd be convenient to learn a method of how to come up with an orthonormal basis Gram Schmidt Orthogonalization helps us in generating one from a given basis for the vector space
Suppose we have the basis $\{u_1, u_2, \dots, u_n\}$
Now take

$$\begin{aligned}
v_1 &= u_1 \\
v_2 &= u_2 - \frac{\langle v_1, u_2 \rangle}{\langle v_1, v_1 \rangle} v_1 \\
&\vdots \\
v_n &= u_n - \sum_{i=1}^{n-1} \frac{\langle v_i, u_n \rangle}{\langle v_i, v_i \rangle} v_i
\end{aligned}$$

5 Spectral Theorem

$$C = \begin{bmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \omega_n \end{bmatrix}$$

5.1 Schur's Theorem

First we'll prove a few propositions, which are required for proving Schur's theorem
Proposition 1: $T \in \mathcal{L}(V)$ and $(v_1, v_2, \dots, v_n)$ is a basis. The following are equivalent

1. Matrix of T w.r.t basis is an upper triangular matrix
2. $Tv_k \in span(v_1, v_2, \dots, v_n)$
3. $span(v_1, v_2, \dots, v_n)$ is invariant under T for each k=1,...,n

Proof. $1 \Leftrightarrow 2$ By definition
$2 \Rightarrow 3$ By definition
$3 \Rightarrow 2$ Prove Properly!

□

Proposition 2: Suppose V is a complex vector space and $T \in \mathcal{L}(V)$, then T has an upper triangular matrix w.r.t some basis of V

Proof. The proof is by induction on the dimension of the vector space
For n=1 Since an operator of dimension 1 is already upper triangular, it's true for this case
Assuming the claim to be true for an n-dimensional matrix, we prove it's true for n+1 too So given an eigenvalue of the matrix λ the matrix T, there's a vector space U s.t
$U = Range(T - \lambda I)$
$Range(T - \lambda)$ doesn't span the whole of V, thus $dim(U) < dim(V)$
Also, we see that U is invariant under T as for any $u \in U$
$Tu = (T - \lambda I)u + \lambda u \in U$ Thus $T_{|U} \in \mathcal{L}(U)$
by induction hypothesis, there's a basis of U wrt which $T|_U$ has an upper triangular matrix. Let's assume the basis is $u_1, \ldots, u_m$. Then
$(T|_u)(u_j) \in span(u_1, \ldots, u_j)$
Extending this to a basis of V, say $(u_1, \ldots, u_m, v_{m+1}, \ldots, v_{n+1})$
For each j we havw $Tv_k = (T - \lambda I)v_k + \lambda v_k$ By definition of U, $(T - \lambda I)v_k \in U$ which shows that

$$Tv_k \in span(u_1, \ldots, u_m, v_{m+1} \ldots, v_{n+1})$$

T has an upper triangular matrix wrt the basis $(u_1, \ldots, u_m, v_{m+1}, \ldots, v_{n+1})$ □

Proposition 3: Suppose $T \in \mathcal{L}(V)$. If T has a an upper triangular matrix w.r.t some basis of V, then T has an upper triangular w.r.t some orthonormal basis of V

Proof. $(\Leftarrow)$
$(\Rightarrow)T$ has an upper triangular matrix w.r.t some basis of V $(v_1, \ldots, v_n)$
By the Proposition 1, $span(v_1, v_2, \ldots, v_k)$ is invariant under T for each k=1,...,n thus we can find an orthinormal basis by using gram schmidt orthogonalization to build an orthonormal basis $(e_1, \ldots, e_n)$ such that $span(e_1, e_2, \ldots, e_k)$ is invariant under T as e_k will only involve terms till v_k. Thus again by Proposition 1, we have that T will have an upper triangular matrix wrt $(e_1, \ldots, e_n)$ □

Schur Theorem: Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ Then T has an upper triangular matrix w.r.t some orthonormal basis
The proof immediately follows from Propositions 2 and 3

5.2 Spectral Theorem for Normal Complex Matrices

Theorem: Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. then V has an orthonormal basis consisting of eigenvectors of T iff T is normal, i.e. $TT^* = T^*T$

