Soham Chatterjee

Email: soham.chatterjee@tifr.res.in

Course: Mathematical Foundations for Computer Sciences

Problem 1

Let V be a vector space over $\mathbb{R}$. Show that the set $V_{\mathbb{C}} = V \times V$ with the operations below is a vector space over $\mathbb{C}$

$$(v_1, v_2) + (v_1', v_2') = (v_1 + v_1', v_2 + v_2')$$
$$(a + bi) \cdot (v_1, v_2) = (av_1 - bv_2, bv_1 + av_2)$$

This is called complexification and (v_1, v_2) is often denoted as $v_1 + v_2 i$. Show that:

- If B is a basis of V, it is also a basis of $V_{\mathbb{C}}$.

- For $\theta \in L(V)$, define the complexified operator $\theta_{\mathbb{C}} \in L(V_{\mathbb{C}})$ so that $\theta_{\mathbb{C}}(v_1 + v_2 i) = \theta(v_1) + \theta(v_2)i$. Show that for any basis B of V, we have $[\theta_{\mathbb{C}}]_B = [\theta]_B$

- For all $\lambda \in \mathbb{R}$, λ is an eigenvalue of θ if and only if it is an eigenvalue of $\theta_{\mathbb{C}}$. For $\lambda \in \mathbb{C}$, λ is an eigenvalue of $\theta_{\mathbb{C}}$ if and only if $\overline{\lambda}$ is an eigenvalue of $\theta_{\mathbb{C}}$ and they have the same multiplicity. Conclude that every real operator over an odd dimensional real vector space has an eigenvalue.

Solution:

- B is a basis of V. Let $\dim V = n$. Suppose $B = \{b_1, \ldots, b_n\}$. We want to show B is also a basis of $V_{\mathbb{C}}$ i.e. the set $B' = \{(b_i, 0) : i \in [n]\}$ is a basis of $V_{\mathbb{C}}$. So we have to show $\langle B_{\mathbb{C}} \rangle = V_{\mathbb{C}}$. From now if B is a basis of V then by $B_{\mathbb{C}}$ we denote the set $\{(b, 0) : b \in B\}$.

  Now $\forall\, i \in [n]$, $(b_i, 0) \in V_{\mathbb{C}}$. Therefore $B_{\mathbb{C}} \subseteq V_{\mathbb{C}}$. Hence $\langle B_{\mathbb{C}} \rangle \subseteq V_{\mathbb{C}}$. Now we have to show that $\langle B_{\mathbb{C}} \rangle \supseteq V_{\mathbb{C}}$. So suppose $(v_1, v_2) \in V_{\mathbb{C}}$. Then $v_1, v_2 \in V$. Hence $\exists!\ \{a_{1,i}\}_{i \in [n]}$ and $\{a_{2,i}\}_{i \in [n]}$ such that

$$v_1 = \sum_{i=1}^{n} a_{1,i} b_i, \qquad v_2 = \sum_{i=1}^{n} a_{2,i} b_i$$

Now for any $v \in V$, $(a + bi) \cdot (v, 0) = (av, bv)$. Therefore we have

$$\sum_{i=1}^{n} \left(a_{1,i} + a_{2,i} i\right)(b_i, 0) = \sum_{i=1}^{n} \left(a_{1,i} b_i, a_{2,i} b_i\right) = \left(\sum_{i=1}^{n} a_{1,i} b_i, \sum_{i=1}^{n} a_{2,i} b_i\right) = (v_1, v_2)$$

Therefore $(v_1, v_2) \in \langle B_{\mathbb{C}} \rangle$. Hence

$$V_{\mathbb{C}} \subseteq \langle B_{\mathbb{C}} \rangle \implies V_{\mathbb{C}} = \langle B_{\mathbb{C}} \rangle$$

Hence B is also a basis of $V_{\mathbb{C}}$

- By the above part we know if B is a basis of V then $B_{\mathbb{C}}$ is basis of $V_{\mathbb{C}}$. Now if $\theta \in L(V)$ then $\theta_{\mathbb{C}} \in L(V_{\mathbb{C}})$ such that $\theta_{\mathbb{C}}(v_1 + v_2 i) = \theta(v_1) + \theta(v_2)i$. So for any $b + 0i \in B_{\mathbb{C}}$ we have

$$\theta_{\mathbb{C}}(b + 0i) = \theta(b) + \theta(0)i = \theta(b) + 0i$$

Let b_j be the jth vector of B. $\exists!\ a_{j,l}$ forall $l \in [n]$ such that $\theta\left(b_j\right) = \sum_{l=1}^{n} a_{j,l} b_l$. Then $[\theta]_B = \left(a_{j,l}\right)_{1 \le j, l \le n}$. Then

$$\theta_C\left(b_j\right) = \theta(b) + 0i = \sum_{l=1}^{n} a_{j,l} b_l + 0i = \sum_{l=1}^{n} \left(a_{j,l} + 0i\right)(b_l + 0i) = \sum_{l=1}^{n} a_{j,l}(b_l + 0i)$$

Therefore $[\theta_{\mathbb{C}}]_B = \left(a_{j,l}\right)_{1 \le j, l \le n}$. Therefore $[\theta_{\mathbb{C}}]_B = [\theta]_B$.

