McGill University - Winter 2019

Final Review - MATH 251

Ralph Sarkis

April 15, 2019

Exercise 1 (1.1). Are the following subspaces of $M_n(\mathbb{R})$ isomorphic?

(i) The upper triangular matrices (U_n) and the lower triangular matrices (L_n).

(ii) The symmetric matrices S_n and U_n.

(iii) The anti-symmetric matrices A_n and S_n.

Solution. We will construct the isomorphism for (i), but for the two others, we only argue informally by comparing dimensions and leave the rigorous proof as an exercise.

(i) We send A to its transpose A^t. It is clear that it maps U_n to L_n and linearity follows from the properties of transposes, namely,

$$0^t = 0, \ \lambda A^t = (\lambda A)^t \text{ and } A^t + B^t = (A+B)^t.$$

To see that it is an isomorphism, it is enough to find an inverse. In this case, this map is its own inverse because $(A^t)^t = A$.

(ii) In a symmetric matrix ($A = A^t$), the entries below the diagonal are determined by the entries above the diagonal, so we only need to know the entries in the upper triangle to define a symmetric matrix. In other words, we have $\frac{n(n+1)}{2}$ degrees of freedom and similarly for U_n because we know the entries below the diagonal are zero. Informally, we have shown S_n and U_n both have dimension $\frac{n(n+1)}{2}$, so they are isomorphic.

(iii) There is an additional question in the sheet asking to show the symmetric matrices are in bijection with anti-symmetric matrices (A_n). However, this is not true as the space A_n does not have the same dimension as S_n. For any matrix in A_n, the entries below the diagonal are still determined by the entries above the diagonal, but the diagonal must be zero because $a = -a \implies a = 0$. This means we only have $\frac{(n-1)n}{2}$ degrees of freedom.

□

Exercise 2 (1.4)**.** Let U be a subspace of a finite dimensional vector space V and W be a complement of U, i.e.: a subspace of V such that $U + W = V$ and $U \cap W = \{0\}$. Fix a basis $w_1, \dots, w_k$ for W and for each $u \in U$, define

$$W(u) = \operatorname{span}\{u + w_1, \dots, u + w_k\}.$$

(i) Show that $W(u)$ is a complement of U.

(ii) Show that for any $u \neq u'$, $W(u) \neq W(u')$.

Solution. (i) We need to show $U + W(u) = V$ and $U \cap W(u) = \{0\}$. For the first part, let $v \in V$, then we can write $v = u' + w$ for $u' \in U$ and $w \in W$, furthermore, we have the decomposition $w = \sum_{i=1}^{k} c_i w_i$. Therefore, we have

$$W(u) \ni w' = \sum_{i=1}^{k} c_i(w_i + u) = w + \left(\sum_{i=1}^{k} c_i\right) u,$$

and if we let $u'' = u' - \left(\sum_{i=1}^{k} c_i\right) u \in U$, then $v = u'' + w'$. we conclude that $U + W(u) = V$.

For the second part, let $w' \in W(u) \cap U$ and decompose it as $\sum_{i=1}^{k} c_i(w_i + u)$, then we have

$$\sum_{i=1}^{k} c_i(w_i + u) - \left(\sum_{i=1}^{k} c_i\right) u = \sum_{i=1}^{k} c_i w_i \in W \cap U.$$

Because $U \cap W = \{0\}$, we conclude $\sum_{i=1}^{k} c_i w_i = 0$ and by independence that each $c_i = 0$, hence $w' = 0$.

(ii) Suppose that $W(u) = W(u')$, we will show that $u = u'$. If both spaces are equal, in particular, we have

$$u + w_1 = \sum_{i=1}^{k} c_i(w_i + u') = \sum_{i=1}^{k} c_i w_i + \left(\sum_{i=1}^{k} c_i\right) u'.$$

Equivalently, we have

$$w_1 - \sum_{i=1}^{k} c_i w_i = \left(\sum_{i=1}^{k} c_i\right) u' - u,$$

and we note that the L.H.S. is in W and the R.H.S. is in U, so both sides must be 0. By independence, we then have $c_1 = 1$ and $c_i = 0$ for $i > 1$, so the R.H.S. becomes $u' - u = 0$, yielding the desired $u = u'$.

