Problem F02.10. Let T be a linear operator on a finite dimensional complex inner product space V such that $T^*T = TT^*$. Show that there is an orthonormal basis of V consisting of eigenvectors of B.

Solution. When T satisfies $T^*T = TT^*$, we call T *normal*.

We prove this by induction on the dimension of the space that T operates on. If T is operating on a 1-dimensional space, the claim is obvious.

Suppose the claim holds for any normal T operating on an $n-1$ dimensional space ($n \geq 2$). By the fundamental theorem of algebra, the characteristic polynomial of T^* has a root λ which is an eigenvalue of T^*. Let v be the corresponding non-zero eigenvector (wlog $||v|| = 1$). Then $v^\perp = \{x \in V : (x, v) = 0\}$ has dimension $n-1$. Also, if $x \in v^\perp$, then

$$(Tx, v) = (x, T^*v) = \overline{\lambda}(x, v) = 0.$$

Thus $v^\perp$ is T-invariant. Then the restriction of T to $v^\perp$ is a normal operator on an $n-1$ dimensional space. Then by our inductive hypothesis, there is an orthonormal basis $\{v_2, \ldots, v_n\}$ for $v^\perp$ consisting of eigenvectors of T. Then $\{v, v_2, \ldots, v_n\}$ is an orthonormal set with n elements and is thus a basis for V. It remains to prove that v is also an eigenvector of T and then we will have the required basis. Since T is normal, so is T^* and thus $T^* - cI$ for every $c \in \mathbb{C}$. Also, for any normal operator S and any vector x, we see

$$||Sx||^2 = (Sx, Sx) = (x, S^*Sx) = (x, SS^*x) = (S^*x, S^*x) = ||S^*x||^2 .$$

Then since v is an eigenvector of T^*, we have $(T^* - \lambda I)v = 0$. Then

$$0 = ||(T^* - \lambda I)v||^2 = ||(T^* - \lambda I)^*v||^2 = \left|\left|(T - \overline{\lambda}I)v\right|\right|^2 .$$

Thus $Tv = \overline{\lambda}v$ so v is also an eigenvector of T. Thus $\{v, v_2, \ldots, v_n\}$ is a basis of V consisting of eigenvectors of T. This completes the induction and the proof.

Problem W02.8. Let $T : V \to W$ and $S : W \to X$ be linear transformations of real finite dimensional vector spaces. Prove that

$$\text{rank}(T) + \text{rank}(S) - \dim(W) \leq \text{rank}(S \circ T) \leq \max\{\text{rank}(T), \text{rank}(S)\}.$$

Solution. By the Rank-Nullity Theorem,

$$\text{rank}(T) + \dim(\ker(T)) = \dim(V), \tag{1}$$
$$\text{rank}(S) + \dim(\ker(S)) = \dim(W), \tag{2}$$
$$\text{rank}(S \circ T) + \dim(\ker(S \circ T)) = \dim(V). \tag{3}$$

Adding the (1), (2) and then subtracting (3) gives

$$\text{rank}(T) + \text{rank}(S) - \text{rank}(S \circ T) + \dim(\ker(T)) + \dim(\ker(S)) - \dim(\ker(S \circ T)) = \dim(W).$$

Let $\{v_1, \ldots, v_\ell\}$ be a basis for $\ker(T)$. Then $(S \circ T)(v_i) = 0$ for each i so $\ker(T) \subset \ker(S \circ T)$. Thus we can extend this to a basis $\{v_1, \ldots, v_\ell, y_1, \ldots, y_k\}$ for $\ker(S \circ T)$. Then for each $j = 1, \ldots, k$, we have

$$0 = (S \circ T)(y_j) = S(T(y_j)).$$

Hence $T(y_j) \in \ker(S)$ for each j. Further if $a_1, \dots, a_k \in \mathbb{C}$ are such that

$$a_1 T(y_1) + \cdots + a_k T(y_k) = 0,$$

Then

$$T(a_1 y_1 + \cdots + a_k y_k) = 0$$

so $a_1 y_1 + \cdots + a_k y_k \in \ker(T)$ so there are $b_1, \dots, b_\ell \in \mathbb{C}$ such that

$$a_1 y_1 + \cdots + a_k y_k = b_1 v_1 + \cdots b_\ell v_\ell \quad \Longrightarrow \quad a_1 y_1 + \cdots + a_k y_k - b_1 v_1 - \cdots - b_\ell v_\ell = 0.$$

But these vectors form a basis for $\ker(S \circ T)$ so in particular, $a_1 = \cdots = a_k = 0$. Thus $\{T(y_1), \dots, T(y_k)\}$ is a linearly independent subset of $\ker(S)$ and so $\dim(\ker(S)) \geq k$. Hence

$$\dim(\ker(T)) + \dim(\ker(S)) \geq \dim(\ker(S \circ T))$$

and so the equation above yields

$$\text{rank}(T) + \text{rank}(S) - \text{rank}(S \circ T) \leq \dim(W)$$

or

$$\text{rank}(T) + \text{rank}(S) - \dim(W) \leq \text{rank}(S \circ T)$$

which is the first half of the inequality.

Now suppose that $\{x_1, \dots, x_m\}$ is a basis for $\text{im}(S \circ T)$. Then there are $u_1, \dots, u_m \in V$ such that $(S \circ T)(u_i) = x_i, i = 1, \dots, m$. Thus $S(T(u_i)) = x_i$ for each i, and so in particular $x_i \in \text{im}(S)$ for each i and so we have m linearly independent vectors in $\text{im}(S)$. This gives $\text{rank}(S \circ T) \leq \text{rank}(S) \leq \max\{\text{rank}(S), \text{rank}(T)\}$. This is the second half of the inequality.

Problem W02.10. Let V be a finite dimensional complex inner product space and $f : V \to \mathbb{C}$ a linear functional. Show that there exists a vector $w \in V$ such that $f(v) = (v, w)$ for all $v \in V$.

Solution. Let $v_1, \dots, v_n$ be an orthonormal basis for V. Given $f \in V^*$, put $f(v_i) = \alpha_i \in \mathbb{C}$. Next set

$$w = \overline{\alpha_1} v_1 + \cdots + \overline{\alpha_n} v_n.$$

For any $v \in V$, there are $\beta_1, \dots, \beta_n \in \mathbb{C}$ such that

$$v = \beta_1 v_1 + \cdots + \beta_n v_n.$$

Then

$$f(v) = \beta_1 f(v_1) + \cdots + \beta_n f(v_n) = \beta_1 \alpha_1 + \cdots + \beta_n \alpha_n$$

and

$$(v, w) = \left(\sum_{i=1}^n \beta_i v_i, \sum_{j=1}^n \overline{\alpha_j} v_j \right) = \sum_{i=1}^n \sum_{j=1}^n (\beta_i v_i, \overline{\alpha_j} v_j) = \sum_{i=1}^n \sum_{j=1}^n \beta_i \alpha_j (v_i, v_j).$$

Thus by orthonormality,

$$(v, w) = \sum_{i=1}^n \beta_i \alpha_i = f(v).$$

Since v was arbitrary, $f(v) = (v, w)$ for all $v \in V$.

Problem W02.11. Let V be a finite dimensional complex inner product space and $T : V \to V$ a linear transformation. Prove that there exists an orthonormal ordered basis for V such that the matrix representation of T in this basis is upper triangular.

Solution. We prove this by induction on the dimension of the space T acts upon. If T is acting on a 1-dimensional space, the claim is obvious.

Suppose the claim holds for linear maps acting on $n-1$ dimensional spaces. Let $\dim(V) = n$. By the fundamental theorem of algebra, there is an eigenvalue $\lambda \in \mathbb{C}$ and corresponding non-zero eigenvector $0 \neq v_1 \in V$ of T^*; wlog $||v_1|| = 1$. Then $v_1^\perp = \{x \in V : (x, v_1) = 0\}$ is an $n-1$-dimensional space. Also for $x \in v_1^\perp$, we have

$$(T(x), v_1) = (x, T^*(v_1)) = (x, \lambda v_1) = \overline{\lambda}(x, v_1) = 0.$$

Hence $v_1^\perp$ is T-invariant. Thus $T\big|_{v_1^\perp}$ is an operator acting on an $n-1$-dimensional space. By our inductive hypothesis, there is an orthonormal basis $\{v_2, \dots, v_n\}$ for $v^\perp$ such that the matrix of T is upper-triangular with respect to this basis. Then $\{v_1, v_2, \dots, v_n\}$ is an orthonormal basis for V. Further, $T(v_1) = \sum_{j=1}^n a_{1j} v_j$ for some $a_{1j} \in \mathbb{C}$ and by assumption $T(v_i) = \sum_{j=i}^n a_{ij} x_j$. Thus the matrix of T with respect to this basis is

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Problem S03.8. Let V be an n-dimensional complex vector space and $T : V \to V$ a linear operator. Suppose that the characteristic polynomial of T has n distinct roots. Show that there is a basis B of V such that the matrix representation of T in the basis B is diagonal.

Solution. Since each root of the characteristic polynomial (and thus each eigenvalue of T) is distinct and since eigenvectors corresponding to different eigenvalues are linearly independent, each eigenspace E_λ is a one-dimensional T-invariant subspace. Let $\lambda_1, \dots, \lambda_n$ be the distinct eigenvalues of T with corresponding eigenvectors $v_1, \dots, v_n$. We know that eigenvectors corresponding to distinct eigenvalues are linearly independent, thus $E_{\lambda_i} \cap E_{\lambda_j} = \{0\}$ whenever $i \neq j$. Further, since we have n-linearly independent vectors, $\{v_1, \dots, v_n\}$ is a basis for V. The matrix of T with respect to this basis is

$$[T] = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & 0 & \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & \lambda_n \end{bmatrix},$$

since $T(v_i) = \lambda_i v_i, i = 1, \dots, n$.

Problem S03.9. Let $A \in M_3(\mathbb{R})$ satisfy $\det(A) = 1$ and $A^t A = I = AA^t$ where I is the identity matrix. Prove that the characteristic polynomial of A has 1 as a root.

Solution. Clearly the characteristic polynomial of A has a real root since it has odd order. Let λ be a real root of the characteristic polynomial. Then λ is an eigenvalue of A. Suppose $0 \neq v \in \mathbb{R}^3$ is a normalized eivengector corresponding to λ. Then

$$\lambda^2 = \lambda^2(v,v) = (\lambda v, \lambda v) = (Av, Av) = (v, A^t A v) = (v,v) = 1.$$

Thus $\lambda = \pm 1$. If $\lambda = 1$, then we are done. If $\lambda = -1$, suppose $\mu, \nu \in \mathbb{C}$ are the other eigenvalues of A. Then $-1 = -\det(A) = -\mu\nu\lambda = \mu\nu$. If μ, ν are not real, they must be a conjugate pair since A is real. But this is impossible, because then $\mu\nu \geq 0$. Thus both μ, ν are real. By the same reasoning as above, $\mu, \nu = \pm 1$. Then $\mu\nu = -1$ forces $\mu = 1, \nu = -1$ (or vice versa). Thus A has 1 as an eigenvalue and so the characteristic polynomial of A has 1 as a root.