- - Let $\lambda \in \mathbb{R}$ is an eigenvalue of $\theta \in L(V)$. Suppose $v \in V$, $v \neq 0$ be eigenvector corresponding to λ. Then in $V_{\mathbb{C}}$ we have the vector $v + 0i$. Then

  $$\theta_{\mathbb{C}}(v + 0i) = \theta(v) + \theta(0)i = \lambda v + 0i = \lambda v + \lambda \cdot 0i = \lambda(v + 0i)$$

  Hence λ is also an eigenvalue of $\theta_{\mathbb{C}}$. Now suppose $\lambda \in \mathbb{R}$ is an eigenvalue of $\theta_{\mathbb{C}}$. Then suppose $v_1 + v_2 i \in V_{\mathbb{C}}$, $v_1 + v_2 \neq 0$ be an eigenvector corresponding to λ. Now

  $$\theta_{\mathbb{C}}(v_1 + v_2 i) = \theta(v_1) + \theta(v_2)i, \ \theta_{\mathbb{C}}(v_1 + v_2 i) = \lambda(v_1 + v_2 i) = \lambda v_1 + \lambda v_2 i \implies \theta(v_1) + \theta(v_2)i = \lambda v_1 + \lambda v_2 i$$

  Hence we get $\theta(v_1) = \lambda v_1$ and $\theta(v_2) = \lambda v_2$. Since $v_1 + v_2 i \neq 0$, either $v_1 \neq 0$ or $v_2 \neq 0$. So there exists at least one eigenvector for λ in V.

  - Suppose $\lambda \in \mathbb{C}$. Now we know $\overline{\overline{\lambda}} = \lambda$. So showing if λ is eigenvalue of $\theta_{\mathbb{C}} \implies \overline{\lambda}$ is eigenvalue of $\theta_{\mathbb{C}}$ is enough since then replacing $\overline{\lambda}$ in place of λ we get that if $\overline{\lambda}$ is eigenvalue of $\theta_{\mathbb{C}} \implies \overline{\overline{\lambda}} = \lambda$ is eigenvalue of $\theta_{\mathbb{C}}$. Now suppose $v_1 + v_2 i \in V_{\mathbb{C}}$, $v_1 + v_2 i \neq 0$ be eigenvector corresponding to λ. Let $\lambda = a + bi$ where $a, b \in \mathbb{R}$. Then

  $$\lambda(v_1 + v_2 i) = (a + bi)(v_1 + v_2 i) = (av_1 - bv_2, bv_1 + av_2) = \theta(v_1) + \theta(v_2)i$$

  Hence we have $\theta(v_1) = av_1 - bv_2$ and $\theta(v_2) = bv_1 + av_2$. Hence

  $$\theta_{\mathbb{C}}(v_2 + iv_1) = \theta(v_2) + \theta(v_1)i = (av_2 + bv_1) + (av_1 - bv_2)i = (a - bi)(v_2 + v_1 i)$$

  Hence $v_2 + v_1 i$ is an eigenvector of $\theta_{\mathbb{C}}$ with corresponding eigenvalue $\overline{\lambda}$. Hence $\omega \lambda$ is an eigenvalue of $\theta_{\mathbb{C}}$. Therefore λ is eigenvalue of $\theta_{\mathbb{C}}$ if and only if $\overline{\lambda}$ is eigenvalue of $\theta_{\mathbb{C}}$.

    Now we want to show that they have the same multiplicity. We shown above that $v_1 + v_2 i$ is eigenvector of λ if and only $v_2 + v_1 i$ is eigenvector of $\overline{\lambda}$. Hence there is a bijection between the eigenvectors of λ and $\overline{\lambda}$. Hence their multiplicities are also same.

  - Suppose V is a vector space over $\mathbb{R}$ such that $\dim V = 2n + 1$ for some $n \in \mathbb{Z}_0$. Now take the complexification of V, $V_{\mathbb{C}}$ and then take the complexified operator of θ, $\theta_{\mathbb{C}}$. Since a basis of V is also a basis of $V_{\mathbb{C}}$ we can say $\dim V = \dim V_{\mathbb{C}}$. Let $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ are the distinct eigenvalues of $\theta_{\mathbb{C}}$. Since $V_{\mathbb{C}}$ is a vector space over $\mathbb{C}$ we can express V as direct sum of generalized eigenspaces i.e. $V = \bigoplus_{i=1}^{k} G_{\theta_{\mathbb{C}}}(\lambda_i)$. Now for any $i \in [m]$, $\dim G_{\theta_{\mathbb{C}}}(\lambda_i) = Mult(\lambda_i)$. Therefore $\sum_{i=1}^{k} Mult(\lambda_i) = 2n + 1$.