$\square$

Exercise 3 (1.5)**.** Let V be a finite dimensional inner product space. Show that if U and W are orthogonal subspaces of V and $\dim U = \dim W = \dim V$, then $U^{\perp}$ and $W^{\perp}$ are orthogonal.

Solution. We claim (1) that W is a complement of U, namely $U \oplus W = V$, and (2) that orthogonal complements are unique. These two claims imply that W and U are orthogonal complements and hence that $U^\perp = W$ and $W^\perp = U$. The result then follows trivially.

(1) First, note that $U \cap W = \{0\}$ because if v is a vector in the intersection, then $\langle v, v \rangle = 0$, but by positive-definiteness of the inner product, v must be 0. Moreover, since we have

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap V) = \dim V - \dim(U \cap W) = \dim V,$$

we can infer that $U + W = V$ and then conclude $U \oplus W = V$. This shows U and W are orthogonal complements.

(2) Let W' be another orthogonal complement of U, then for any $w' \in W' \subset V$, we can write it uniquely (because U and W are complements) as $w' = u + w$ with $u \in U$ and $w \in W$. Now, we have

$$0 = \langle u, w' \rangle = \langle u, u + w \rangle = \langle u, u \rangle + \langle u, w \rangle = \langle u, u \rangle\,,$$

and by positive-definiteness, this implies $u = 0$, hence $w = w'$. Along with the symmetric argument and since w was arbitrary, we conclude that $W = W'$.

□

Exercise 4 (1.6)**.** Let $v_1, \dots, v_m \in V$ be independent and $w \in W$, what are the possible values of the dimension of $V(w) = \operatorname{span}\{v_1 + w, \dots, v_m + w\}$.

Solution. First, note that since $V(0)$ is the span of m independent vectors, it has dimension m. Moreover, recall from the proof of 2.(i) that for any w in the complement of $V(0)$, $V(w)$ is also a complement of the complement of $V(0)$, hence it has the same dimension m.

Next, since $V(w)$ is always the span of m vectors, it cannot have a dimension bigger than m. It can have a dimension $m - 1$ if, for instance $w = -v_1$. To see this, observe that $\{v_i - v_1 \mid i \geq 2\}$ is still linearly independent because the original set is independent and hence

$$0 = \sum_{i=2}^{m} c_i(v_i - v_1) = \sum_{i=2}^{m} c_i v_i - \left(\sum_{i=2}^{m} c_i\right) v_1 \Leftrightarrow c_i = 0, \forall i \geq 2.$$

Next, we claim that the dimension is at least $m - 1$. Let T be the linear map that sends v_i to v_i for each i and sends all other independent vectors to 0, it is clear that $\dim(\operatorname{Im}(T)) = m$. Let S be the linear map that sends v_i to w for each i and sends all other independent vectors to 0, it is clear that (if $w \neq 0$ a case that was already dealt with) $\dim(\operatorname{Im}(S)) = 1$.

Furthermore, notice that $\operatorname{Im}(T + S) = V(w)$ and recall that

$$\dim(\operatorname{Im}(T + S)) \geq \dim(\operatorname{Im}(T)) - \dim(\operatorname{Im}(S)).$$

Since $\dim(\operatorname{Im}(T)) = m$ and $\dim(\operatorname{Im}(S)) = 1$, the claim follows. □

Exercise 5 (N)**.** A linear map is injective if and only if its kernel is trivial.

Solution. Suppose T is not injective, then $T(u) = T(v)$ for $u \neq v$. By linearity, this implies $T(u - v) = 0$ and since $u - v \neq 0$, this shows the kernel is not trivial.

If the kernel is non-trivial, then two vectors are mapped to 0 and T is clearly not injective. □

Exercise 6 (N)**.** Let $T : V \to W$ be a linear map between two vector spaces of the same dimension n, show that T is injective if and only if it is surjective.

Proof. ($\Rightarrow$) If T is injective, then it is easy to check that a basis $\{v_1, \ldots, v_n\}$ of V is sent to a linearly independent set $\{T(v_1), \ldots, T(v_n)\} \subseteq W$. Since $\dim W = n$ this set must be spanning and hence T is surjective.