Proof. $(\Rightarrow)$ Suppose V has an orthonormal basis consisting of eigenvectors of T. Then T has a diagonal with respect to this basis. Thus $T*$ will also have a diagonal matrix with respect to this basis. Since diagonal matrices commute, T and T^*

($\Leftarrow$)Suppose T is normal. We have an orthonormal basis, say $(v_1, \ldots, v_n)$ w.r.t which T has an upper triangular matrix

$$T[v_1 \ v_2 \ldots v_n] = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{n1} \end{bmatrix}$$

Claim: T is actually a diagonal matrix w.r.t this basis

$$\begin{aligned} \|Tv_1\| =& \langle Tv_1, Tv_1 \rangle = |a_{11}|^2 \\ & \langle Tv_1, Tv_1 \rangle = \langle T^* v_1, T^* v_1 \rangle = |a_{11}|^2 + \ldots + |a_{1n}|^2 \end{aligned}$$

Since T is normal ,$\|Tv_1\| = \|T^* v_1\|$, we have $a_{1i} = 0 \ \forall i \neq 1$

$$\begin{aligned} \|Tv_2\| =& \langle Tv_2, Tv_2 \rangle = |a_{22}|^2 \\ & \langle Tv_1, Tv_1 \rangle = \langle T^* v_1, T^* v_1 \rangle = |a_{22}|^2 + \ldots + |a_{2n}|^2 \end{aligned}$$

we have $a_{2i} = 0 \ \forall i \neq 2$
Continuing in a similar manner, we get $a_{ij} = 0, i \neq j$ hence giving a diagonal matrix $\square$

It is observed that any matrix which has distinct eigenvalues is diagonalizable, where the diagonal matrix consists of the eigenvalues of A, one can see that through the following.
Take a matrix A,suppose U is an invertible matrix with columns $u_1, \ldots, u_n$ and D the diagonal matrix with diagonal entries as eigenvvalues of A

$$\begin{aligned} A &= UDU^{-1} \\ AU &= UD \\ A[u_1 \ldots u_n] &= [u_1 \ldots u_n] \begin{bmatrix} \gamma_1 & 0 & \ldots & 0 \\ 0 & \gamma_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \gamma_n \end{bmatrix} \\ A[u_1 \ldots u_n] &= [\gamma_1 u_1 \ldots \gamma_n u_n] \\ \Rightarrow Au_k &= \gamma u_k \ \forall \ k = 1, \ldots, n \end{aligned}$$

Hence, we can see that each u_k is the eigenvector corresponding to γ_k

6 Jordan Normal Form

It's seen that one can obtain a diagonal form for matrices with unique eigenvalues.

However for matrices with repeated eigenvalues, we have the jordan normal form where the matrix is divided into Jordan blocks which are of the form

$$A = \begin{bmatrix} \gamma_1 & 1 & 0 & 0 \\ 0 & \gamma_1 & \ddots & 0 \\ \vdots & & \ddots & 1 \\ 0 & \dots & 0 & \gamma_1 \end{bmatrix}$$

Where the size of a block is determined by the multiplicity of the repeated eigenvalue So we end up getting a matrix of the following form

$$\begin{bmatrix} \gamma_1 & & & \\ & \begin{bmatrix} \gamma_2 & 1 & \\ & \gamma_2 & 1 \\ & & \gamma_2 \end{bmatrix} & & \\ & & \ddots & \\ & & & \begin{bmatrix} \gamma_k & 1 \\ & \gamma_k \end{bmatrix} \end{bmatrix}$$

The proof follows in a similar way as for the unique eigenvalue case Where the eigenvectors (denoted by u) for the repeated cases are calculated by the previous eigenvector, i.e

$$(A - \gamma_r I)u_{r_k} = u_{r_{k-1}}$$

Computing the powers of a matrix in Jordan form is easy

$$A^n = \begin{bmatrix} \gamma_1^n & \binom{n}{1}\gamma^{n-1} & \dots & \binom{n}{k}\gamma^{n-k} \\ & \gamma^n & \dots & \binom{n}{k-1}\gamma^{n-k+1} \\ & & \ddots & \vdots \\ & & & \gamma^n \end{bmatrix}$$