    Now for each $\lambda \in \mathbb{C}$ by the above part we have that λ is eigenvalue iff $\overline{\lambda}$ is eigenvalue with same multiplicities. Hence we can pair an eigenvalue with its conjugate $(\lambda, \overline{\lambda})$. From each pair $2Multi(\lambda)$ is contributed to the dimension. For any complex number with nonzero imaginary part along with its conjugate always an even number is contributed to the dimension. But the dimension is odd. So there is an eigenvalue with imaginary part is zero so that its conjugate is the number itself and therefore contribute to the dimension calculation only once. Hence there is an real eigenvalue λ of $\theta_{\mathbb{C}}$. Therefore by the second part we have that λ is an eigenvalue of θ. Hence θ has an real eigenvalue if the vector space has odd dimension over $\mathbb{R}$.

    ■

Problem 2

Let there be n goods and let the price of good $i \in [n]$ be c_i. Suppose there m orders, where each order $j \in [m]$ is a subset $S_j \subseteq [n]$ saying which goods have been ordered. Suppose there exists $t > 0$ such that no good is ordered more than t times. Formally we have, for all $i \in [n]$:

$$|\{j \in [m] \mid i \in S_j\}| \leq t$$

Imagine that the government demonetizes all currency below ₹100 and thus all the prices have to be rounded to a multiple of 100. Show that there is a way of rounding the prices to c_i' that are multiples of 100 such that for all $j \in [m]$. the cost of order j changes by at most ₹$100t$. Formally, for all $j \in [m]$:

$$\left| \sum_{i \in S_j} c_i - \sum_{i \in S_j} c_i' \right| \leq 100t$$

Solution: For any set $S \subseteq [n]$, $Cost(S) = \sum_{i \in S} c_i$. Let $\mathcal{F}$ denote the set of orders.

Observation: Now since at most $100t$ error can be tolerated we can only think about the sets which have more than t elements since for sets with less than or equal can deviate at most $100t$ from the original cost of the set. Therefore from now on we will only focus on the sets with more than t elements.

Now each element appears in at most t sets. So if we add the cardinalities of the sets with more than t elements it can be at most nt. But since each set has more than t elements number of such sets is less than n.

Let $\mathcal{A}$ be the set of active goods. Initially we start with all goods inactive. So at time $T = 0$ we have $|\mathcal{A}| = 0$. Let at time T the set

$$\mathcal{F}_T := \{S - \mathcal{A} : S \in \mathcal{F}, \ |S - \mathcal{A}| > t \ \& \ \text{For } S_1 - \mathcal{A} = S_2 - \mathcal{A} \text{ keep two copies}\}$$

Let M_T is the $|\mathcal{F}_T| \times (n - |\mathcal{A}|)$ matrix where the ith row corresponds to i^{th} set of $\mathcal{F}_T$ where each entry of M is from $\{0, 1\}$. Now in i^{th} row M in j^{th} column has 1 if and only if j^{th} good is in the ith set of $\mathcal{F}_T$. So we will now describe an algorithm which at each time set a value multiple of 100 to at least one of the cost of the n goods then make that good active and we will not change the new cost of the value while keeping the total cost for each order in $\mathcal{F}_0$ to be the same as the total cost with respect to the original costs. This naturally suggests T can be at most n. Now P_T be the total costs of each orders with respect to the cost function at time C_T. So at time $T = 0$ we have $C_0 = C$, the given costs of the goods and $P_0 = M_0 C$.

At time T we have $\mathcal{F}_T$. So we restrict M_{T-1} to only the rows of $\mathcal{F}_{T-1}$ and also we remove the corresponding columns of $\mathcal{A}$ and after restricting M_{T-1} we have the matrix M_T. We denote the ith element of $\mathcal{F}_T$ by S_i .We also remove the rows of P_0 in $\mathcal{A}$ and then for i^{th} element we make

$$P_T(i) = P(i) - \sum_{j \in S_i \cap \mathcal{A}} C_{T-1}(j)$$

We made P_T like this to ensure $M_T \, C_{T-1}|_{\mathcal{F}_T} = P_T$ where $C_{T-1}|_{\mathcal{F}_T}$ is restriction of C_{T-1} by removing the rows not in $\mathcal{F}_T$.

Now number of rows of M_T is $|\mathcal{F}_T|$ and number of columns of M_T is $n - |\mathcal{A}|$. Now for each set $\mathcal{F}_T$ it has more than t elements. Therefore

$$t \cdot |\mathcal{F}_T| \leq \{(i, S) : S \in \mathcal{F}_T, \ i \in S\} \leq (n - |\mathcal{A}|) \cdot t \implies |\mathcal{F}_T| < n - \mathcal{A}|$$