($\Leftarrow$) If T is surjective, then let $\{w_1, \ldots, w_n\}$ be a basis for W and for $1 \leq i \leq n$, let $v_i \in T^{-1}(w_i)$. It is easy to check that because the w_i's are independent, the v_i's must also be independent. Then, we infer that $\ker(T) = 0$ because non-zero vector is a non-trivial linear combination of the v_i's and hence is mapped to a non-trivial linear combination of the w_i's which cannot be zero. This shows T is injective. □

Proof 2 (uses material not covered in class). Alternatively, we can use the first isomorphism theorem. It says that $V/\ker(T) \cong \mathrm{Im}(T)$ and if we consider the dimensions of these spaces, we get $\dim(V) - \dim(\ker(T)) = \dim(\mathrm{Im}(T))$. Now, we observe

$$T \text{ is injective} \Leftrightarrow \dim(\ker(T)) = 0 \Leftrightarrow \dim(\mathrm{Im}(T)) = \dim(V) \Leftrightarrow T \text{ is surjective.}$$

□

Exercise 7 (N)**.** Let T be a linear operator on a vector space V such that for any $v \in V$, there exists $n \in \mathbb{N}$ such that $T^n(v) = 0$. Is T necessarily nilpotent? State a sufficient condition for T to be nilpotent.

Solution. This operator is not necessarily nilpotent. For instance, let $T : \mathbb{F}[x] \to \mathbb{F}[x]$ be the differentiation map, namely, $T(p(x)) = p'(x)$. It follows from basic properties of the derivative that T is linear. For any $n \in \mathbb{N}$, T^n is not zero because $T^n(x^n) = n!$.

The particularity of this counter example is that $\mathbb{F}[x]$ is an infinite dimensional vector space (over $\mathbb{F}$) and we claim that if we restrict v to be finite dimensional, then a linear operators with the above property is nilpotent.

Indeed, if we fix a basis $\{v_1, \ldots, v_k\}$, then for each $i \in [n]$, we have an integer n_i such that $T^{n_i}(v_i) = 0$. Letting $n = \max_{i \in [n]} n_i$, we obtain $T^n(v_i) = 0$ for any i. Then, by linearity and the fact that $\{v_1, \ldots, v_n\}$ spans V, we can conclude that $T^n = 0$. □

Exercise 8 (2.4)**.** Let V be a finite dimensional vector space over a field $\mathbb{F}$, W be a subspace of V and $v \notin W$. Show that there is a linear map $\phi : V \to \mathbb{F}$ such that $\phi(v) = 1$ and $\phi(w) = 0$ for all $w \in W$.

Solution. Recall that to define a linear map $\phi : V \to \mathbb{F}$, it is enough to describe where ϕ sends a basis of V. We construct a basis that will make it easy to define ϕ with the desired properties. First, let $\{w_1, \ldots, w_n\}$ be a basis of W, since $v \notin W$, v is independent of this basis and the set $\{w_1, \ldots, w_n, v\}$ is still independent. Then, we can complete the basis with vectors $\{v_1, \ldots, v_m\}$. Finally, we define ϕ by sending each w_i to 0, v to 1 and each v_i to 0 (this value does not really matter). □

Exercise 9 (N)**.** Fix the bases $\{1, x, x^2, x^3\}$ and $\{1\}$ of $\mathbb{P}_3$ and $\mathbb{R}$ respectively. Give the matrix entries of the linear transformation $T : \mathbb{P}_3 \to \mathbb{R}$ given by

$$p(x) \mapsto \int_{-1}^{8} p(x)dx.$$

Solution. The representation of T with respect to these bases is

$$\begin{bmatrix} T(1) & T(x) & T(x^2) & T(x^3) \end{bmatrix},$$

so the entries are straightforward to compute. We find $T(1) = 9$, $T(x) = \frac{63}{2}$, $T(x^2) = \frac{513}{3}$ and $T(x^3) = \frac{4095}{4}$. □

Exercise 10 (4.1)**.** Which matrices are only similar to themselves?