7 Application of diagonalization: Recurrence Relations

7.1 2×2 case

Special CASE: Fibonacci Series

We know that the fibonacci series is given by:

$$R_{n+1} = R_n + R_{n-1}$$

To compute the n+1th term in terms of R_0 and R_1
We can express it through a matrix

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} R_1 \\ R_0 \end{bmatrix} = \begin{bmatrix} R_{n+1} \\ R_n \end{bmatrix}$$

This is of the form $A^n R = X$
The eigenvalues are $\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_1 = \frac{1-\sqrt{5}}{2}$
So $A^n = U \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}^n U^{-1}$
So we can say that the nth term will be some linear combination of the power of the 2 eigenvalues

$$
\begin{aligned}
R_n &= a\left(\frac{1+\sqrt{5}}{2}\right)^n + b\left(\frac{1-\sqrt{5}}{2}\right)^n \\
R_0 &= a + b = 0 \\
&\Rightarrow a = -b \\
R_1 &= a\left(\frac{1+\sqrt{5}}{2}\right) + b\left(\frac{1-\sqrt{5}}{2}\right) = 1 \\
&\Rightarrow a = \frac{1}{\sqrt{5}}, b = -\frac{1}{\sqrt{5}} \\
\therefore R_n &= \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right] \\
&= \frac{1}{\sqrt{5}} \sum_{i=1,odd}^{n} \left[\binom{n}{i}\frac{\sqrt{5}^i}{2^n}\right]
\end{aligned}
$$

CASE I: $\lambda_1 \neq \lambda_2$

$$R_{n+1} = cR_n + dR_{n-1}$$

We can express it through a matrix

$$\begin{bmatrix} c & d \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} R_1 \\ R_0 \end{bmatrix} = \begin{bmatrix} R_{n+1} \\ R_n \end{bmatrix}$$

This is of the form $A^n R = X$
The eigenvalues are $\lambda_1 = \frac{c+\sqrt{c^2+4d}}{2}, \lambda_1 = \frac{c-\sqrt{c^2+4d}}{2}$ Here the matrix can easily diagonalized ($A = UDU^{-1}$) where U is the matrix containing its eigenvectors and D is a diagonal matrix comprising the eigenvalues

$$A = UDU^{-1} = U \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} U^{-1}$$

Hence $A^n = UD^nU^{-1} = U \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} U^{-1}$
So we can say that the nth term will be some linear combination of the power

of the 2 eigenvalues

$$R_n = a\lambda_1^n + b\lambda_2^n$$
$$R_0 = a + b$$
$$R_1 = a\lambda_1 + b\lambda_2$$

Therefore

$$a = \frac{R_1 - \lambda_2 R_0}{\lambda_1 - \lambda_2}, b = \frac{R_1 - \lambda_1 R_0}{\lambda_2 - \lambda_1}$$
$$\therefore R_n = \frac{1}{\lambda_1 - \lambda_2}\left[(R_1 - \lambda_2 R_0)\lambda_1^n - (R_1 - \lambda_1 R_0)\lambda_2^n\right]$$
$$R_n = (\lambda_2^n - \lambda_1^n)\left[R_0 - \frac{R_1}{\lambda_1 - \lambda_2}\right]$$

CASE II: $\lambda_1 = \lambda_2$

For the eigenvalues to be equal, we can see from the previous case that $d = -\frac{c^2}{4}$
Hence we have

$$R_{n+1} = cR_n - \frac{c^2}{4}R_{n-1}$$

, giving the matrix representation

$$\begin{bmatrix} c & \frac{-c^2}{4} \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} R_1 \\ R_0 \end{bmatrix} = \begin{bmatrix} R_{n+1} \\ R_n \end{bmatrix}$$

On computation we found

$$A^n = \begin{bmatrix} (n+1)\frac{c^n}{2^n} & -n\frac{c^{n+1}}{2^{n+1}} \\ (n)\frac{c^{n-1}}{2^{n-1}} & -(n-1)\frac{c^n}{2^n} \end{bmatrix}$$