So we want to find a new vector x_T such that $M_T x_T = 0$. And by the above we have number of equations is less than number of variables. So we can actually get a vector $x_T \notin \langle C_{T-1} \rangle$ and $\|xTT\|_1 \leq 1$ that is the l_1 norm of x_T is less than 1 such that $M_T x_T = 0$. Then we can use $C_{T-1} + \gamma x_T$ such that $\exists \ i \in [n] - \mathcal{A}$ such that $C_{T-1}(i) + \gamma x_T(i)$ is the closest multiple of 100 of c_i. So we take $C_T = C_{T-1} + \gamma X_T$. We make i^{th} good active and add it to $\mathcal{A}$. Now we remove the sets S in $\mathcal{F}_T$ which have $\leq t$ many inactive good after updating $\mathcal{A}$ and now we get $\mathcal{F}_{T+1}$. So we have now $\mathcal{F}_{T+1}, M_T, C_T, P_T, \mathcal{A}$. So if $\mathcal{F}_{T+1} \neq \emptyset$ we can do the

same algorithm one more time and we can keep doing this until at some time $T = k \leq n$ we have $\mathcal{F}_k = \emptyset$. Then we have all the sets with $\leq t$ inactive goods. Now we set the costs of the unset variables to be the closest multiple of the of the corresponding values in C_k. And by the first observation the deviation of the costs for each order will be at most $100t$. ∎

> ## Problem 3
> Let $\theta \in L(V)$ and $\theta^* \in L(V^*)$ be its dual. For any basis B of V, show that there exists a basis B^* of V^* such that $[\theta]_B = [\theta^*]_{B^*}$.
> Hint: It may be helpful to use another question in this assignment to solve this one.

Solution:

V is a vector space over $\mathbb{F}$. Let $\dim V = n$. Suppose B is a basis of V. Let $\tilde{B}$ denote the dual vectors of B which is also basis of V^*. Now we know $[\theta^*]_{\tilde{B}} = [\theta]_B^T$. Hence now we need to show that there is a basis B^* such that changing into that basis we get the matrix for θ to be $[\theta^*]_{B^*} = [\theta^*]_{\tilde{B}}^T$ then we will have $[\theta^*]_{B^*} = [\theta]_B$. So the problem is now reduced to showing that for any $\theta \in L(V)$ where V is an $n-$dimensional vector space $\mathbb{C}$ if B is a basis of V then there exists a basis B' such that $[\theta]_{B'} = [\theta]_B^T$.

Suppose $\lambda_1, \ldots, \lambda_m$ are the eigenvalues of θ. Then we know there exists a basis B_J of V such that $[\theta]_{B_J}$ is block diagonal where every block J_i corresponds to one of the λ_i and it has dimension $Mult(\lambda_i) = d_i$ and the block has λ_i's on the diagonal and 0's or 1's above the diagonal i.e.

$$J_i = \begin{bmatrix} \lambda & t_{i,1} & & & \\ & \lambda_i & t_{i,2} & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & t_{i,d_i-1} \\ & & & & \lambda_i \end{bmatrix} \implies [\theta]_{B_J} = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{bmatrix}$$

Now for each $i \in [m]$, J_i can be written as block matrices of the form $\begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}$ for each $t_{i,k} = 0$. For $i \in [m]$ let g_i be number of such blocks for λ_i. Then suppose $J_{\lambda_i,j}$ for $j \in [g_i]$ denote the jth block matrix of the above form. So we have

$$J_i = \begin{bmatrix} J_{i,1} & & \\ & \ddots & \\ & & J_{i,g_i} \end{bmatrix} \implies [\theta]_{B_J} = \begin{bmatrix} \begin{bmatrix} J_{1,1} & & \\ & \ddots & \\ & & J_{1,g_1} \end{bmatrix} & & \\ & \ddots & \\ & & \begin{bmatrix} J_{m,1} & & \\ & \ddots & \\ & & J_{m,g_m} \end{bmatrix} \end{bmatrix}$$

Hence there exists $n \times n$ invertible P matrix such that

$$[\theta]_{B_J} = P^{-1}[\theta]_B P$$

We need a basis B' such that $[\theta]_{B'} = [\theta]_B^T$ i.e. we need to find there exists $n \times n$ invertible matrix Q such that $[\theta]_B^T = Q^{-1}[\theta]_B Q$. Now if we can find an $n \times n$ invertible matrix B such that $B^{-1}[\theta]_{B_J} B = [\theta]_{B_J}^T$ Then we have

$$P^{-1}[\theta]_B P = [\theta]_{B_J} \implies P^T[\theta]_B^T(P^{-1})^T = P^T[\theta]_B^T(P^T)^{-1} = [\theta]_{B_J}^T = B^{-1}[\theta]_{B_J}B = B^{-1}(P^{-1}[\theta]_B P)B$$
$$\implies [\theta]_B^T = \left[(P^T)^{-1}B^{-1}P^{-1}\right][\theta]_B\left[PBP^T\right]$$

Therefore from P and B we get a base changing matrix $Q = PBP^T$ such that $[\theta]_B^T = Q^{-1}[\theta]_B Q$. Hence we only have to find B now.