Solution. A matrix M is similar to another matrix N if and only if there exist an invertible matrix P such that $PMP^{-1} = N$. Thus, a matrix is only similar to itself if and only if for any invertible matrix P, $PMP^{-1} = M$. So the question is in fact asking which matrices is in the center of $GL_n(\mathbb{F})$. You have already seen that matrices that commute with every other matrices are scalar matrices, namely diag$(\lambda, \ldots, \lambda)$ for $\lambda \in \mathbb{F}$. We claim that this is also the center of $GL_n(\mathbb{F})$.

Let M be in the center of $GL_n(\mathbb{F})$. A first step towards the proof is to notice that permutation matrices have an inverse (the matrix corresponding to the inverse permutation), so they belong to $GL_n(\mathbb{F})$. We infer that for any such matrix P, we have $PMP^{-1} = PP^{-1}M = M$. Now, choosing a permutation σ, we notice that conjugation by P_σ will permute the rows according to σ and the columns according to σ^{-1}.

Consider the simple involution $(i\ j)$ and its corresponding matrix $P_{i,j}$, since $(i\ j)^2$ is the identity, $P_{i,j}^{-1} = P_{i,j}$, so we have $P_{i,j}MP_{i,j} = M$. Comparing the coordinates of both sides, we can conclude

$$M[i][j] = M[j][i] \text{ and } M[i][i] = M[j][j].$$

Since this works for any i, j. We conclude that M is symmetric and all the element of its diagonal are the same. Now, if we were working with $\mathbb{F} = \mathbb{R}$, since any symmetric matrix can be diagonalized, we would be done (why?).

In the general case, the above suggests that finding other ways to extract coordinates of M will help completing the proof. Another simple kind of matrices are the elements of the standard basis of $M_n(\mathbb{F})$, namely $E_{i,j}$ that has a one at coordinate (i, j) and zero everywhere else. Notice that $E_{i,j}M$ extracts the i-th row of M and $ME_{i,j}$ extracts the

j-th column, thus if we have $E_{i,j}M = ME_{i,j}$, comparing the contents of both matrices lets us infer that all the elements but the diagonal are zero. Since we already know all the diagonal elements are equal (from above), we conclude that M is a scalar matrix.

The problem with that argument is that $E_{i,j}$ is clearly not invertible, so we cannot assume that M commutes with it. A small fix for that is to notice that $I_n + E_{i,j}$ is invertible. Indeed, if $i = j$, the determinant is two and otherwise we can use the formula for the determinant of upper/lower triangular matrices to see that the determinant is one. Finally, we can conclude with the following equivalence:

$$(I_n + E_{i,j})M = M(I_n + E_{i,j}) \Leftrightarrow M + E_{i,j}M = M + ME_{i,j} \Leftrightarrow E_{i,j}M = ME_{i,j}.$$

□

Exercise 11 (N)**.** Show that if E is nilpotent (i.e.: there exists $k \in \mathbb{N}$ such that $E^k = 0$), then $I_n + E$ is invertible.

Solution. We can find the inverse of $I_n - E$ using the following clever telescoping sum:

$$I_n = I_n - E^k = (I_n - E)(I + E + E^2 + \cdots + E^{k-1}).$$

Now, the result follows because if E is nilpotent, then $-E$ is nilpotent.

Notice that this can also be used to show that $I_n + E_{i,j}$ is invertible in the proof of the last exercise because $E_{i,j}^2 = 0$. □

Exercise 12 (4.6)**.** Let A be an invertible $n \times n$ matrix with integer entries. Show that A^{-1} has integer entries if and only if $\det A = \pm 1$.

Solution. ($\Rightarrow$) If a matrix has integer entries, then its determinant is also an integer, hence $\det A$ and $\det A^{-1}$ are both integers. Moreover, we have $\det A \det A^{-1} = \det(AA^{-1}) = \det I = 1$ and ± 1 are the only invertible integers.