Problem S03.10. Let V be a finite dimensional real inner product space and $T : V \to V$ a hermitian linear operator. Suppose the matrix representation of T^2 in the standard basis has trace zero. Prove that T is the zero operator.

Solution. Let $\dim(V) = n$ and let A be the matrix of T in the standard basis. Since T is hermitian, so is A and thus by the spectral theorem, there is an orthonormal basis $\{v_1, \dots, v_n\}$ for $\mathbb{R}^n$ consisting of eigenvectors of A. Let $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ be the corresponding eigenvalues (repeats are allowed and the eigenvalues are real since A is hermitian). Then

$$Av_i = \lambda_i v_i \quad \Longrightarrow \quad A^2 v_i = \lambda_i A v_i = \lambda_i^2 v_i.$$

Thus (λ_i^2, v_i) is an eigenpair for A^2 which is the matrix of T^2. We are given that the trace of the matrix is zero, but the trace is the sum of the eigenvalues. Hence

$$\sum_{i=1}^n \lambda_i^2 = 0 \quad \Longrightarrow \quad \lambda_1 = \cdots = \lambda_n = 0;$$

again this holds since all eigenvalues of a hermitian operator are real. Then $Av_i = 0, i = 1, \dots, n$. But this means A sends a basis for $\mathbb{R}^n$ to zero. This is only possible if A is the zero matrix. Thus T is the zero transformation.

Problem F03.9. Consider a 3×3 real symmetric matrix with determinant 6. Assume that $(1, 2, 3)$ and $(0, 3, -2)$ are eigenvectors with corresponding eigenvalues 1 and 2 respectively.

(a) Give an eigenvector of the form $(1, x, y)$ which is linearly independent from the two vectors above.

(b) What is the eigenvalue of this eigenvector?

Solution. We answer the questions in the reverse order. The product of the eigenvalues equals the determinant, so the third eigenvalue is 3. This answers (b).

By the spectral theorem, the eigenspaces corresponding to distinct eigenvalues will be orthogonal. Here all eigenvalues are distinct. Since the first two eigenvectors span a two dimensional space, any vector orthogonal to both will necessarily be a third eigenvector. Taking the cross product of the two vectors gives a vector which is orthogonal to both. We see

$$(1,2,3)\times(0,3,-2) = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 1 & 2 & 3 \\ 0 & 3 & -2 \end{vmatrix} = (-13,2,3).$$

Then, $v = (1,-2/13,3/13)$ is an eigenvector of the desired form. This answers (a).

Problem F03.10.

(a) Take $t \in \mathbb{R}$ such that t is not an integer multiple of π. Prove that if

$$A = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

then there is no invertible real matrix B such that BAB^{-1} is diagonal.

(b) Do the same for

$$A = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$$

where $\lambda \in \mathbb{R} - \{0\}$.

Solution.

(a) A doesn't have any real eigenvalues so it cannot be diagonalizable in $M_2(\mathbb{R})$. Indeed, any eigenvalue λ of A would need to satisfy

$$(\cos(t)-\lambda)^2 + \sin^2(t) = 0 \quad \Longrightarrow \quad \cos(t)-\lambda = \pm i\sin(t) \quad \Longrightarrow \quad \lambda = \cos(t) \pm i\sin(t)$$

which is not real since t is not an integer multiple of π.

(b) The only eigenvalue of A is 1 and it has algebraic multiplicity 2. However, the only eigenvalue corresponding to this eigenvalue (up to scaling) is $v = (1,0)$. Thus the geometric multiplicity of the eigenvalue is 1. Since A has an eigenvalues whose algebraic and geometric multiplicities are unequal, A is not diagonalizable.

Problem S04.7. Let V be a finite dimensional real vector space and $U, W \subset V$ be subspaces of V. Show both of the following:

(a) $U^0 \cap W^0 = (U+W)^0$

(b) $(U \cap W)^0 = U^0 + W^0$

Solution.

(a) Let $f \in U^0 \cap W^0$. Then $f \in U^0$ and $f \in W^0$. Take $x \in U + W$. Then $x = u + w$ for some $u \in U, w \in W$. We see

$$f(x) = f(u + w) = f(u) + f(w) = 0 + 0 = 0.$$

Thus $f \in (U + W)^0$.

Now take $f \in (U + W)^0$. For any $u \in U$, we have $u \in U + W$. Then $f(u) = 0$. Hence, since u was arbitrary, $f \in U^0$. Similarly, $f(w) = 0$ for all $w \in W$ so $f \in W^0$. Thus $f \in U^0 \cap W^0$.

We conclude that $U^0 \cap W^0 = (U + W)^0$.

(b) Let $f \in (U \cap W)^0$. For any $Z \subset V$, define $f_Z \in V^*$ by $f_Z(v) = f(v), v \in Z$, $f_Z(v) = 0, v \in V - Z$. Then

$$f = f_U + f_{V-U}.$$

Clearly $f_{V-U} \in U^0$. We must show $f_U \in W^0$. Let $w \in W$. If $w \in U$, then $f_{V-U}(w) = 0$, and $f(w) = 0$ since $w \in U \cap W$. Thus $0 = f(w) = f_U(w) + f_{V-U}(w) = f_U(w)$. Otherwise $w \notin U$, in which case $f_U(w) = 0$ by definition. Hence, $f_U(w) = 0$ for all $w \in W$ and so $f_U \in W^0$. Then

$$f = f_U + f_{V-U} \in U^0 + W^0.$$

Take $f \in U^0 + W^0$. Then $f = f_1 + f_2$, $f_1 \in U^0, f_2 \in W^0$. For any $v \in U \cap W$, we have $v \in U$ and $v \in W$. Then

$$f(v) = f_1(v) + f_2(v) = 0 + 0 = 0.$$

Hence $f(v) = 0$ for all $v \in U \cap W$ so $f \in (U \cap W)^0$.

We conclude $(U \cap W)^0 = U^0 + W^0$.

Problem S04.9. Let V be a finite dimensional real inner product space and $T : V \to V$ a linear operator. Show the following are equivalent:

(a) $(Tx, Ty) = (x, y)$ for all $x, y \in V$,

(b) $||Tx|| = ||x||$ for all $x \in V$,

(c) $T^*T = I$, where T^* is the adjoint of T and $I : V \to V$ is the identity map,

(d) $TT^* = I$.

Solution. (a) $\implies$ (b) by taking $x = y$.

Assume (b) is true. Then for any $x, y \in V$,

$$\begin{aligned}(x,x) + 2(x,y) + (y,y) &= (x+y, x+y) \\ &= (Tx + Ty, Tx + Ty) \\ &= (Tx, Tx) + 2(Tx, Ty) + (Ty, Ty) \\ &= (x,x) + 2(Tx, Ty) + (y,y).\end{aligned}$$

Thus $2(x,y) = 2(Tx,Ty)$ and so (b) $\implies$ (a).

Assume (a), (b) are true. Consider, for any $v \in V$,

$$\begin{aligned}((T^*T-I)v,(T^*T-I)v) &= (T^*Tv-v,T^*Tv-v)\\ &= (T^*Tv,T^*Tv)-(T^*Tv,v)-(v,T^*Tv)+(v,v)\\ &= (T^*Tv,T^*Tv)-(Tv,Tv)-(Tv,Tv)+(v,v)\\ &= (T^*Tv,T^*Tv)-(v,v)-(v,v)+(v,v) \quad \text{[by (a)]}\\ &= (T^*Tv,T^*Tv)-(v,v)\\ &= (Tv,TT^*Tv)-(v,v)\\ &= (v,T^*Tv)-(v,v) \quad \text{[by (a)]}\\ &= (Tv,Tv)-(v,v) = 0 \quad \text{[by (b)]}.\end{aligned}$$

Thus $(T^*T-I)v = 0$ for all $v \in V$ so $T^*T = I$. Thus (b) $\implies$ (c).

Assume (c) is true. Then for any $v \in V$,

$$(v,v) = (v,Iv) = (v,T^*Tv) = (Tv,Tv).$$

Hence (c) $\implies$ (b).

Thus far we have that (a),(b),(c) are equivalent. Assume these hold.Then (a) implies

$$(TT^*x,TT^*y) = (T^*x,T^*y) = (TT^*x,y) \quad \implies \quad (TT^*x,(TT*-I)y) = 0$$

for all $x,y \in V$. From (c) we see that T,T^* are bijective, thus so is TT^*. Hence $(TT^*-I)y$ is orthogonal to all of V so $(TT^*-I)y = 0$. However y was arbitrary, so this implies that $TT^*-I = 0$ so $TT^* = I$. Thus (a),(b),(c) imply (d).

Assume (d) is true. Then for any $x,y \in V$, we have

$$(x,y) = (x,Iy) = (x,TT^*y) = (T^*x,T^*x).$$

Then

$$(T^*Tx,T^*Ty) = (Tx,Ty) = (T^*Tx,y) \quad \implies \quad (T^*Tx,(T^*T-I)y) = 0.$$

This implies (c) (and thus (a),(b)) by the same reasoning as above.

Problem F04.10. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and for $\lambda \in \mathbb{C}$, define the subspace $V(\lambda) = \{v \in V : (T-\lambda I)^n v = 0 \text{ for some } n \geq 1\}$. This is called the generalized eigenspace of λ.

1. Prove there for each $\lambda \in \mathbb{C}$ there is a fixed $N \in \mathbb{N}$ such that $V(\lambda) = \ker\left((T-\lambda I)^N\right)$.
2. Prove that if $\lambda \neq \mu$ then $V(\lambda) \cap V(\mu) = \{0\}$. Hint: use the fact that

$$I = \frac{T-\lambda I}{\mu-\lambda} + \frac{T-\mu I}{\lambda-\mu}.$$

Solution.

1. Let $v_1, \ldots, v_k$ be a basis for $V(\lambda)$. For $i = 1, \ldots, k$, define $N_i \in \mathbb{N}$ to be equal to the least $n \in \mathbb{N}$ such that $(T - \lambda I)^n v_i = 0$. Take $N = \max\{N_1, \ldots, N_k\}$. Then $(T - \lambda I)^N v_i = 0$ for all $i = 1, \ldots, k$. Since we can build any vector in $V(\lambda)$ from a linear combination of $v_1, \ldots, v_k$, we see $V(\lambda) = \ker\left((T - \lambda I)^N\right)$.