Now if we can have a base changing matrix $B_{i,j}$ for $i \in [m]$ and $j \in [g_i]$ such that $B_{i,j}^{-1} J_{i,j} B_{i,j} = J_{i,j}^T$. Then we can construct

$$B = \begin{bmatrix} B_{1,1} & & & & & & \\ & \ddots & & & & & \\ & & B_{1,g_1} & & & & \\ & & & \ddots & & & \\ & & & & B_{m,1} & & \\ & & & & & \ddots & \\ & & & & & & B_{m,g_m} \end{bmatrix}$$

and then we will have $B_{-1}[\theta]_{B_J} B = [\theta]_{B_J}^T$. Therefore we only have to find $B_{i,j}$ for $J_{i,j}$ for any $i \in [m]$ and $j \in [g_j]$.

So now by we will not mention the indices and will construct an invertible matrix B such that for any $J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$ we have $B^{-1} J B = J^T$. For that take

$$B^{-1} = B = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \implies B^{-1} J B = J^T$$

Hence we get a base changing matrix Q such that $[\theta]_B^T = Q^{-1}[\theta]_B Q$. Hence there exists a basis B' such that changing basis make the matrix of θ to be transpose. Therefore for any $\theta \in L(V)$ for any basis B of V there is a basis $\tilde{B}$ of V^* such that $[\theta^*]_{B^*} = [\theta^*]_{\tilde{B}}^T \implies [\theta]_B = [\theta^*]_{\tilde{B}}$. ■

Problem 4 Solving linear recurrence equations

Let $A = (a_0, a_1, \dots)$ be a sequence of complex numbers. We say that it satisfies a linear recurrence relation with constant coefficients if there is a number $t > 0$ and constants $c_1, \dots, c_t$ such that, for all $i \geq t$, we have:
$$a_i = c_1 a_{i-1} + c_2 a_{i-2} + \cdots + c_t a_{i-t}$$
Observe that if A satisfies such a recurrence relation, then A is determined by $(a_1, \dots . a_{t-1})$. In this exercise, we compute all solutions to an arbitrary such recurrence relation.

- For all $i \geq 0$, define the vector $b_i = (a_i, a_{i+1}, \dots, a_{i+t-1})^T \in \mathbb{C}^i$. Describe a linear operator $\theta : \mathbb{C}^t \to \mathbb{C}^t$ that, for all $i \geq 0$, maps b_i to b_{i+1}.

- Show how to compute the powers of a matrix in Jordan normal form, and use it to show that if $\lambda_1, \dots, \lambda_m$ are the eigenvalues of θ with multiplicities $d_1, \dots, d_m$ respectively, then sequence A satisfied the recurrence if and only if there exists polynomials $q_1, \dots, q_m$ such that $\deg(q_j) \leq d_j$ for all $j \in [m]$ and for all $i \geq 0$, we have:
$$a_i = \sum_{j=1}^m q_j(n) \lambda_j^n$$

- Let $F = (f_0, f_1, f_2, \dots)$ be the Fibonacci sequence. Namely, F satisfies $f_0 = 0$, $f_1 = 1$, and the linear recurrence relation $f_i = f_{i-1} + f_{i-2}$ for all $i \geq 2$. Compute a formula for this sequence by computing the polynomials q_j.

Solution:

- Take the standard basis $\{e_i \mid i \in [t]\}$ of $\mathbb{C}^t$ where $e_i = (\underbrace{0\ 0\ \cdots\ 0}_{(i-1)\ \text{times}}\ 1\ 0\ 0\ \cdots\ 0)^T$. Then define

$$\theta(e_{i+1}) = e_i + c_{i+1}e_t \ \forall \ i \in [t-1] \quad \text{and} \quad \theta(e_1) = c_1 e_1$$

  Since there is only one $\theta \in L(\mathbb{C}^t)$ such that these properties are satisfied we have defined θ uniquely. Now for any $i \geq 0$ we have

$$\begin{aligned}
\theta(b_i) = \theta\left(\sum_{j=0}^{t-1} a_{i+j}e_{j+1}\right) &= \sum_{j=0}^{t-1} a_{i+j}\theta(e_{j+1}) \\
&= \sum_{j=1}^{t-1} a_{i+j}(e_j + c_{j+1}e_t) + a_i\theta(e_1) \\
&= \sum_{j=1}^{t-1} a_{i+j}e_j + \left(\sum_{j=0}^{t-1} a_{i+j}c_{j+1}\right)e_t = \sum_{j=1}^{t-1} a_{i+j}e_j + a_{i-t}e_t \\
&= \sum_{j=0}^{t-1} a_{i+j+1}e_{j+1} = b_{i+1}
\end{aligned}$$

  Hence this linear map has the property that for all $i \geq 0$, b_i gets mapped to b_{i+1}. The matrix will be

$$[\theta] = \begin{bmatrix} 0 & & & \\ \vdots & & I_{t-1} & \\ 0 & & & \\ c_1 & c_2 & \cdots & c_t \end{bmatrix}$$

- Suppose $\dim V = n$. The eigenvalues of θ are $\lambda_1, \ldots, \lambda_m$ with multiplicities $d_1, \ldots, d_m$. Hence let J_i denote the jordan block corresponding to λ_i. Then the Jordan normal form of matrix of θ looks like