($\Leftarrow$) One can view the inversion of a matrix as solving multiple systems of linear equation. Indeed, if A^{-1} has columns unknown columns $c_1, \ldots, c_n$, then we want to solve

$$A \begin{bmatrix} c_1 & \cdots & c_n \end{bmatrix} = I_n \Leftrightarrow Ac_i = e_i, \forall i \in [n],$$

where the e_i's are the standard basis vector. Now, we fix $i \in [n]$ and use Cramer's rule to solve the system $Ac_i = e_i$. Letting A_j be the matrix formed by replacing the j-th column by the vector e_i, we have

$$c_i[j] = \frac{\det A_j}{\det A}.$$

Notice that A_j has integer entries, so its determinant is an integer. We infer that $c_i[j]$ is also an integer because $\det A = \pm 1$. Since i and j were arbitrary, we conclude that all the entries of A^{-1} are integers. □

Exercise 13 (4.7)**.** Let A be an $n \times n$ matrix with real entries, $v_1, \ldots, v_n$ be an orthonormal basis of $\mathbb{R}^n$ (with respect to the usual inner product) and M be a new $n \times n$ matrix with

$$M_{i,j} = (Av_i) \cdot v_j.$$

Compute the determinant of M in terms of the determinant of A.

Solution. For any two vectors in $v, w \in \mathbb{R}^n$, another way to compute the dot product is $v \cdot w = w^t v$, where we are multiplying a $1 \times n$ and $n \times 1$ together. Thus, we have $M_{i,j} = v_j^t A v_i$ and if we let V be the matrix whose columns are the v_i's, then we have

$$M^t = V^t A V \Leftrightarrow \det(M) = \det(V)^2 \det(A).$$

We claim that $\det(V) = \pm 1$ implying that $\det(M) = \det(A)$. In fact V is an orthogonal matrix satisfying $V^t V = I_n$ because

$$(V^t V)_{i,j} = v_j^t v_i = v_i \cdot v_j = \delta_{i,j},$$

and thus its determinant must lie in $\{\pm 1\}$. $\square$

Exercise 14 (5.2)**.** Let $\Delta(c_0, \ldots, c_{n-1})$ be the characteristic polynomial of

$$C(c_0, \ldots, c_{n-1}) = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & & & \vdots & \\ & & & 0 & 1 \\ -c_0 & -c_1 & \cdots & -c_{n-2} & -c_{n-1} \end{bmatrix}.$$

Show that

$$\Delta(c_0, \ldots, c_{n-1}) = t^n + c_{n-1} t^{n-1} + \cdots + c_1 t + c_0.$$

Solution. We proceed by induction, the base case being $\Delta(c_0) = t + c_0$ or $\Delta(c_0, c_1) = t^2 + c_1 t + c_0$. Next, we use the Laplace decomposition along the first row to find the determinant of

$$C(c_0, \ldots, c_{n-1}) = \begin{bmatrix} t & -1 & & & \\ 0 & t & -1 & & \\ & & & \vdots & \\ & & & t & -1 \\ c_0 & c_1 & \cdots & c_{n-2} & t + c_{n-1} \end{bmatrix}.$$

We obtain

$$t\Delta(c_1, \cdots, c_{n-1}) + (-1)^{n+1}(-1)^{n-1} c_0 = t\Delta(c_1, \cdots, c_{n-1}) + c_0,$$

and the result readily follows. $\square$

Exercise 15 (6.1)**.** Let $A \in M_n(\mathbb{C})$ and $f(t) \in \mathbb{C}[t]$. Show that if λ is an eigenvalue of A, then $f(\lambda)$ is an eigenvalue of $f(A)$.

Solution. Write $f(t) = \sum_{i=1}^{m} a_i t^i$, then $f(A) = \sum_{i=1}^{m} a_i A^i$. Let v be an eigenvector associated to λ, that is $Av = \lambda v$, then

$$f(A)v = \sum_{i=1}^{m} a_i A^i v = \sum_{i=1}^{m} \lambda^i v = f(\lambda)v.$$

□

Exercise 16 (N)**.** Show that a $m \times m$ matrix T has at most n distinct eigenvalues.