2. By part 1., there are $N_1, N_2 \in \mathbb{N}$ such that
$$V(\lambda) = \ker\left((T - \lambda I)^{N_1}\right) \quad \text{and} \quad V(\mu) = \ker\left((T - \mu I)^{N_2}\right).$$
Take $M = \max\{N_1, N_2\}$. We see that
$$I = I^{2M} = \left(\frac{T - \lambda I}{\mu - \lambda} + \frac{T - \mu I}{\lambda - \mu}\right)^{2M}$$
Since $(T - \lambda I)$ and $(T - \mu I)$ commute with each other, by the binomial theorem, if we expand the right hand side above, each summand will have a factor of $(T - \lambda I)$ or $(T - \mu I)$ which has a power greater than or equal to M. Thus if $v \in V(\lambda) \cap V(\mu)$, then
$$v = Iv = \left(\frac{T - \lambda I}{\mu - \lambda} + \frac{T - \mu I}{\lambda - \mu}\right)^{2M} v = 0.$$
Hence $V(\lambda) \cap V(\mu) = \{0\}$.

Problem S05.1. Given $n \geq 1$, let $\mathrm{tr} : M_n(\mathbb{C}) \to \mathbb{C}$ denote the trace operator:
$$\mathrm{tr}(A) = \sum_{k=1}^{n} A_{kk}, \quad A \in M_n(\mathbb{C}).$$

(a) Determine a basis for the kernel of tr.

(b) For $X \in M_n(\mathbb{C})$, show that $\mathrm{tr}(X) = 0$ if and only if there are matrices $A_1, \ldots, A_m$, $B_1, \ldots, B_m \in M_n(\mathbb{C})$ such that
$$X = \sum_{j=1}^{m} A_j B_j - B_j A_j.$$

Solution.

(a) Since $\dim(M_n(\mathbb{C})) = n^2$ and since the range of tr is $\mathbb{C}$ which has dimension 1 as a vector space over itself, we know from the Rank-Nullity theorem that $\dim\ker(\mathrm{tr}) = n^2 - 1$. Thus to find a basis for the kernel, it is sufficient to find $n^2 - 1$ linearly independent matrices on which the trace operator vanishes. For $i = 1, \ldots, n, j = 1, \ldots n,$, define $E_{ij} \in M_n(\mathbb{C})$ be such that the entry in the i^{th} row and j^{th} column is 1 and the other entries are 0. Then the matrices $E_{ij}, i \neq j$ are clearly linearly independent and have 0 trace. This constitutes $n^2 - n$ linearly independent matrices with 0 trace. For $n - 1$ more, consider $F_k = E_{11} - E_{kk}, k = 2, \ldots, n$. These are also linearly independent. Thus the collection
$$\{E_{ij} : i, j \in \{1, \ldots, n\}, i \neq j\} \cup \{F_k : k = 2, \ldots, n\}$$
forms a basis for the kernel of tr.

(b) It is well-known that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, $A, B \in M_n(\mathbb{C})$. Thus if X has the specified form,

$$\mathrm{tr}(X) = \sum_{j=1}^{m} \mathrm{tr}(A_jB_j) - \mathrm{tr}(B_jA_j) = \sum_{j=1}^{m} \mathrm{tr}(A_jB_j) - \mathrm{tr}(A_jB_j) = 0.$$

Conversely, if X has trace 0, we must have some complex numbers $\alpha_{ij}, 1 \le i, j \le n, i \neq j$ and $\beta_k, k = 2, \ldots, n$ such that

$$X = \sum_{i,j=1, i\neq j}^{n} \alpha_{ij}E_{ij} + \sum_{k=2}^{n} \beta_k F_k.$$

Further, we see that

$$E_{ij} = E_{ii}E_{ij} \quad \text{and} \quad 0 = E_{ij}E_{ii}.$$

Thus

$$\alpha_{ij}E_{ij} = E_{ii}(\alpha_{ij}E_{ij}) - (\alpha_{ij}E_{ij})E_{ii}, \quad i, j = 1, \ldots, n, i \neq j.$$

Further

$$F_k = E_{k1}E_{1k} - E_{1k}E_{k1}, \quad k = 2, \ldots, n.$$

Thus

$$X = \sum_{i,j=1, i\neq j}^{n} E_{ii}(\alpha_{ij}E_{ij}) - (\alpha_{ij}E_{ij})E_{ii} + \sum_{k=2}^{n} (\beta_k E_{k1})E_{1k} - E_{1k}(\beta_k E_{k1}).$$

This is the desired form up to renaming matrices and indices.

Problem S05.2. Let V be a finite-dimensional vector space and let V^* denote the dual space of V. For a set $W \subset V$, define

$$W^\perp = \{f \in V^* : f(w) = 0 \text{ for all } w \in W\} \subset V^*$$

and for a set $U \subset V^*$, define

$$^\perp U = \{v \in V : f(v) = 0 \text{ for all } f \in U\} \subset V.$$

(a) Show that for any $W \subset V$, $^\perp(W^\perp) = \mathrm{Span}\,(W)$.

(b) Let W be a subspace of V. Give an explicit isomorphism bewteen $(V/W)^*$ and $W^\perp$. Show that it is an isomorphism.

Solution.

(a) Let $W \subset V$. Take $x \in \mathrm{Span}\,(W)$. Then there are scalars $\alpha_1, \ldots, \alpha_n$ and vectors $w_1, \ldots, w_n \in W$ such that

$$x = \alpha_1 w_1 + \cdots + \alpha_n w_n.$$

Then for any $f \in W^\perp$ we have

$$f(x) = \alpha_1 f(w_1) + \cdots + \alpha_n f(w_n) = \alpha_1 \cdot 0 + \cdots + \alpha_n \cdot 0 = 0.$$

Since f was arbitrary, $f(x) = 0$ for all $f \in W^\perp$. Hence by definition $x \in^\perp (W^\perp)$. Thus $\text{Span}\,(W) \subset^\perp (W^\perp)$.

Now take $x \in W$, then for all $f \in W^\perp$, we have $f(x) = 0$. Hence, $x \in^\perp (W^\perp)$. However, $\text{Span}\,(W)$ is the smallest vector space containing W. Thus ${}^\perp(W^\perp) \subset \text{Span}\,(W)$.

Thus ${}^\perp(W^\perp) = \text{Span}\,(W)$.

(b) Define $\phi : (V/W)^* \to W^\perp$ by

$$\phi(f) = g_f, \quad f \in (V/W)^*$$

where g_f is the functional which sends x to $f(x+W)$. Then for any $w \in W$, we have

$$[\phi(f)](w) = g_f(w) = f(w+W) = f(0) = 0.$$

Thus $\phi(f)$ is indeed in $W^\perp$ when $f \in (V/W)^*$. Further, if $\phi(f) = 0$, then

$$[\phi(f)](x) = 0, \quad \Longrightarrow \quad f(x+W) = 0 \text{ for all } x \in V.$$

Hence f is the zero functional. Thus $\phi(f) = 0$ implies $f = 0$ which tells us that ϕ is injective. Then since $\dim((V/W)^*) = \dim(V/W) = \dim(V) - \dim(W) = \dim(W^\perp)$, the injectivity of ϕ implies bijectivity, so ϕ is an isomorphism.

Problem S05.3. Let A be a Hermitian-symmetric $n \times n$ complex matrix. Show that if $(Av, v) \geq 0$ for all $v \in \mathbb{C}^n$ then there exists and $n \times n$ matrix T such that $A = T^*T$.

Solution. By the spectral theorem, there is an orthonormal basis $\{v_1, \dots, v_n\}$ of $\mathbb{C}^n$ consisting of eigenvectors of A. Let $\lambda_1 \dots, \lambda_n$ be the corresponding eigenvalues. Since A is Hermitian, all λ_i are real. Further, for each i,

$$0 \leq (Av_i, v_i) = (\lambda_i v_i, v_i) = \lambda_i (v_i, v_i) = \lambda_i.$$

Thus we can define a matrix T such that $Tx_i = \sqrt{\lambda_i} x_i$ (recall that a linear operator is uniquely identified by how it treats basis vectors). Then $T^* x_i = \overline{\sqrt{\lambda_i}} x_i = \sqrt{\lambda_i} x_i$ since T is clearly diagonal with respect to this basis. Then

$$Ax_i = \lambda_i x_i = \sqrt{\lambda_i}\left(\sqrt{\lambda_i} x_i\right) = \sqrt{\lambda_i} T^* x_i = T^*\left(\sqrt{\lambda_i} x_i\right) = T^*(Tx_i) = (T^*T)x_i.$$

Then since A and T^*T agree on basis elements, they are the same matrix.

Problem S05.4. We say that $\mathcal{I} \subset M_n(\mathbb{C})$ is a two-sided ideal in $M_n(\mathbb{C})$ if

(i) for all $A, B \in \mathcal{I}$, $A + B \in \mathcal{I}$,

(ii) for all $A \in \mathcal{I}$ and $B \in M_n(\mathbb{C})$, AB and BA are in $\mathcal{I}$.

Show that the only two-sided ideals in $M_n(\mathbb{C})$ are $\{0\}$ and $M_n(\mathbb{C})$ itself.

Solution. Let $\mathcal{I}$ be a two sided ideal in $M_n(\mathbb{C})$. Suppose there is a non-zero matrix $A \in \mathcal{I}$. Then by multiplying A by elementary matrices we can transform $A \to \overline{A}$ which has a non-zero entry in its first row and column and is still in $\mathcal{I}$. Next, by multiplying on the left and right by a matrix B such that $(B)_{11} = 1$ and $(B)_{ij} = 0$ otherwise. We arrive at a matrix $A^* \in \mathcal{I}$ which is zero everywhere except for a nonzero entry in the first row and column. Again, multiplying by an elementary matrix, we can reduce A^* to a matrix $A_{11} \in \mathcal{I}$ which has a 1 in the first row and first column and zeroes elsewhere. Performing similar steps, we can create matrices $A_{ii} \in \mathcal{I}$ which have a 1 in the i^{th} row and i^{th} column and zeroes elsewhere. Adding all these matrices together, we see have the identity matrix $I \in \mathcal{I}$. Then for any $C \in M_n(\mathbb{C})$, we have $CI = C \in \mathcal{I}$. Hence $I = M_n(\mathbb{C})$. Thus the only two-sided ideals in $M_n(\mathbb{C})$ are $\{0\}$ and $M_n(\mathbb{C})$.

Problem F05.6.