$J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{bmatrix}$ with respect to some basis. Hence there exists $n \times n$ invertible matrix P such

  that $[\theta] = P^{-1}JP$. Therefore $[\theta]^n = (P^{-1}JP)^n = P^{-1}J^nP$. Therefore enough to calculate the power of J. Now

$$J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_m \end{bmatrix} \implies J^n = \begin{bmatrix} J_1^n & & \\ & \ddots & \\ & & J_m^n \end{bmatrix}$$

  Hence we only have to calculate power of each Jordan blocks. Now for any Jordan block with eigenvalue λ there can be 1 or 0 in the positions just above the diagonal. Let λ has geometric multiplicity l. Then there will be $l-1$ zeros in the positions just above the diagonal. Hence we can break J_λ into smaller blocks of the form

$$J_{\lambda,i} = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

  For $i \in [l]$. Then we have

$$J_\lambda^n = \begin{bmatrix} J_{\lambda,1}^n & & \\ & \ddots & \\ & & J_{\lambda,l}^n \end{bmatrix}$$

6

Now for each $i \in [l]$ we have $J_{\lambda,i} = \lambda I + N$ where N has 1's on the positions just above the diagonal. Then

$$J_{\lambda,i}^n = (\lambda I + N)^n = \sum_{j=0}^{n} \binom{n}{j} \lambda^{n-j} N^j$$

Hence we have finally

$$J_{\lambda,i}^n = \begin{bmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \cdots & \cdots & \binom{n}{p}\lambda^{n-p} \\ & \lambda^n & \binom{n}{1}\lambda^{n-1} & \cdots & \cdots & \binom{n}{p-1}\lambda^{n-p+1} \\ & & \ddots & \ddots & \vdots & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \lambda^n & \binom{n}{1}\lambda^{n-1} \\ & & & & & \lambda^n \end{bmatrix}$$

Hence from this we can also conclude that every entry of $[\theta]^n$ is linear combinations of $\{\lambda_i^n, n\lambda_i^n, \ldots, n^{d_i}\lambda_i^n\}_{i \in [m]}$. Now we know for all $n \geq 0$

$$b_{n+t-1} = [\theta]^n b_0$$

We know every element of $[\theta]^n$ is linear combinations of $\{\lambda_i^n, n\lambda_i^n, \ldots, n^{d_i}\lambda_i^n\}_{i \in [m]}$. Therefore for all $i \in [t]$ $\exists b_{i,j,t_j} \in \mathbb{C}$ such that

$$a_n = \sum_{i \in [t]} a_{i-1} \sum_{j \in [m], 0 \leq t_j \leq d_j} b_{i,j,t_j} n^{t_j} \lambda_j^n = \sum_{j \in [m]} \left[\sum_{i \in [t], 0 \leq t_j \leq d_j} a_{i-1} b_{i,j,t_j} n^{t_j} \right] \lambda_j^n$$

So define the polynomial $q_j(x) = \sum_{i \in [t], 0 \leq t_j \leq d_j} a_{i-1} b_{i,j,t_j} x^{t_j}$. We know $\deg q_j \leq d_j$. Then we have

$$a_n = \sum_{j=1}^{m} q_j(n) \lambda_j^n$$

Now suppose the opposite. There exists $q_1, \ldots, q_m$ such that $a_n = \sum_{j=1}^{m} q_j(n)\lambda^n$ for all $n \geq 0$. Each q_j can be expressed as $q_j(n) = \sum_{t_j=0}^{d_j} p_{j,t_j} n^{t_j}$. Now from the above discussion it is visible that q_j is uniquely determined by the first $a_0, \ldots, a_{t-1}$ values. So the coefficients p_{j,t_j} for all j, t_j is uniquely determined by $a_0, \ldots, a_{t-1}$ and through a linear transformation from $(a_0, \ldots, a_{t-1}) \mapsto (p_{j,t_j})_{j \in [m], 0 \leq t_j \leq d_j}$. Since $\sum_{j=1}^{t} d_j = t$ we have an injective linear map from $\mathbb{C}^t$ to $\mathbb{C}^t$. Hence it is surjective. Therefore if we have an isomorphism and we can take the inverse map and from $(p_{j,t_j})_{j \in [m], 0 \leq t_j \leq d_j}$ we can get back $(a_0, \ldots, a_{i-t})$. And then produce the whole sequence by the recurrence relation but using the formula that $a_n = \sum_{j=1}^{m} q_j(n)\lambda^n$ for all $n \geq 0$.

- We have the recurrence relation $F(n) = F_{n-1} + F_{n-2}$ with $F_1 = F_2 = 1$. So we can represent this with matrices like following:

$$\begin{bmatrix} f_n \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f_{n-1} \\ f_{n-2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^2 \begin{bmatrix} f_{n-2} \\ f_{n-3} \end{bmatrix} = \cdots \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-2} \begin{bmatrix} f_2 \\ f_1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^{n-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Denote $\overline{f}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\overline{f}_k = \begin{bmatrix} f_{k+1} \\ f_k \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Therefore we have $\overline{f}_n = A^n \overline{f}_0$.