Solution. Let $\{\lambda_1, \dots, \lambda_n\}$ be distinct eigenvalues of T with associated eigenvectors $\{v_1, \dots, v_n\} \subseteq \mathbb{F}^m$. We claim that $\{v_1, \dots, v_n\}$ are linearly independent and show it using induction. It is clear for $n = 1$, so we assume it is true for $n - 1$. Suppose there exists a non-trivial linear combination $\sum_{i=1}^{n} \alpha_i v_i = 0$, then we have

$$0 = T\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i T(v_i) = \sum_{i=1}^{n} \alpha_i \lambda_i v_i.$$

This also leads to

$$\sum_{i=1}^{n} \alpha_i \lambda_i v_i - \lambda_n \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n-1} \alpha_i \lambda_n v_i = 0.$$

By our induction hypothesis, we know $\{v_1, \dots, v_{n-1}\}$ is independent, so the above combination is trivial, namely, $\alpha_i \lambda_n = 0$. If $\lambda_n = 0$, then

$$\sum_{i=1}^{n} \alpha_i \lambda_i v_i = \sum_{i=1}^{n-1} \alpha_i \lambda_i v_i = 0,$$

and by independence of the vectors involved, we have $\alpha_i \lambda_i = 0$ for each $1 \leq i \leq n - 1$. Moreover, since each λ_i cannot be zero (they are distinct from λ_n), we must have $\alpha_i = 0$. We conclude that $\alpha_n \neq 0$, but this implies $v_n = 0$ because

$$0 = \sum_{i=1}^{n} \alpha_i v_i = 0 = \alpha_n v_n.$$

However, this contradicts the definition of an eigenvector (i.e.: it cannot be zero).

If $\lambda_n \neq 0$, then one of the λ_i's can be zero and we end up in the case studied above. If none of them are zero, then all the α_i's must be zero and we can conclude as we did above. Since we got n independent vectors, n cannot be bigger than m (the number of entries in each vector). □

Solution 2. Note that if λ is an eigenvalue of T, then $\ker(T - \lambda\mathbb{1})$ is a non-trivial subspace of $\mathbb{F}^m$. Moreover, if λ and μ are distinct eigenvalues, then

$$\ker(T - \lambda\mathbb{1}) - \ker(T - \mu\mathbb{1}) = \{0\}.$$

This holds because $T(v) = \lambda v = \mu v$ implies $(\lambda - \mu)v = 0$ implies $\lambda = \mu$ when v is non-zero.

Therefore, if $\{\lambda_1, \ldots, \lambda_n\}$ are distinct eigenvalues of T, then

$$\ker(T - \lambda_1\mathbb{1}) \oplus \cdots \oplus \ker(T - \lambda_n\mathbb{1})$$

is a subspace of $\mathbb{F}^m$ and its dimension is the sum of the dimension of its terms. Each term is a non-trivial subspace (has dimension ≥ 1), the dimension of the sum is at least n, this implies that $n \leq \dim(\mathbb{F}^m) = m$. □

Solution 3. The eigenvalues are precisely the roots of the characteristic polynomial which has degree m. Therefore, there can only be m of them. □

Exercise 17 (6.4)**.** Let A be an $n \times n$ matrix with complex entries satisfying $A^k = 0$ for some $k \in \mathbb{N}$.

(i) Find the eigenvalues of A.

(ii) Find the characteristic polynomial of A.

(iii) Show that $A^n = 0$.

Solution. (i) If λ is an eigenvalue of A, then by exercise 15, the λ^k is an eigenvalue of A^k. However, since $A^k = 0$, we must have $\lambda^k = 0$, or equivalently $\lambda = 0$ (because a field has no zero-divisors).

(ii) Recall that the roots of the characteristic polynomial of A are precisely the eigenvalues of A, thus $\Delta_A(t)$ will be a power of t. Since the degree of Δ_A is the size of the matrix, we infer that $\Delta_A(t) = t^n$ because any other polynomial will have other roots (we are working in an algebraically closed field).

(iii) Follows from the Cayley-Hamilton theorem.

□

Exercise 18 (N)**.** Give a general construction for two $n \times n$ matrices M and N that are not similar, but have the same characteristic polynomial.

Solution. To simplify our life, we can ensure that the matrices $tI_n - M$ and $tI_n - N$ have determinants that are easy to compute (for instance uppertriangular matrices). Moreover, recall from the previous exercise nilpotent matrices all have the same characteristic polynomial, so we will also try to have M and N nilpotent. We claim the following matrices have the desired properties.