Since we know $\lambda = c/2$, the matrix can be re-written as

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

Hence, $R_n = \lambda^{n-1}[nR_1 - (n-1)\lambda R_0] = n\lambda^{n-1}R_1 - (n-1)\lambda^n R_0$
Another way of finding the n^{th} term we use the jordan canonical form, thus

$$A = UDU^{-1} = U\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}U^{-1}$$

$$\Rightarrow A^n = UD^nU^{-1} = U\begin{bmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{bmatrix}U^{-1}$$

So we can find the nth term as a general combination of the $\lambda^n, n\lambda^{n-1}$

$$\begin{aligned}
R_n &= a\lambda^n + bn\lambda^{n-1} \\
R_0 &= a \\
R_1 &= a\lambda + b \\
\Rightarrow b &= R_1 - R_0\lambda \\
\Rightarrow R_n &= \lambda^n R_0 + n(R_1 - R_0\lambda)\lambda^{n-1} \\
&= n\lambda^{n-1}R_1 - (n-1)\lambda^n R_0
\end{aligned}$$

7.2 3×3 case

We'll consider the 3 term case, that is

$$R_{n+2} = aR_{n+1} + bR_n + cR_{n-1}$$

$$\Rightarrow \begin{bmatrix} a & b & c \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}^n \begin{bmatrix} R_2 \\ R_1 \\ R_0 \end{bmatrix} = \begin{bmatrix} R_{n+2} \\ R_{n+1} \\ R_n \end{bmatrix}$$

CASE I: $\lambda_1 \neq \lambda_2 \neq \lambda_3$

Here we can diagonalize the matrix using spectral theorem hence rewriting the nth form matrix as

$$R_{n+2} = aR_{n+1} + bR_n + cR_{n-1}$$

$$U \begin{bmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{bmatrix} U^{-1} \begin{bmatrix} R_2 \\ R_1 \\ R_0 \end{bmatrix} = \begin{bmatrix} R_{n+2} \\ R_{n+1} \\ R_n \end{bmatrix}$$

$$\begin{aligned}
R_n &= \alpha\lambda_1^n + \beta\lambda_2^n + \gamma\lambda_3^n \\
\Rightarrow R_0 &= \alpha + \beta + \gamma \\
R_1 &= \alpha\lambda_1 + \beta\lambda_2 + \gamma\lambda_3 \\
R_2 &= \alpha\lambda_1^2 + \beta\lambda_2^2 + \gamma\lambda_3^2
\end{aligned}$$

We have 3 equations for the three unknown coefficients s.t.

$$\begin{bmatrix} \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} R_2 \\ R_1 \\ R_0 \end{bmatrix}$$

which gives us a solvable equation set as determinant is non-zero

Tribonacci Case

$$R_{n+2} = R_{n+1} + R_n + R_{n-1}$$
$$R_{n+2} = aR_{n+1} + bR_n + cR_{n-1}$$
$$\Rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}^n \begin{bmatrix} R_2 \\ R_1 \\ R_0 \end{bmatrix} = \begin{bmatrix} R_{n+2} \\ R_{n+1} \\ R_n \end{bmatrix}$$

CASE II: $\lambda_1 = \lambda_2 \neq \lambda_3$

$$\begin{aligned} \Rightarrow & x^3 - ax^2 - bx - c = (x-\lambda_1)^2(x-\lambda_2) \\ \Rightarrow & a = 2\lambda_1 + \lambda_2 \\ & b = -2\lambda_1\lambda_2 - \lambda_1^2 \\ & c = \lambda_1^2\lambda_2 \end{aligned}$$

This will give us a matrix with 2 jordan blocks

$$U \begin{bmatrix} \lambda_1^n & n\lambda_1^{n-1} & 0 \\ 0 & \lambda_1^n & 0 \\ 0 & 0 & \lambda_2^n \end{bmatrix} U^{-1} \begin{bmatrix} R_2 \\ R_1 \\ R_0 \end{bmatrix} = \begin{bmatrix} R_{n+2} \\ R_{n+1} \\ R_n \end{bmatrix}$$

$$\begin{aligned} & R_n = \alpha\lambda_1^n + \beta n\lambda_1^{n-1} + \gamma\lambda_2^n \\ \Rightarrow & R_0 = \alpha + \gamma \\ & R_1 = \alpha\lambda_1 + \beta + \gamma\lambda_2 \\ & R_2 = \alpha\lambda_1^2 + 2\beta\lambda_1 + \gamma\lambda_2^2 \end{aligned}$$