(a) Prove that if P is a real-coefficient polynomial and A a real-symmetric matrix, then λ is an eigenvalue of A if and only if $P(\lambda)$ is an eigenvalue of $P(A)$.

(b) Use (a) to prove that if A is real symmetric, then A^2 is non-negative definite.

(c) Check part (b) by verifying directly that $\det(A^2)$ and $\text{trace}(A^2)$ are non-negative when A is real-symetric.

Solution.

(a) Suppose λ is an eigenvalue of A. Then there is an eigenvector $v \neq 0$ such that $Av = \lambda v$ Then

$$A^2v = A(Av) = A(\lambda v) = \lambda Av = \lambda^2 v.$$

Similarly, $A^k v = \lambda^k v$ for $k = 3, 4, 5, \ldots$. Let P be the polynomial

$$P(x) = a_0 + a_1 x + \cdots + a_m x^m.$$

Then

$$\begin{aligned} P(A)v &= a_0 Iv + a_1 Av + \cdots + a_m A^m v \\ &= a_0 v + a_1 \lambda v + \cdots + a_m \lambda^m v \\ &= (a_0 + a_1\lambda + \cdots + a_m \lambda^m)v = P(\lambda)v. \end{aligned}$$

So $P(\lambda)$ is an eigenvalue of $P(A)$ (note, we didn't need the assumption that A is real-symmetric for this direction).

By the spectral theorem, A is (orthogonally) similar to a diagonal matrix D. Say $A = UDU^*$ where U is orthogonal. Then

$$\begin{aligned} P(A) &= a_0 I + a_1 UDU^* + \cdots + a_m (UDU^*)^m \\ &= a_0 UU^* + a_1 UDU^* + \cdots + a_m UD^m U^* \\ &= U(a_0 I + a_1 D + \cdots + a_m D^m)U^* = UP(D)U^*. \end{aligned}$$

Thus $P(A)$ is similar to $P(D)$. Now if $P(\lambda)$ is an eigenvalue of $P(A)$ then it is also an eigenvalue of $P(D)$. But $P(D)$ is diagonal so $P(\lambda)$ lies on the diagonal of $P(D)$ and so λ lies on the diagonal of D and thus is an eigenvalue of D. Then λ is also an eigenvalue of A since A is similar to D (note, this argument also didn't require that A is symmetric, just that A is diagonalizable).

(b) If A is $n \times n$, then there is an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors $\{v_1, \ldots, v_n\}$ of A. Suppose the corresponding eigenvalues are $\lambda_1, \ldots, \lambda_n$ (possibly with repititions). All these eigenvalues are real since A is symmetric. Also, $\lambda_1^2, \ldots, \lambda_n^2$ are eigenvalues of A^2 corresponding to the same eigenvalues. Finally, for any $x \in \mathbb{R}^n$, $x = \alpha_1 v_1 + \cdots + \alpha_n v_n$ for some $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Thus gives

$$x^t A^2 x = (\alpha_1 v_1 + \cdots + \alpha_n v_n)^t (\alpha_1 \lambda_1^2 v_1 + \cdots + \alpha_n \lambda_n^2 v_n) = \sum_{i=1}^n \alpha_i^2 \lambda_i^2, \quad \text{by orthogonality.}$$

this last expression is clearly non-negative since all λ_i and α_i are real. Thus A^2 is non-negative definite.

(c) We see

$$\det(A^2) = \det(AA) = \det(A)\det(A) = (\det(A))^2 \geq 0.$$

Also, the entries of A^2 are

$$(A^2)_{ij} = \sum_{k=1}^n a_{ik} a_{kj}.$$

Thus the diagonal entries are

$$(A^2)_{ii} = \sum_{k=1}^n a_{ik} a_{ki} = \sum_{k=1}^n a_{ik}^2, \quad \text{by symmetry.}$$

Thus all diagonal entries of A^2 are non-negative and so the trace is non-negative.

Problem F05.9. Suppose U, W are subspaces of a finite-dimensional vector space V.

(a) Show that $\dim(U \cap W) = \dim(U) + \dim(W) - \dim(\mathrm{Span}\,(U, W))$.

(b) Let $n = \dim(V)$. Use part (a) to show that if $k < n$ then an intersection of k subspaces of dimension $n-1$ always has dimension at least $n-k$.

Solution.

(a) This is called the dimension formula and is proven as follows.

Let $\{x_1, \ldots, x_k\}$ be a basis for $U \cap W$. Extend this separately to a basis

$$\{x_1, \ldots, x_k, u_1, \ldots, u_\ell\} \text{ of } U \text{ and}$$
$$\{x_1, \ldots, x_k, w_1, \ldots, w_m\} \text{ of } W.$$

Then $\dim(U \cap W) = k$, $\dim(U) = k + \ell$ and $\dim(W) = k + m$. So it remains to prove that $\dim(\operatorname{Span}(U, W)) = k + \ell + m$; the result will follow. To do this, we just throw all the vectors together and if there is any justice in the world

$$\{x_1, \ldots, x_k, u_1, \ldots, u_\ell, w_1, \ldots, w_m\}$$

will form a basis for $\operatorname{Span}(U, W)$. Take any $y \in \operatorname{Span}(U, W)$. Then y can be built from vectors in U and W. But each of those vectors can be built by basis vectors of U and W and so y can be built using the basis vectors of U and W. That is

$$\{x_1, \ldots, x_k, u_1, \ldots, u_\ell\} \cup \{x_1, \ldots, x_k, w_1, \ldots, w_m\} = \{x_1, \ldots, x_k, u_1, \ldots, u_\ell, w_1, \ldots, w_m\}$$

forms a spanning set for $\operatorname{Span}(U, W)$. Take scalars $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_\ell, \gamma_1, \ldots, \gamma_m$ such that

$$\underbrace{\alpha_1 x_1 + \cdots + \alpha_k x_k}_{:=x} + \underbrace{\beta_1 u_1 + \cdots + \beta_\ell u_\ell}_{:=u} + \underbrace{\gamma_1 w_1 + \cdots + \gamma_m w_m}_{:=w} = 0.$$

Then $w = -x - u \in U$. Also, it is clear $w \in W$. Then $w \in U \cap W$. Hence there are scalars $\mu_1, \ldots, \mu_k$ such that

$$w = \mu_1 x_1 + \cdots + \mu_k x_k.$$

Then

$$\gamma_1 w_1 + \cdots + \gamma_m w_m - \mu_1 x_1 - \cdots - \mu_k x_k = 0.$$

But these vectors form a basis for W and thus a linealry independent set. Thus $\gamma_1 = \cdots = \gamma_m = 0$ (and the same for μ_i but these won't matter). Then $w = 0$ so $x + u = 0$. But the vectors comprising x and u form a basis for U, thus $\alpha_1 = \cdots = \alpha_k = \beta_1 = \cdots = \beta_\ell = 0$. Thus

$$\{x_1, \ldots, x_k, u_1, \ldots, u_\ell, w_1, \ldots, w_m\}$$

is a linearly independent set.

From this we see that there is a basis for $\operatorname{Span}(U, W)$ which has $k + \ell + m$ elements so $\dim(\operatorname{Span}(U, W)) = k + \ell + m$ and the proof is completed.

(b) We prove the claim by induction on k. If $k = 1$, the result is trivial.

Suppose the result holds from some $k \geq 1$. Let $V_1, \ldots, V_k, V_{k+1}$ be subspaces of V of dimension $n - 1$. Then

$$\dim\left(\cap_{i=1}^{k+1} V_i\right) = \dim\left(V_{k+1} \cap \left(\cap_{i=1}^{k} V_i\right)\right) = \dim\left(V_{k+1}\right) + \dim\left(\cap_{i=1}^{k} V_k\right) - \dim\left(\operatorname{Span}\left(V_{k+1}, \cap_{i=1}^{k} V_i\right)\right),$$

by the dimensionality formula. But $\operatorname{Span}\left(V_{k+1}, \cap_{i=1}^{k} V_i\right)$ has dimension at most n, V_{k+1} has dimension $n - 1$ and by our inductive hypothesis, $\cap_{i=1}^{k} V_i$ has dimension at least $n - k$. Then

$$\dim\left(\cap_{i=1}^{k+1} V_i\right) \geq n - 1 + n - k - n = n - k - 1 = n - (k + 1).$$

Thus the claim holds for $k + 1$. This completes the proof.

Problem F05.10.

(a) For $n = 2, 3, 4, \ldots$, is there an $n \times n$ matrix A with $A^{n-1} \neq 0$ and $A^n = 0$?

(b) Is there an $n \times n$ upper triangular matrix A with $A^n \neq 0$ and $A^{n+1} = 0$?

Solution.

(a) Yes. Let A be the matrix which $A_{i,i+1} = 1$ for $i = 1, \ldots, n-1$ and $A_{ij} = 0$ otherwise. That is, 1 on the first superdiagonal and 0 elsewhere.

(b) No. Suppose $A^{n+1} = 0$. Suppose that λ is an eigenvalue of A with nonzero eigenvector v. Then

$$Av = \lambda v \quad \Longrightarrow \quad A^2 v = \lambda A v = \lambda^2 v \quad \Longrightarrow \quad \cdots \quad \Longrightarrow \quad A^{n+1} v = \lambda^n A v = \lambda^{n+1} v.$$

But $A^{n+1} = 0$ so $\lambda^{n+1} v = 0$. Then $v \neq 0$ implies that $\lambda^{n+1} = 0$ which gives $\lambda = 0$. Thus all eigenvalues of A are zero. Then the minimal polynomial of A has only zero as a root and thus $m_A(x) = x^k$ for some $k \in \mathbb{N}$. However, the degree of m_A is at most n. So

$$m_A(A) = 0 \quad \Longrightarrow \quad A^k = 0, \quad \text{for some } k \leq n \quad \Longrightarrow \quad A^n = 0.$$

[Note: the assumption that A is upper triangular is unnecessary.]

Problem S06.7. Prove that if $a, \lambda \in \mathbb{C}$ with $a \neq 0$, then

$$T = \begin{pmatrix} 1 & a & 0 \\ 0 & 1 & a \\ 0 & 0 & \lambda \end{pmatrix}$$

is not diagonalizable.

Solution. A matrix is diagonalizable if and only if the geometric multiplicity of each eigenvalue is equal to the algebraic multiplicity of the eigenvalue. Here 1 is an eigenvalue of T of algebraic multiplicity 2 but

$$(T - I)v = 0 \quad \Longrightarrow \quad \begin{pmatrix} 0 & a & 0 \\ 0 & 0 & a \\ 0 & 0 & \lambda - 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Then $v_2 = v_3 = 0$ and so $v = (1, 0, 0)^t$ spans the eigenspace corresponding to the eigenvalue 1. Hence the geometric multiplicity of this eigenvalue is only 1. Hence the matrix is not diagonalizable.