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It is easy to see that both matrices have characteristic polynomial t^4 and since $M^2 = 0$ and $N^2 \neq 0$, we can conclude that these matrices are not similar. $\square$

Exercise 19 (N)**.** Diagonalize the following matrix.

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

Solution. First, since A acts on e_4 independently, it is enough to diagonalize

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

Then, we can find the characteristic polynomial $(t-2)(t_2)(t-1)$. Since there are only two eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 2$, we are not yet sure if this matrix is diagonalizable. We need to find three independent eigenvectors. For λ_1, we find $\begin{bmatrix} 0 & -1 & 1 \end{bmatrix}^t$ and for λ_2, we find

$$v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } v_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

In a more systematic way, one could find the eigenvectors of λ_2 by solving the following system.

$$\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \\ 2z \end{bmatrix}.$$

We find that v_1, v_2 and v_3 are independent, and so B is indeed diagonalizable with diagonal form $\operatorname{diag}(1,2,2)$. $\square$

Exercise 20 (N, max-min principle, very challenging question)**.** Let V be an inner product space (over $\mathbb{C}$), and $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ when represented with the standard basis with $\lambda_1 \leq \cdots \leq \lambda_n$. Show that

$$\lambda_j = \max_{\mathcal{S}} \min_{\substack{\psi \in \mathcal{S}^\perp \\ \|\psi\|=1}} \langle \psi, A\psi \rangle,$$

where the maximum is taken over subspaces $\mathcal{S} \subseteq V$ of dimension $n-j$.

Proof. Fix $j \in [n]$ and $\mathcal{S}$ a subspace of dimension $n-j$, we can decompose $V = \mathcal{S} \oplus \mathcal{S}^\perp$. Hence, the space perpendicular to $\mathcal{S}$ has dimension j and therefore must intersect the span of $e_j, \ldots, e_n$ which has dimension $n-j+1$. Let ψ be in this intersection, without loss of generality, it has norm one and we can write it as $\psi = \sum_{i=j}^{n} a_i e_i$. Note that the unit length of ψ yields

$$1 = \langle \psi, \psi \rangle = \sum_{i=j}^{n} |a_i|^2,$$

thus we infer (because $\lambda_j \geq \lambda_i$ for $i \geq j$)

$$\langle \psi, A\psi \rangle = \sum_{i=j}^{n} \lambda_i |a_i|^2 \leq \sum_{i=j}^{n} \lambda_j |a_i|^2 = \lambda_j.$$

This means that for any subspace $\mathcal{S}$ of dimension $n - j$,

$$\inf\{\langle \psi, A\psi \rangle \mid \psi \perp \mathcal{S}, \|\psi\| = 1\} \leq \lambda_j.$$

Now, since e_j has unit length, is perpendicular to the span of $\{e_{j+1}, \ldots, e_n\}$ which has dimension $n - j$ and is such that $\langle e_j, Ae_j \rangle = \lambda_j$, we can conclude the min-max principle

$$\sup\{\inf\{\langle \psi, A\psi \rangle \mid \psi \perp \mathcal{S}, \|\psi\| = 1\} \mid \dim(\mathcal{S}) = n - j\} = \lambda_j.$$

□

Exercise 21 (N)**.** Let M and N be two square matrices of dimension n over a field $\mathbb{F}$. Is is true that MN and NM have the same characteristic polynomial? What about their minimal polynomials? [Hint: look at the first line of the proof.]

Proof. We construct two new matrices:

$$A = \begin{bmatrix} tI_n & M \\ N & I_n \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} I_n & 0 \\ -N & tI_n \end{bmatrix}.$$

We also compute their product:

$$AB = \begin{bmatrix} tI_n - MN & tM \\ 0 & tI_n \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} tI_n & 0 \\ 0 & tI_n - NM \end{bmatrix}.$$

Observe that using the formula for the determinant of upper triangular block matrices that

$$\det(tI_n - MN)\det(tI_n) = \det(AB) = \det(BA) = \det(tI_n)\det(tI_n - NM).$$

We conclude after dividing by t^n that MN and NM have the same characteristic polynomial.

This is not true for the minimal polynomial as is inferred from the following counterexample:

$$M = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

□