We have 3 equations for the three unknown coefficients s.t.

$$\begin{bmatrix} \lambda_1^2 & 2\lambda_1 & \lambda_2^2 \\ \lambda_1 & 1 & \lambda_2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} R_2 \\ R_1 \\ R_0 \end{bmatrix}$$

CASE III: $\lambda_1 = \lambda_2 = \lambda_3$

For this case we need to have a perfect cube characteristic equation

$$\begin{aligned} \Rightarrow & x^3 - ax^2 - bx - c = (x-\lambda)^3 \\ \Rightarrow & a = 3\lambda \\ & b = -3\lambda^2 \\ & c = \lambda^3 \end{aligned}$$

Using the Jordan Canonical form

$$\begin{bmatrix} \lambda^n & n\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} \\ 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{bmatrix}$$

So we can say that R_n will have the following form

$$
\begin{aligned}
& R_n = \alpha\lambda^n + \beta n\lambda^{n-1} + \gamma\binom{n}{2}\lambda^{n-2} \\
\Rightarrow\ & R_0 = \alpha & \alpha = R_0 \\
& R_1 = \alpha\lambda + \beta & \beta = R_1 - \lambda R_0 \\
& R_2 = \alpha\lambda^2 + 2\beta\lambda + \gamma & \gamma = \lambda^2 R_0 - 2\lambda R_1 + R_2 \\
& R_n = \frac{(n-1)(n-2)}{2}\lambda^n R_0 - n(n-2)\lambda^{n-1}R_1 + \frac{n(n-1)}{2}\lambda^{n-2}R_2
\end{aligned}
$$

7.3 Special n×n Case

Recurrance Relations can involve more than 2 terms. They'll have the general form

$$R_{n+r+1} = \sum_{i=0}^{r} \alpha_i R_{n+i}$$

and can be represented by the following matrix

$$\begin{bmatrix} \alpha_r & \dots & \alpha_0 \\ \begin{bmatrix} & & \\ & I_{r\times r} & \\ & & \end{bmatrix} & & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \end{bmatrix}^n \begin{bmatrix} R_r \\ R_{r-1} \\ \vdots \\ R_0 \end{bmatrix} = \begin{bmatrix} R_{n+r} \\ R_{n+r-1} \\ \vdots \\ R_n \end{bmatrix}$$

Here we consider the case where all eigenvalues in a general k+1 term Recurrance Relation are equal.
Here the coefficients, α_i will have a special form
Since $x^{k+1} - \sum_{i=0}^{k} \alpha_i x^i = (x-\lambda)^{k+1}$

$$R_{n+k+1} = \sum_{i=0}^{k} (-1)^i \binom{k+1}{k+1-i} \lambda^i R_{n+i}$$

There, we get
$R_n = \left[\sum_{i=0}^{k}(-1)^i\binom{n}{i}\right] R_0 + \left[\sum_{i=1}^{k}(-1)^i\binom{i}{1}\binom{n}{i}\right] R_1 + \dots$

$$R_n = \sum_{0\leqslant j\leqslant i\leqslant k}\sum (-1)^i \binom{i}{j}\binom{n}{i} R_j$$

8 Based on the paper

8.1 Nilpotent Matrix Properties

We have a nilpotent matrix C with nilpotency k such that:

$$
\begin{aligned}
&C : \mathbb{C}^n \to \mathbb{C}^n \\
&\gamma_1 = dim(Ker(C)) \\
&\gamma_2 = dim(Ker(C^2) \cap Ker(C^{\perp})) \\
&\vdots \\
&\gamma_k = dim(Ker(C^k) \cap Ker((C^{k-1})^{\perp}))
\end{aligned}
$$