Problem S06.8. A linear transformation T is called *orthogonal* if it is non-singular and $T^t = T^{-1}$. Prove that if $T : \mathbb{R}^{2n+1} \to \mathbb{R}^{2n+1}$ is orthogonal, then there is $v \in \mathbb{R}^{2n+1}$ such that $T(v) = v$ or $T(v) = -v$.

Solution. The characteristic polynomial $p_T(t)$ is an odd degree polynomial over $\mathbb{R}$ so there is a real root λ. Then there is a nonzero vector $v \in \mathbb{R}^{2n+1}$ such that $T(v) = \lambda v$ so

$$\lambda^2(v,v) = (\lambda v, \lambda v) = (T(v), T(v)) = (v, T^tT(v)) = (v, Iv) = (v,v).$$

Since $(v,v) \neq 0$, this gives $\lambda^2 = 1$ so $\lambda = \pm 1$ and hence $T(v) = \pm v$.

Problem S06.9. Let S be a real symmetric matrix.

(a) Prove that all eigenvalues of S are real.

(b) State and prove the spectral theorem.

Solution.

(a) Let $\lambda \in \mathbb{C}$ be an eigenvalue of S with nonzero eigenvector v.

$$\lambda(v,v) = (\lambda v, v) = (Sv, v) = (v, Sv) = (v, \lambda v) = \overline{\lambda}(v,v).$$

Hence since $v \neq 0$, we have $\lambda = \overline{\lambda}$ so $\lambda \in \mathbb{R}$.

(b) **Spectral Theorem.** Let S be a symmetric matrix in $M_n(\mathbb{F})$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Then there is an orthonormal basis for $\mathbb{F}^n$ consisting of eigenvectors of S. In particular, S is orthogonally diagonalizable.

Proof. We prove this by induction on n where n is the dimension of the space that S operates on.

For $n = 1$, the statement is obvious since the operator S simply scales vectors.

Assume that if a symmetric matrix operates on an $(n-1)$-dimensional space, then the statement holds. If S is an $n \times n$ matrix, then it operates on $\mathbb{F}^n$. By fundamental theorem of algebra, there is an eigenvalue λ of S and by the above proof, $\lambda \in \mathbb{R}$. Let $0 \neq v \in \mathbb{F}^n$ be a corresponding eigenvector; without loss of generality, $||v|| = 1$. Then $v^\perp = \{x \in \mathbb{F}^n : (x,v) = 0\}$ is an $(n-1)$-dimensional subspace of $\mathbb{F}^n$. Also, if $x \in v^\perp$, then

$$(Sx, v) = (x, Sv) = (x, \lambda v) = \overline{\lambda}(x,v) = 0.$$

Hence $v^\perp$ is an S-invariant subspace. Thus $S\big|_{v^\perp}$ is a symmetric matrix operating on an $(n-1)$-dimensional space. By inductive hypothesis, there is an orthonormal basis $\{x_1, \ldots, x_{n-1}\}$ of $v^\perp$ which consists of eigenvectors of $S\big|_{v^\perp}$ and thus eigenvectors of S. Then $\{x_1, \ldots, x_{n-1}, v\}$ is an orthonormal set in $\mathbb{F}^n$ (and thus a basis for $\mathbb{F}^n$) which consists of eigenvectors of S. In particular, if $\lambda_1, \ldots, \lambda_{n-1}, \lambda$ are the corresponding eigenvalues (there can be repetitions) and we put $P = [x_1 \cdots x_{n-1}\, v]$, then

$$SP = P \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \lambda_{n-1} & \\ & & & & \lambda \end{bmatrix} := PD,$$

so $S = PDP^{-1}$ and since $\{x_1, \ldots, x_{n-1}, v\}$ is an orthonormal set $P^{-1} = P^t$ so $S = PDP^t$ and thus S is orthogonally diagonalizable.

Problem S06.10. Let Y be an arbitrary set of commuting matrices in $M_n(\mathbb{C})$. Prove that there exists a non-zero vector $v \in \mathbb{C}^n$ which is a common eigenvector of all matrices in Y.

Solution. Let $A \in Y$. Then by the fundamental theorem of algebra, A has an eigenvalue λ. Let $E_\lambda = \ker(A - \lambda I)$. Let $0 \neq x \in E_\lambda$. Then for arbitrary $B \in Y$,

$$Ax = \lambda x \quad \Longrightarrow \quad BAx = B(\lambda x) \quad \Longrightarrow \quad A(Bx) = \lambda(Bx).$$

Hence $Bx \in E_\lambda$. Thus E_λ is B invariant. Hence by the fundamental theorem of algebra, $B\big|_{E_\lambda}$ has an eigenvalue and thus an eigenvector $v \neq 0$. This v is a simultaneous eigenvector of A and B. Since B was arbitrary, v is a simultaneuous eigenvector of all matrices in Y.

Problem W06.7. Let V be a complex inner product space. State and prove the Cauchy-Schwarz inequality for V.

Solution.
Cauchy-Schwarz Inequality. Let V be an inner product space over $\mathbb{C}$ and $v, w \in V$. Then

$$|(v,w)| \leq ||v|| \, ||w||$$

with equality if and only if v, w are linearly dependent.

Proof. If $w = 0$, the inequality is trivially satisfied (it is actually equality and v, w are linearly dependent so all statements hold). Assume $w \neq 0$. For any $c \in \mathbb{C}$, we have

$$0 \leq (v - cw, v - cw) = (v,v) - \overline{c}(v,w) - c\overline{(v,w)} + |c|^2 (w,w).$$

Putting $c = \frac{(v,w)}{(w,w)}$ (possible since $w \neq 0$) we see

$$0 \leq (v,v) - \frac{|(v,w)|^2}{(w,w)} - \frac{|(v,w)|^2}{(w,w)} + \frac{|(v,w)|^2 (w,w)}{(w,w)^2} \quad \Longrightarrow \quad \frac{|(v,w)|^2}{(w,w)} \leq (v,v).$$

Multiplying by (w,w), we see

$$|(v,w)|^2 \leq (v,v)(w,w) = ||v||^2 + ||w||^2 \quad \Longrightarrow \quad |(v,w)| \leq ||v|| \, ||w|| \, .$$

We also notice that equality holds if and only if

$$(v - cw, v - cw) = 0$$

which holds if and only if $v = cw$.

Problem W06.8. Let $T : V \to W$ be a linear transformation of finite dimensional innter product spaces. Show that there exists a unique linear transform $T^t : W \to V$ such that

$$(Tv, w)_W = (v, T^t w)_V \quad \text{for all } v \in V, w \in W.$$

Solution. We prove uniqueness first. Suppose there are linear maps $R, S : W \to V$ both satisfying the condition. Then for all $v \in V, w \in W$,

$$(Tv, w)_W = (v, Rw)_V = (v, Sw)_V \qquad \Longrightarrow \qquad (v, (R-S)w)_V = 0.$$

Thus $(R-S)w = 0$ since it is orthogonal to all of V. However, w was arbitrary so this implies $R = S$.

For existence, let $\{v_1, \ldots, v_n\}$ and $\{w_1, \ldots, w_m\}$ be orthonormal bases of V and W respectively. If T has matrix A with respect to these bases, define T^t to have matrix A^t with respect to these basis. Then T^t satisfies the conditon.

Problem W06.9. Let $A \in M_3(\mathbb{R})$ be invertible and satisfy $A = A^t$ and $\det(A) = 1$. Prove that 1 is an eigenvalue of A.

Solution. The conclusion isn't actually true. The matrix

$$A = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

is symmetric and has determinant 1 but doesn't have 1 as an eigenvalue.

It is likely that we were supposed to assume that $A^t = A^{-1}$ rather than $A^t = A$. In this case, see **S03.9**.

Problem S07.2. Let U, V, W be n-dimensional vector spaces and $T : U \to V, S : V \to W$ be linear transformations. Prove that is $S \circ T : U \to W$ is invertible, then both T, S are invertible.

Solution. First suppose that T is not invertible. Then the kernel of T is nontrivial, so there is a non-zero vector $u \in U$ such that T(u)= 0. But then $S \circ T(u) = 0$, so the kernel of $S \circ T$ is nontrivial and so $S \circ T$ is not invertible; a contradiction. Thus T is invertible.

Now assume that S is not invertible. Then the kernel of S is nontrivial so there is a non-zero vector $v \in V$ such that $S(v) = 0$. However, T is invertible, and thus surjective so there is $u \in U$ such that $T(u) = v$ (and $u \neq 0$ since $v \neq 0$). Then $S \circ T(u) = S(T(u)) = S(v) = 0$, so the kernel of $S \circ T$ is nontrivial and $S \circ T$ is not invertible; a contradiction. Hence S is invertible.

Problem S07.3. Consider the space of infinite sequences of real numbers

$$S = \{(a_0, a_1, a_2, \ldots) : a_n \in \mathbb{R}, n = 0, 1, 2, \ldots\}.$$

For each pair if real numbers A and B, prove that the set of solutions $(x_0, x_1, x_2, \ldots)$ of the linear recursion $x_{n+2} = Ax_{n+1} + Bx_n, \quad n = 0, 1, 2, \ldots$ is a subspace of S of dimension 2.

Solution. Let S_0 be set of all solutions to the recurrence relation. We first show that S_0 is indeed a subspace of S. Take $x, y \in S_0$. Put $z = x + y$. Then

$$z_{n+2} = x_{n+2} + y_{n+2} = Ax_{n+1} + Bx_n + Ay_{n+1} + By_n = A(x_{n+1} + y_{n+1}) + B(x_n + y_n) = Az_{n+1} + Bz_n.$$

Thus $z \in S_0$. Next for $\alpha \in \mathbb{R}$ and $x \in S_0$, put $w = \alpha x$. Then

$$w_{n+2} = \alpha x_{n+2} = \alpha(Ax_{n+1} + Bx_n) = A(\alpha x_{n+1}) + B(\alpha x_n) = Aw_{n+1} + Bw_n.$$

Thus $w \in S_0$. Hence S_0 is closed under addition and multiplication by scalars and is thus a subspace of S.