Now clearly A has full rank and $\forall\, k \in \mathbb{N}$, $\overline{f}_k \in \mathbb{R}^2$. So we will find the eigenvalues of A to find an eigenbasis.

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

So if $t^2 - t - 1 = 0$ then

$$t = \frac{1 \pm \sqrt{1+4}}{2} = \frac{1 \pm \sqrt{5}}{2}$$

So denote $\varphi = \frac{1+\sqrt{5}}{2}$ and $\psi = \frac{1-\sqrt{5}}{2}$. Now let $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ be an eigenvector corresponding to φ. Then

$$AX = \begin{bmatrix} x_1 + x_2 \\ x_1 \end{bmatrix} = \varphi \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Therefore $x_1 = \varphi x_2$. Therefore take $x_2 = 1$ then we have $x_1 = \varphi$. So $X = \begin{bmatrix} \varphi \\ 1 \end{bmatrix}$. Similarly we have $Y = \begin{bmatrix} \psi \\ 1 \end{bmatrix}$ is an eigenvector of A corresponding to ψ.

Now we want to express $\overline{f}_0$ as a linear combination of X, Y. Notice

$$\frac{1}{\sqrt{5}}(X - Y) = \frac{1}{\sqrt{5}} \begin{bmatrix} \varphi - \psi \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \frac{1=\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2} \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \sqrt{5} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \overline{f}_0$$

Therefore

$$\overline{f}_n = A^n \overline{f}_0 = A^n \left(\frac{1}{\sqrt{5}}(X - Y) \right) = \frac{1}{\sqrt{5}}(AX - AY) = \frac{1}{\sqrt{5}}(\varphi^n X - \psi^n Y) = \frac{1}{\sqrt{5}} \varphi^n \begin{bmatrix} \varphi \\ 1 \end{bmatrix} - \frac{1}{\sqrt{5}} \psi^n \begin{bmatrix} \psi \\ 1 \end{bmatrix}$$

Therefore $f_n = \frac{\phi^n - \psi^n}{\sqrt{5}}$.

∎

Problem 5

Let U and V be vector spaces over a field $\mathbb{F}$ and $\theta \in L(U,V)$. If A and B are bases of U and V respectively, we use $_B[\theta]_A$ to denote the matrix of θ with respect to bases A and B. If $U = V$ and $A = B$, we simply write $[\theta]_B$

- Let U, V, W be vector spaces over a field $\mathbb{F}$ with bases A, B, C respectively. Let $\theta \in L(U,V)$ and $\phi \in L(V,W)$. Show that $_C[\phi \circ \theta]_A = {_C[\phi]_B} \circ {_B[\theta]_A}$.

- Let V be vector space and $I \in L(V)$ be the identity transformation. Let A and B be bases of V. Show that $_B[I]_A$ is invertible. Conversely, show that if A is a basis of V and M is an invertible matrix, then there exists another basis B of V such that $M = {_B[I]_A}$. Thus, invertible matrices correspond to basis changes.

- Using these, show that if M and N are square matrices and N is invertible, then $rank(MN) = rank(M)$. Deduce that converting a matrix to reduced row echelon form preserves its rank. This justifies the use of reduced row echelon forms.

Solution:

- Let $A = \{a_1, \ldots, a_n\}$, $B = \{b_1, \ldots, b_m\}$ and $C = \{c_1, \ldots, c_k\}$. Let $\exists! \, \alpha_{i,j} \in \mathbb{F}$ for all $i \in [n]$ and $j \in [m]$ and $\exists! \, \beta_{j,l}$ for all $j \in [m]$ and $l \in [k]$ such that

$$\theta(a_i) = \sum_{j=1}^n \alpha_{i,j} b_j \quad \text{and} \quad \phi(b_j) = \sum_{l=1}^k \beta_{j,l} c_l$$

Therefore $_B[\theta]_A(i,j) = \alpha_{j,i}$ and $_C[\phi]_B(j,l) = \beta_{l,j}$. Now

$$\phi \circ \theta(a_i) = \phi \left(\sum_{j=1}^n \alpha_{i,j} b_j \right) = \sum_{j=1}^n \alpha_{i,j} \phi(b_j) = \sum_{j=1}^n \alpha_{i,j} \sum_{l=1}^k \beta_{j,k} c_l = \sum_{l=1}^k \left(\sum_{j=1}^n \alpha_{i,j} \beta_{j,l} \right) c_l$$

8

Therefore $_C[\phi \circ \theta]_A(i,l) = \sum_{j=1}^{n} \alpha_{i,j}\beta_{j,l}$ for all $l \in [k]$ and $i \in [n]$. Now $_C[\phi \circ \theta]_B(i,l) = \sum_{j=1}^{n} \alpha_{i,j}\beta_{j,l}$ is the product of lth row of $_C[\phi]_B$ and i^{th} column of $_B[\theta]_A$. Hence we have $_C[\phi \circ \theta]_A = {}_C[\phi]_B \circ {}_B[\theta]_A$