Clearly, $\gamma_1 + \gamma_2 + \ldots + \gamma_k = n$
Claim: $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_k \geq 1$

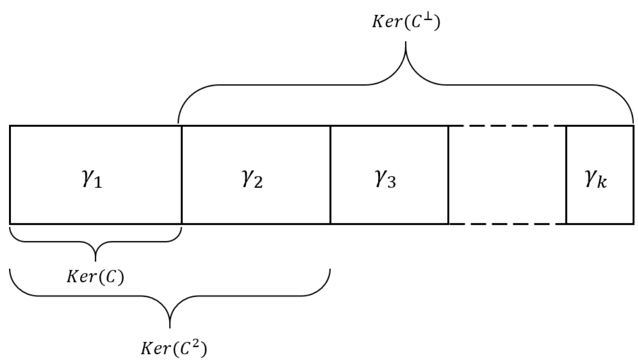

Figure 1: Visualization of the sets

Proof. Take a basis for γ_2 say $(y_1, y_2, \ldots, y_{\gamma_2})$

$$
\begin{aligned}
(y_1, y_2, \ldots, y_{\gamma_2}) &\in Ker(C^2) \\
\therefore\ (Cy_1, \ldots, Cy_{\gamma_2}) &\in Ker(C) \\
\Rightarrow\ dim(Cy_1, \ldots, Cy_{\gamma_2}) &\leq \gamma_1
\end{aligned}
$$

$$
\begin{aligned}
\text{Since } (Cy_1, \ldots, Cy_{\gamma_2}) \text{ are LI in } Ker(C) \text{as } &\sum_{i=1}^{\gamma_2} a_i C y_i = 0 \\
&\Rightarrow \sum_{i=1}^{\gamma_2} C(a_i y_i) = 0 \\
&\Leftrightarrow a_i = 0\ \forall\ \imath = 1, \ldots, n
\end{aligned}
$$

$$
\Rightarrow\ dim(Cy_1, \ldots, Cy_{\gamma_2}) = \gamma_2
$$

$\Rightarrow \gamma_2 \leq \gamma_1$
We can prove this for other dimensions in a similar way

Now assume $\gamma_r < 1$ *i.e.* $\gamma_r = \ldots = \gamma_k = 0$
This would mean $C^r = 0$ but this is not so
Thus all $\gamma_i \geq 1$ $\square$

Now we divide C into block matrices, s.t

$$C = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$$

$$C^k = \begin{bmatrix} P^k & \square \\ \square & S^k \end{bmatrix}$$

$\because C^k = 0$
$\Rightarrow P^k = S^k = 0$
Thus P and Q are nilpotent matrices with nilpotency less than or equal to k
So we

8.2 PAPER

Given a matrix C s.t

$$[c_{ij}] = \begin{cases} \frac{1}{\omega_i - \omega_j}, & \text{if } i \neq j \\ \sum\limits_{\substack{k=1 \\ k \neq i}}^{n}, & \text{if } i = j \end{cases}$$

From the definition, it's not easy to see that this matrix would be nilpotent. To make it easier to see, we shall associate a map from the vector space of Polynomials (denoted by $\mathcal{P}$) of degree less than n.

$$T : \mathcal{P} \to \mathbb{C}^n$$

Take $g(x) \in \mathcal{P}$
Claim: T is surjective

Proof. We divide the proof into 2 parts
`Onto:` Since for any given vector $[\omega_1, \omega_2, \ldots, \omega_n]^t$ we'll have some polynomial function which takes the value ω_i at some point in $x_i \in \mathbb{C}$
`One-One:` Assume $g_1(\omega_i) = g_2(\omega_i)\ \forall\ i = 1, \ldots, n$
Where $g_1(x) = \sum_{i=0}^{n-1} a_i x^i,\ g_2(x) = \sum_{i=0}^{n-1} b_i x^i$
$\therefore (g_1 - g_2)(\omega_i) = 0\ \forall\ i = 1, \ldots, n$
Since g_1 and g_2 are polynomials of degree less than n and we have n linearly

independent equations

$$\begin{bmatrix} 1 & \omega_1 & \omega_1^2 & \dots & \omega_1^n \\ 1 & \omega_2 & \omega_2^2 & & \omega_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega_n & \omega_n^2 & \dots & \omega_n^n \end{bmatrix} \begin{bmatrix} a_0 - b_0 \\ a_1 - b_1 \\ \vdots \\ a_{n-1} - b_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