Next we show that if $u, v \in S_0$ are such that $u_0 = 1, u_1 = 0$, $v_0 = 0, v_1 = 1$ then u, v form a basis for S_0. It is clear that they are linearly independent and for $x = (x_0, x_1, x_2, \ldots) \in S_0$, we have $x = x_0 u + x_1 v$. Indeed, $x_0 u + x_1 v$ agrees with x in the first and second entries. Assume that it agrees with x in the $(n+1)^{\text{th}}$ and $(n+2)^{\text{th}}$ entries for some $n \geq 0$. Then

$$x_{n+1} = x_0 u_{n+1} + x_1 v_{n+1} \quad \text{and} \quad x_{n+2} = x_0 u_{n+2} + x_1 v_{n+2}.$$

Consider

$$\begin{aligned} x_{n+3} &= Ax_{n+2} + Bx_{n+1} \\ &= A(x_0 u_{n+2} + x_1 v_{n+2}) + B(x_0 u_{n+1} + x_1 v_{n+1}) \\ &= x_0(Au_{n+2} + Bu_{n+1}) + x_1(Av_{n+2} + Bv_{n+1}) \\ &= x_0 u_{n+3} + x_1 v_{n+3}. \end{aligned}$$

Hence by induction it is true that $x = x_0 u + x_1 v$. Thus u, v span S_0. Hence $\dim(S_0) = 2$.

Problem S07.4. Suppose that A is a symmetric $n \times n$ real matrix and let $\lambda_1, \ldots, \lambda_\ell$ be the distinct eigenvalues of A. Find the sets

$$X = \left\{ x \in \mathbb{R}^n : \lim_{k \to \infty} \left(x^t A^{2k} x \right)^{1/k} \text{ exists} \right\}$$

and

$$L = \left\{ \lim_{k \to \infty} \left(x^t A^{2k} x \right)^{1/k} : x \in X \right\}.$$

Solution. We prove that $X = \mathbb{R}^n$ and $L = \{0, \lambda_1^2, \ldots, \lambda_\ell^2\}$.

By taking the zero vector for x, it is clear that the zero vector is in X and zero is in L. We'll focus on non-zero vectors henceforth.

Since A is symmetric, by the spectral theorem there is an orthonormal basis $\{x_1, \ldots, x_n\}$ for $\mathbb{R}^n$ consisting of eigenvectors of A. Suppose that $\mu_1, \ldots, \mu_n \in \mathbb{R}$ are the corresponding eigenvalues [the eigenvalues are real since A is symmetric; also repititions are allowed so this is a slightly different list from $\lambda_1, \ldots, \lambda_\ell$]. Let $0 \neq x \in \mathbb{R}^n$. Then there are $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ (not all zero) such that

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n.$$

Then

$$A^{2k} x = \alpha_1 A^{2k} x_1 + \cdots + \alpha_n A^{2k} x_n = \alpha_1 \mu_1^{2k} x_1 + \cdots + \alpha_n \mu_n^{2k} x_n$$

so by orthogonality,

$$x^t A^{2k} x = \alpha_1^2 \mu_1^{2k} + \cdots + \alpha_n \mu_n^{2k}.$$

From here, it is clear that

$$x^t A^{2k} x \geq \alpha_j^2 \mu_j^{2k}$$

for all $j = 1, \ldots, n$.

Let $m \in \{1, \ldots, n\}$ be an index satisfying $\alpha_m \neq 0$, and if $\alpha_j \neq 0$ for some $j = 1, \ldots, n$, then $|\mu_j| \leq |\mu_m|$. That is, m is the index of the largest of the μ's which has a non-zero coefficient in the representation of x in this basis (note: such m always exists; it is not necessarily unique, but that won't matter). Then

$$x^t A^{2k} x = \alpha_1^2 \mu_1^{2k} + \cdots + \alpha_n \mu_n^{2k} \leq n\alpha_m^2 \mu_m^{2k}.$$

Then using our two bounds, we see that

$$\alpha_m^{2/k} \mu_m^2 \leq \left(x^t A^{2k} x\right)^{1/k} \leq (n\alpha_m^2)^{1/k} \mu_m^2.$$

Thus by the squeeze theorem, $\lim\limits_{k\to\infty} \left(x^t A^{2k} x\right)^{1/k} = \mu_m^2$.

This shows that $X = \mathbb{R}^n$. It also shows that for every non-zero $x \in \mathbb{R}^n$, we have $\lim\limits_{k\to\infty} \left(x^t A^{2k} x\right)^{1/k} = \mu_j^2$ for some $j = 1, \ldots, n$. Thus the eigenvalues of A are the only possible values for the limit. To see that each eigenvalue is indeed achieved, notice that

$$\left(x_j^t A^{2k} x_j\right)^{1/k} = \mu_j^2$$

for all $k \in \mathbb{N}$. Thus $\lim\limits_{k\to\infty} \left(x_j^t A^{2k} x_j\right)^{1/k} = \mu_j^2$ for each $j = 1, \ldots, n$.

Problem S07.5. Let T be a normal linear operator on a finite dimensional complex inner product space V. Prove that if v is an eigenvector of T, then v is also an eigenvector of the adjoint T^*.

Solution. First, consider if S is any normal operator and $x \in \mathbb{C}^n$, then

$$(Sx, Sx) = (x, S^*Sx) = (x, SS^*x) = (S^*x, S^*x).$$

Thus $||Sx|| = ||S^*x||$ for all $x \in \mathbb{C}^n$.

Let v be an eigenvector of T with corresponding eigenvalue $\lambda \in \mathbb{C}$. Since T is normal, so is $(T - \lambda I)$. Then

$$0 = ||(T - \lambda I)v|| = ||(T - \lambda I)^* x|| = \left|\left|(T^* - \overline{\lambda} I)v\right|\right|.$$

Hence $T^* v = \overline{\lambda} v$ so v is an eigenvector of T^* corresponding to eigenvalue $\overline{\lambda}$.

Problem F07.3. Let V be a vector space and T a linear transformation such that Tv and v are linearly dependent for every $v \in V$. Prove that T is a scalar multiple of the identity transformation.

Solution. If $\dim(V) = 1$, the result is trivial. Assume that $\dim(V) > 1$ Suppose $x, y \in V$ are linearly independent. Then there are scalars α, β, γ such that

$$Tx = \alpha x, \quad Ty = \beta y, \quad T(x + y) = \gamma(x + y).$$

Then
$$0 = (\gamma - \alpha)x + (\gamma - \beta)y.$$
But x, y were taken to be linearly independent, so $\alpha = \gamma = \beta$. Since the same argument would work for any linearly independent vectors, we see that $Tv = \alpha v$ for all $v \in V$. Thus $T = \alpha I$.

Problem F07.12.

(a) Suppose that $x_0 < x_1 < \ldots < x_n$ are points in $[a, b]$. Define linear functionals on $\mathbb{P}^n$ (the space of all polynomials of degree less than or equal to n) by
$$\ell_j(p) = p(x_j), \;\; j = 0, 1, \ldots, n, \;\; p \in \mathbb{P}^n.$$
Show that the set $\{\ell_j\}_{j=0}^n$ is linearly independent.

(b) Show that there are unique coefficients $c_j \in \mathbb{R}$ such that
$$\int_a^b p(t)dt = \sum_{j=0}^n c_j \ell_j(p)$$
for all $p \in \mathbb{P}^n$.

Solution.

(a) Let $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ be such that
$$\alpha_0 \ell_0 + \alpha_1 \ell_1 + \cdots + \alpha_n \ell_n = 0.$$
Put
$$p_i(x) = \prod_{m=0, m \neq i}^n (x - x_m), \quad i = 0, 1, \ldots, n.$$
Then each p_i is a polynomial of degree n and $\ell_i(p_i) \neq 0$ since $x_i \neq x_j$ when $i \neq j$. However, $\ell_j(p_i) = 0$ when $i \neq j$ since $(x - x_j)$ is a factor of p_i when $i \neq j$. Thus for each $i = 0, 1, \ldots, n$,
$$\alpha_0 \ell_0(p_i) + \alpha_1 \ell_1(p_i) + \ldots + \alpha_n \ell_n(p_i) = 0 \quad \Longrightarrow \quad \alpha_i \ell_i(p_i) = 0 \quad \Longrightarrow \quad \alpha_i = 0.$$
Hence $\ell_0, \ell_1, \ldots, \ell_n$ are linearly independent.

(b) For any finite dimensional vector space V, we know $\dim(V) = \dim(V^*)$. Here $\mathbb{P}^n$ has dimension $n+1$ and we have found $n+1$ linearly independent members of $(\mathbb{P}^n)^*$. Thus $\ell_0, \ell_1, \ldots, \ell_n$ form a basis for $(\mathbb{P}^n)^*$. Since $\ell(p) = \int_a^b p(x)dx, \;\; p \in \mathbb{P}^n$ defines a linear functional on $\mathbb{P}^n$, we know there are unique $c_0, c_1, \ldots, c_n \in \mathbb{R}$ such that $\ell = c_0 \ell_0 + c_1 \ell_1 + \cdots + c_n \ell_n$. Hence
$$\int_a^b p(x)dx = \sum_{j=0}^n c_j \ell_j(p), \quad \text{for all } p \in \mathbb{P}^n.$$

Problem S08.8. Assume that V is an n-dimensional vector space and that T is a linear transformation $T : V \to V$ such that $T^2 = T$. Prove that every $v \in V$ can be written uniquely as $v = v_1 + v_2$ such that $T(v_1) = v_1$ and $T(v_2) = 0$.

Solution. Note that if $v \in \text{im}(T)$, then $v = T(w)$ for some $w \in V$ and so

$$T(v) = T^2(w) = T(w) = v.$$

Thus T fixes members of its image so $T(v - T(v)) = 0$ for all $v \in V$.

For any $v \in V$, write $v = T(v) + (v - T(v)) := v_1 + v_2$. It's clear that this is one way to represent v in the desired way. Assume that $v = v_1 + v_2 = x_1 + x_2$, where $T(v_1) = v_1, T(x_1) = x_1$ and $T(v_2) = 0, T(x_2) = 0$. Then

$$T(v) = T(v_1) + T(v_2) = T(x_1) + T(x_2) \quad \Longrightarrow \quad T(v_1) = T(x_1) \quad \Longrightarrow v_1 = x_1.$$

Then $v_1 + v_2 = v_1 + x_2 \quad \Longrightarrow \quad v_2 = x_2$. This gives uniqueness of such a representation.

Problem S08.9. Let V be a finite-dimensional vector space over $\mathbb{R}$.

(a) Show that if V has odd dimension and $T : V \to V$ is a real linear transformation, then T has a non-zero eigenvector $v \in V$.

(a) Show that for every even positive integer n, there is a vector space V over $\mathbb{R}$ of dimension n and a real linear transformation $T : V \to V$ such that there is no non-zero $v \in V$ that satisfies $T(v) = \lambda v$ for some $\lambda \in \mathbb{R}$.

Solution.