-
  - Let $A = \{a_1, \ldots, a_n\}$, $B = \{b_1, \ldots, b_n\}$. Then $\exists! \; \alpha_{i,j}$ for all $i,j \in [n]$ such that

$$a_i = \sum_{j=1}^{n} \alpha_{i,j}b_j \implies {}_B[I]_A(i,j) = \alpha_{i,j}$$

  Now we also have that $\exists! \; \beta_{i,j}$ for all $i,j \in [n]$ such that

$$b_i = \sum_{j=1}^{n} \beta_{i,j}a_j \implies {}_A[I]_B(i,j) = \beta_{i,j}$$

  Now we have $_A[I]_B {}_B[I]_A = {}_A[I]_A$ by the above part. Now $_A[I]_A$ is just the identity matrix. Therefore

$$\det\{\}_A[I]_A \neq 0 \implies \det(_A[I]_B {}_B[I]_A) = \det(_A[I]_B)\det(_B[I]_A) \neq 0 \implies \det(_A[I]_B), \det(_B[I]_A) \neq 0$$

  Therefore $\det(_B[I]_A) \neq 0$ and hence $_B[I]_A$ is invertible.

  - Let $A = \{a_1, \ldots, a_n\}$, $B = \{b_1, \ldots, b_n\}$ where we want to find b_i for $i \in [n]$. Now since A is basis $\exists! \; \alpha_{i,j} \in \mathbb{F}$ for all $i,j \in [n]$ such that $b_i = \sum_{j=1}^{n} \alpha_{i,j}a_j$. Denote T be the $n \times n$ matrix such that $T = (\alpha_{i,j})_{1 \leq i,j \leq n}$. Now suppose $M = (m_{i,j})_{1 \leq i,j \leq n}$. We want $M = {}_B[I]_A$ i.e. $a_i = \sum_{j=1}^{n} m_{i,j}b_j$. Then we have

$$a_i = \sum_{j=1}^{n} m_{i,j}b_j = \sum_{j=1}^{n} m_{i,j} \sum_{k=1}^{n} \alpha_{j,k}a_k = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} m_{i,j}\alpha_{j,k} \right) a_k$$

  Hence we have $\sum_{j=1}^{n} m_{i,j}\alpha_{j,k} = 1$ if $i = k$ and 0 otherwise. Now $\sum_{j=1}^{n} m_{i,j}\alpha_{j,k}$ is product of ith row of M and k^{th} column of T. Hence we have $MT = I$ where I is the identity matrix. Therefore $T = M^{-1}$. Hence we get B by taking the inverse of M, M^{-1}. And let $M^{-1} = (\alpha_{i,j})_{1 \leq i,j \leq n}$ then $b_i = \sum_{j=1}^{n} \alpha_{i,j}a_j$ for all $i \in [n]$. Hence we get a basis B such that $M = {}_B[I]_A$.

Hence invertible matrices correspond to basis changes

- Let M and N are both $n \times n$ matrices. Now consider the vector space $\mathbb{F}^n$. Take the standard basis $A = \{e_i \mid i \in [n]\}$ where e_i has 1 at i^{th} position and rest are 0. Now by the above part we know there exists a basis $B = \{b_1, \ldots, b_n\}$ of $\mathbb{F}^n$ such that $N = {}_B[I]_A$ where $I \in L(\mathbb{F}^n)$ is the identity map. Suppose $M = (m_{i,j})_{1 \leq i,j \leq n}$. Then define the linear map $\theta \in L(\mathbb{F}^n)$ such that

$$\theta(b_i) = \sum_{j=1}^{n} m_{i,j}b_j \quad \forall \; i \in [n]$$

Hence there exists unique such $\theta \in L(\mathbb{F}^n)$ and hence $[\theta]_B = M$. Therefore we have by the 1st part

$$MN = [\theta]_B {}_B[I]_A = {}_B[\theta \circ I]_A = {}_B[\theta]_A$$

But we know for a linear map $\phi \in L(V, V)$ and a basis S of , $\dim(\text{Im}\,\phi) = rank([\phi]_S)$. Now if we change basis of image space the dimension of $\text{Im}\,\theta$ still remains same. Hence we have $rank(MN) = rank(_B[\theta]_A) = \dim(\text{Im}\,\theta)$ and $rank(M) = rank([\theta]_B) = \dim(\text{Im}\,\theta)$. Therefore we have $rank(MN) = rank(M)$.

Since elementary guassian row operations on a matrix M are just multiplying an invertible matrix from left to M. This is because swapping i^{th} and j^{th} rows operation is just multiplying from left by the matrix which almost looks like identity matrix but just the ith and j^{th} rows swapped. For multiplying i^{th} row by $a \in \mathbb{F}$ and then adding to j^{th} row the matrix with 1's on the diagonal and a at $(i,j)^{th}$ position is multiplied to M from left. Therefore converting a matrix M to its reduced row echelon form is just multiplying M with an invertible matrix N from left. Now by the above discussion $rank(NM) = rank(M^T N^T) = rank(M^T) = rank(M)$. Therefore converting a matrix to reduced row echelon form preserves its rank.

$\blacksquare$