This equation is of the form $AX = 0$, which has a non-trivial solution iff $|A| = 0$
But for that to happen, $\omega_i = \omega_j$ for some $i \neq j$, which isn't true in all cases of $\mathbb{C}^n$
Thus the map is one-one

□

So we have a map $T : \mathcal{P} \to \mathbb{C}^n$ $g(x)$ with the vector as $[g(\omega_1), g(\omega_2), \dots, g(\omega_n)]^t$
The matrix C^t gets identified with the operator $\widetilde{C}$ on polynomials where

$$\widetilde{C}g(x) = \sum_{i=1}^{n} \frac{g(x) - g(\omega_i)}{x - \omega_i} - g'(x)$$

Claim: $T\widetilde{C}g(x) = CTg(x)$

Proof.

$$g(x) \in \mathcal{P},\ g(x) = \sum_{i=0}^{n-1} a_i x^i$$

$$\text{Let } C^t T g(x) = [c(\omega_1), \dots, c(\omega_n)]^t$$

$$c(\omega_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{g(\omega_j)}{\omega_j - \omega_i} + \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{g(\omega_i)}{\omega_i - \omega_k}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{g(\omega_i) - g(\omega_j)}{\omega_i - \omega_j}$$

$$\text{Let } T\widetilde{C}g(x) = [c'_1, \dots, c'_2]^t$$

$$c'(\omega_i) = \sum_{j=1}^{n} \frac{g(x_i) - g(\omega_i)}{\omega_i - x_j} - g'(x_i)$$

$$\text{So for } x = (\omega_1, \dots, \omega_n)^t$$

$$c'(\omega_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{g(\omega_i) - g(\omega_i)}{\omega_i - \omega_j} + g'(\omega_i) - g'(\omega_i)$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{g(\omega_i)}{\omega_i - \omega_j}$$

$$= c(\omega_i)$$

$c(\omega_i) = c'(\omega_i) \ \forall \ i = 2, \dots, n$ $\square$

It can be seen the $\widetilde{C}$ is degree reducing, hence $\widetilde{C}$ would become zero after sometime hence implying $\widetilde{C}$ (and thus C) is nilpotent

So now, we find $ker(\widetilde{C})$

$$\Rightarrow \sum_{i=1}^{n} \frac{g(x) - g(\omega_i)}{x - \omega_i} - g'(x) = 0$$

It's clear to see that only a constant polynomial will belong to the kernel
So

$$\begin{aligned} Ker(\widetilde{C}) &= span(1) \\ &\Rightarrow \gamma_1 = 1 \\ Ker(\widetilde{C}^2) &= span(1, x) \\ &\Rightarrow \gamma_2 = 1 \end{aligned}$$

So we have

$$\gamma_i = 1 \ \forall \ i = 1, \dots, n$$

Since $\widetilde{C}$ can be associated with C, its kernels will show similar behaviour

Acknowledgements

I'd like to wholeheartedly thank everyone who helped me throughout the period of my intern
Professor B.V Rajarama Bhat, who guided me throughout the period, correcting errors in the results and explaining the concepts that were not clear to me. He readily took out time to help me out and
I'd like to also thank Miss Nirupama Rao, for helping me out and providing me with reading material for the topics that I had to cover.I'd also like to thank her to motivate me to work further on the questions I was required to solve. Mr Sagar Shrivastav, my fellow internee, who offered his help whenever asked. His work was also very enlightening and motivated me to work harder on mine.

References

- Linear Algebra Done Right, Second Edition, Sheldon Axler
- R. Pereira, Differentiators and the geometry of polynomials, (English summary) J. Math. Anal. Appl. 285 (1) (2003) 336348, MR2000158 (2004g:15041)
- B.V. Rajarama Bhat, Mithun Mukherjee, Integrators of matrices, Linear Algebra and its Applications 426 (2007) 7182