(a) The characteristic polynomial p_T of T is a polynomial over $\mathbb{R}$ of degree $\dim(V)$ which is odd. Every odd ordered polynomial over $\mathbb{R}$ has a root in $\mathbb{R}$. We know this root (say $\lambda \in \mathbb{R}$) is an eigenvalue of T and thus there is a non-zero eigenvector $v \in V$ such that $T(v) = \lambda v$. Hence T has an eigenvector.

(b) Let $V = \mathbb{R}^n$ for positive even n and let T be the left multiplication operator for

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

It is easy to see that T has characteristic polynomial $p_T(t) = (-t)^n + (-1)^n$. But n is even, so $p_T(t) = t^n + 1$. This equation has no roots in $\mathbb{R}$ when n is even and so there is no real eigenvalue and thus no non-zero $v \in \mathbb{R}^n$ such that $T(v) = \lambda v$ for some $\lambda \in \mathbb{R}$.

Problem F08.7. Suppose that T is a complex $n \times n$ matrix and that $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of T with corresponding eigenvectors $v_1, \ldots, v_k$. Show that $v_1, \ldots, v_k$ are linearly independent.

Solution. We use induction.

First, it is clear that $\{v_1\}$ is a linearly independent set because eigenvectors are necessarily non-zero.

Now assume that $\{v_1, \ldots, v_\ell\}$ are linearly independent for some $2 \le \ell < k$. Let $\alpha_1, \ldots, \alpha_\ell, \alpha_{\ell+1} \in \mathbb{C}$ be such that

$$\alpha_1 v_1 + \cdots + \alpha_\ell v_\ell + \alpha_{\ell+1} v_{\ell+1} = 0.$$

Then

$$(T - \lambda_{\ell+1} I)(\alpha_1 v_1 + \cdots + \alpha_\ell v_\ell + \alpha_{\ell+1} v_{\ell+1}) = 0.$$

But $(T - \lambda_{\ell+1} I)v_i = (\lambda_i - \lambda_{\ell+1})v_i$ for $i = 1, \ldots, \ell$ and $(T - \lambda_{\ell+1} I)v_{\ell+1} = 0$. Thus

$$\alpha_1(\lambda_1 - \lambda_{\ell+1})v_1 + \cdots + \alpha_\ell(\lambda_\ell - \lambda_{\ell+1})v_\ell = 0.$$

But these vectors are linearly independent by our inductive hypothesis, so $\alpha_1(\lambda_1 - \lambda_{\ell+1}) = \cdots = \alpha_\ell(\lambda_\ell - \lambda_{\ell+1}) = 0$. However, since the eigenvalues are distinct, this implies that $\alpha_1 = \cdots = \alpha_\ell = 0$. Hence

$$0 = \alpha_1 v_1 + \cdots + \alpha_\ell v_\ell + \alpha_{\ell+1} v_{\ell+1} = \alpha_{\ell+1} v_{\ell+1} \quad \Longrightarrow \quad \alpha_{\ell+1} = 0.$$

Thus $\{v_1, \ldots, v_\ell, v_{\ell+1}\}$ is a linearly independent set.

By induction, we conclude that $\{v_1, \ldots, v_k\}$ are linearly independent.

Problem F08.8. Must the eigenvectors of a linear transformation $T : \mathbb{C}^n \to \mathbb{C}^n$ span $\mathbb{C}^n$?

Solution. No. Take any non-diagonalizable matrix $A \in M_n(\mathbb{C})$ and let $T(x) = Ax, x \in \mathbb{C}^n$. For example, if

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

Then, up to multiplication by scalars, the eigenvectors of A (and thus T) are

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad v_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

which clearly do not span $\mathbb{C}^3$.

Problem F08.9.

(a) Prove that any linear transformation $T : \mathbb{C} \to \mathbb{C}$ must have an eigenvector.

(b) Is (a) true for any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$?

Solution.

(a) By the fundamental theorem of algebra, $p_T(t) = \det(T - tI)$ has a root $\lambda \in \mathbb{C}$. Then $\det(T - \lambda I) = 0$ so $(T - \lambda I)$ is a singular transformation. Hence there is a non-zero $v \in \mathbb{C}^n$ such that $(T - \lambda I)(v) = 0$ or $T(v) = \lambda v$; that is, v is an eigenvector of T.

(b) No. Let

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

and define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(x) = Ax, x \in \mathbb{R}^2$. Suppose T has an eigenvector $0 \neq v \in \mathbb{R}^2$. Then there is $\lambda \in \mathbb{R}$ such that $T(v) = \lambda v$. If $v = (v_1, v_2)$, this implies that $-v_2 = \lambda v_1$ and $v_1 = \lambda v_2$ where at least one of v_1, v_2 is non-zero. Asusme $v_1 \neq 0$. Then by the second equation, neither λ nor v_2 are zero. Also, by the first equation

$$-v_2 = \lambda v_1 = \lambda^2 v_2 \quad \Longrightarrow \quad \lambda^2 = -1,$$

a contradiction to the fact that $\lambda \in \mathbb{R}$. Hence there is no eigenvector for T.

Problem F08.11. Consider the Poisson equation with periodic boundary condition

$$\frac{\partial^2 u}{\partial x^2} = f, \quad x \in (0,1),$$
$$u(0) = u(1).$$

A second order accurate approximation to the problem is given by $A\mathbf{u} = \Delta x^2 \mathbf{f}$ where

$$A = \begin{bmatrix} -2 & 1 & 0 & \cdots & 0 & 1 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & 0 & \cdots \\ & & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 & -2 & 1 \\ 1 & 0 & \cdots & 0 & 1 & -2 \end{bmatrix},$$

$\mathbf{u} = [u_0, u_1, \ldots, u_{n-1}]^t, \mathbf{f} = [f_0, f_1, \ldots, f_{n-1}]^t$ and $u_i \approx u(x_i)$ with $x_i = i\Delta x, \Delta x = 1/n$ and $f_i = f(x_i)$ for $i = 0, 1, \ldots, n-1$.

(a) Show that A is singular.

(b) What conditions must $\mathbf{f}$ satisfy so that a solution exists?

Solution.

(a) Notice $A\mathbf{v} = 0$ when $\mathbf{v} = [1 \ 1 \ \cdots \ 1]^t$, so A has a nontrivial null space and is thus singular (in fact, up to scalar multiplication, this is the only vector in the null space of A).

(b) We need $\mathbf{f}$ in the range of A for a solution to exist. The Fredholm alternative tells us that $\text{range}(A) = \text{null}(A^t)^\perp$. But $A^t = A$, so we simply need $\mathbf{f} \in \text{null}(A)^\perp$. By (a), this is equivalent to

$$(\mathbf{v}, \mathbf{f}) = 0 \quad \Longleftrightarrow \quad f_0 + f_1 + \cdots + f_{n-1} = 0.$$

Problem F08.12. Consider the least square problem

$$\min_{x\in\mathbb{R}^n} ||Ax - b||$$

where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$ and $m \geq n$. Prove that if x and $x + \alpha z$ ($\alpha \neq 0$) are both minimizers, then $z \in \text{null}(A)$.

Solution. Let $b = b_1 + b_2$ be the projection of b onto $\text{im}(A)$; i.e., b_1 is the unique vector in $\text{im}(A)$ such that $b_1 - b \in \text{im}(A)^\perp$. We know that b_1 is the closest vector to b which lies in $\text{im}(A)$. Then the minimizers of $||Ax - b||$ are exactly those vectors $x \in \mathbb{R}^n$ such that $Ax = b_1$. Let x be such a minimizer. Then for any $y \in \mathbb{R}^n$, $Ay \in \text{im}(A)$ and so $(Ay, b_1 - b) = 0$. But then

$$0 = (Ay, Ax - b) = (y, A^*(Ax - b)).$$

Since y is arbitrary, this implies that $A^*(Ax - b)$ is orthogonal to all of $\mathbb{R}^n$ so $A^*(Ax - b) = 0$ or $A^*Ax = A^*b$. That is, any minimizer x of $||Ax - b||$ must satisfy the normal equations: $A^*Ax = A^*b$.

Assume that both x and $x + \alpha z,\ \alpha \neq 0$ are minimizers of $||Ax - b||$. Then

$$A^*b = A^*Ax = A^*A(x + \alpha z) \quad \Longrightarrow \quad \alpha A^*Az = 0 \quad \Longrightarrow \quad A^*Az = 0.$$

Taking the inner product of A^*Az with z, we see

$$(z, A^*Az) = 0 \quad \Longrightarrow \quad (Az, Az) = 0 \quad \Longrightarrow \quad ||Az||^2 = 0 \quad \Longrightarrow \quad Az = 0.$$

Thus $z \in \text{null}(A)$.

Problem F09.4. Let V be a finite dimensional inner product space and let U be a subspace of V. Show that $\dim(U) + \dim(U^\perp) = \dim(V)$.

Suppose $\dim(U) = n$. Then $\dim(V) = n + k$ for some $k \in \mathbb{N}_0$. If $k = 0$, then $U = V$ and the result hold trivially. Assume $k > 0$. Let $\{x_1, \ldots, x_n\}$ form an orthonormal basis for U. We can extend this to an orthonormal basis $\{x_1, \ldots, x_n, y_1, \ldots, y_k\}$ for V. If we can prove that $y_1, \ldots, y_k$ is an orthonormal basis for $U^\perp$, then we will be done.

Since the set $\{x_1, \ldots, x_n, y_1, \ldots, y_k\}$ is orthonormal, for any $j = 1, \ldots, k$ and $i = 1, \ldots, n$, we have $(y_j, x_i) = 0$. Hince y_j is orthogonal to each basis member for U and hence orthogonal to U; that is, $y_j \in U^\perp$ for each j.

Take $y \in U^\perp$. Then, $y \in V$ so there are scalars $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_k$ such that

$$y = \alpha_1 x_1 + \cdots + \alpha_n x_n + \beta_1 y_1 + \cdots + \beta_k y_k.$$

Since each $x_i \in U$, we have $(y, x_i) = 0$. Alternately taking the inner product of y with each x_i, we find $\alpha_i = 0$ for each i. Thus

$$y = \beta_1 y_1 + \cdots + \beta_k y_k.$$

Hence $\{y_1, \ldots, y_k\}$ is a spanning set for $U^\perp$; it is also a linearly independent set since it is part of the basis for V. Thus $\{y_1, \ldots, y_k\}$ is a basis for $U^\perp$ so $\dim(U^\perp) = k$ so the claim is proven.

Problem F09.5. Show that if $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ are all different and some $b_1, \ldots, b_n \in \mathbb{R}$ satisfy

$$\sum_{i=1}^{n} b_i e^{\alpha_i t} \text{ for all } t \in (-1,1)$$

then necessarily $b_1 = \cdots = b_n = 0$.

Solution. Let $T : C^\infty(-1,1) \to C^\infty(-1,1)$ be defined by

$$(Tf)(t) = f'(t), t \in (-1,1).$$

We see that for each α_i, if we define $f_i \in C^\infty(-1,1)$ by $f_i(t) = e^{\alpha_i t}$, $t \in (-1,1)$ then $Tf_i = \alpha_i f_i$. Hence the functions f_i are eigenvectors of T corresponding to different eigenvalues. But eigenvectors of a linear operator corresponding to different eigenvalues are linearly independent. Hence

$$b_1 f_1 + \cdots + b_n f_n = 0 \quad \Longrightarrow \quad b_1 = \cdots = b_n = 0$$

which is the desired conclusion.

Problem F09.12. Let $n \geq 2$ and let V be an n-dimensional vector space over $\mathbb{C}$ with a set of basis vectors $e_1, \ldots, e_n$. Let T be the linear transformation of V satisfying

$$T(e_i) = e_{i+1}, \;\; i = 1, \ldots, n-1 \quad \text{and} \quad T(e_n) = e_1.$$

(a) Show that T has 1 as an eigenvalue. Find an eigenvector with eigenvalue 1 and show that it is unique up to scaling.

(b) Is T diagonalizable?

Solution.

(a) Let $x = e_1 + e_2 + \cdots + e_n \in V$. Then $T(x) = T(e_1) + T(e_2) + \cdots + T(e_n) = e_2 + e_3 + \cdots + e_1 = e_1 + e_2 + \cdots + e_n = x$. Further $x \neq 0$ since the basis vectors are linearly independent. Thus x is an eigenvector of T with eigenvalue 1. Conversely, if v is an eigenvector of T corresponding to eigenvalue 1, then $T(v) = v$. Also there are $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ such that $v = \alpha_1 e_1 + \cdots + \alpha_n e_n$. Then

$$\alpha_1 e_1 + \cdots + \alpha_n e_n = v = T(v) = \alpha_1 e_2 + \alpha_2 e_3 \cdots + \alpha_n e_1$$

which gives

$$(\alpha_1 - \alpha_n)e_1 + (\alpha_2 - \alpha_1)e_2 + \cdots (\alpha_n - \alpha_{n-1})e_n = 0.$$

But these vectors are linearly independent so

$$\alpha_1 - \alpha_n = 0, \alpha_2 - \alpha_1 = 0, \cdots, \alpha_n - \alpha_{n-1} = 0$$

which gives $\alpha_1 = \cdots = \alpha_n$. Call this value $\alpha \in \mathbb{C}$. Then $v = \alpha e_1 + \cdots + \alpha e_n = \alpha x$. Hence the eigenvector is unique up to scaling.

(b) Since $T(e_j) = e_{j+1}$ for all $j = 1, \dots, n-1$ and $T(e_n) = e_1$, the matrix $[T]_B$ is given by

$$[T]_B = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & & \\ 0 & 1 & 0 & \cdots & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

Then $p_T(t) = \det([T]_B - tI) = \pm(t^n - 1)$. But $t^n - 1$ has distinct roots for all n. Thus the eigenvalues of T are distinct and so T is diagonalizable.

Problem S10.1. Let $u_1, \dots, u_n$ be an orthonormal basis for $\mathbb{R}^n$ and let $y_1, \dots, y_n$ be a collection of vectors in $\mathbb{R}^n$ such that $\sum_{j=1}^n ||y_j||^2 < 1$. Show that $u_1 + y_1, \dots, u_n + y_n$ form a basis for $\mathbb{R}^n$.

Solution. Since invertible linear maps send one basis to another, it suffices to show that there is an invertible linear map $L : \mathbb{R}^n \to \mathbb{R}^n$ such that $L(u_j) = u_j + y_j$ for all $j = 1, \dots, n$. Define a linear map, $T : \mathbb{R}^n \to \mathbb{R}^n$ by $T(u_j) = -y_j, j = 1, \dots, n$ (a linear map is uniquely determined by how it treats basis elements).

Let $x \in \mathbb{R}^n$. Then there are scalars $\alpha_1, \dots, \alpha_n$ such that $x = \sum_{j=1}^n \alpha_j u_j$. Then

$$\begin{aligned} ||T(x)|| &= \left|\left| \sum_{j=1}^n \alpha_j T(u_j) \right|\right| \\ &= \left|\left| \sum_{j=1}^n \alpha_j y_j \right|\right| \\ &\le \sum_{j=1}^n |\alpha_j| \, ||y_j|| \\ &\le \left(\sum_{j=1}^n |\alpha_j|^2 \right)^{1/2} \left(\sum_{j=1}^n ||y_j||^2 \right)^{1/2}. \end{aligned}$$

But by orthogonality of $u_1, \dots, u_n$, we have $||x|| = \left(\sum_{j=1}^n |\alpha_j|^2 \right)^{1/2}$. Thus we have that $||T|| \le \sum_{j=1}^n ||y_j||^2 < 1$. But $||T|| < 1$ implies that $I - T$ is invertible. Further, we see that $(I - T)(u_j) = u_j + y_j, j = 1, \dots, n$. Hence $u_1 + y_1, \dots, u_n + y_n$ form a basis for $\mathbb{R}^n$.

Problem S10.3. Let S, T be two normal transformations in the complex finite dimensional inner product space V such that $ST = TS$. Prove that there is a basis for V consisting of vectors which are simultaneous eigenvectors of S and T.

Solution. Suppose that $n = \dim(V)$. Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of S. Then by the spectral theorem, the eigenvectors of S can be taken to form an orthonormal basis for V which means that

$$E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k} = V.$$

Consider, for $v \in E_{\lambda_i}$, we have $Sv = \lambda_i v$. Then

$$TSv = T(\lambda_i v) \quad \implies \quad STv = \lambda_i Tv \quad \implies \quad S(Tv) = \lambda_i (Tv).$$

Thus $Tv \in E_{\lambda_i}$. Hence E_{λ_i} is T-invariant. Then $T\big|_{E_{\lambda_i}}$ is a normal operator on E_{λ_i} so by the spectral theorem, there is an orthonormal basis $v_1^{(i)}, \ldots, v_{\ell_i}^{(i)}$ for E_{λ_i} consisting of eigenvectors of T. Then

$$\bigcup_{i=1}^{k} \bigcup_{j=1}^{\ell_i} \left\{ v_j^{(i)} \right\}$$

is a basis of V consisting of simultaneous eigenvectors of both S and T.

Problem S10.4.

(i) Let A be a real symmetric $n \times n$ matrix such that $x^t A x \leq 0$ for every $x \in \mathbb{R}^n$. Prove that trace$(A) = 0$ implies $A = 0$.

(ii) Let T be a linear transformation in the complex finite dimensional vector space V with an inner product. Suppose that $TT^* = 4T - 3I$ where I is the identity transformation. Prove that T is Hermitian positive definite and find all possible eigenvalues of T.

Solution. I'm have no idea why these two questions are grouped into one problem. As far as I can see, they have nothing to do with each other and have elementary solutions which are completely independent of the other.

(i) Let e_i be the standard basis vectors. Then

$$e_i^t A e_i \leq 0 \quad \implies \quad a_{ii} \leq 0.$$

Then if the sum of the diagonal entries is zero, each entry must be zero; i.e., $a_{ii} = 0, i = 1, \ldots n$. Next,

$$(e_i + e_j)^t A (e_i + e_j) \leq 0 \quad \implies \quad a_{ii} + a_{ij} + a_{ji} + a_{jj} \leq 0 \quad \implies \quad a_{ij} + a_{ji} \leq 0.$$

But since A is symmetric, this implies $a_{ij} \leq 0$. Also

$$(e_i - e_j)^t A (e_i - e_j) = a_{ii} - a_{ij} - a_{ji} + a_{jj} \leq 0$$

which gives $a_{ij} \geq 0$. Thus $a_{ij} = 0$. Since i, j were arbitrary, this gives $A = 0$.

(ii) We see $(TT^*)^* = (T^*)^* T^* = TT^*$. Thus

$$(4T - 3I)^* = 4T - 3I \quad \implies \quad 4T^* - 3I = 4T - 3I \quad \implies \quad T^* = T$$

so T is Hermitian. Then

$$T = \tfrac{1}{4} TT^* + \tfrac{3}{4} I = \tfrac{1}{4} T^2 + \tfrac{3}{4} I.$$

Then for any $x \in V$,

$$(Tx, x) = \tfrac{1}{4}(T^2x, x) + \tfrac{3}{4}(x, x) = \tfrac{1}{4}(Tx, Tx) + \tfrac{3}{4}(x, x) \geq 0$$

since inner products are positive definite. Further, if $(Tx, x) = 0$ then $(Tx, Tx) = 0$ and $(x, x) = 0$. The latter can only happen if $x = 0$. Thus T is positive definite.

The functional equation for T gives

$$T^2 - 4T + 3I = 0 \quad \Longrightarrow \quad (T - I)(T - 3I) = 0.$$

Then by the Cayley-Hamilton Theorem, the characteristic polynomial of T must divide $(t - 1)(t - 3)$. Thus the only possible eigenvalues of T are 1 and 3.

Problem S10.6. Let $A = \begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix}$.

(i) Find a Jordan form J of A and a patrix P such that $P^{-1}AP = J$.

(ii) Compute A^{100} and J^{100}.

(iii) Find a formula for a_n when $a_0 = a, a_1 = b$ and $a_{n+1} = 4a_n - 4a_{n-1}$.

Solution.

(i) The characteristic polynomial of A is $p_A(x) = x(x-4) + 4 = x^2 - 4x + 4 = (x-2)^2$. Thus the sole eigenvalue of A is 2. We see $N = A - 2I$ is not the zero matrix. Thus there is $x \neq 0$ so that $Nx \neq 0$. We see by inspection that $x = (1, 0)^t$ gives $Nx = (2, 1)^t$. Put

$$P = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then

$$P^{-1} = -\begin{pmatrix} 0 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}$$

and

$$P^{-1}AP = \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & -4 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} := J$$

(ii) We see

$$J^2 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 0 & 4 \end{pmatrix},$$

and

$$J^3 = \begin{pmatrix} 4 & 4 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 12 \\ 0 & 8 \end{pmatrix}.$$

From these, it is reasonable to guess that

$$J^n = \begin{pmatrix} 2^n & n2^{n-1} \\ 0 & 2^n \end{pmatrix}.$$

Indeed, assuming this holds for J^n, we see

$$J^{n+1} = \begin{pmatrix} 2^n & n2^{n-1} \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2^{n+1} & 2^n + n2^n \\ 0 & 2^{n+1} \end{pmatrix} = \begin{pmatrix} 2^{n+1} & (n+1)2^n \\ 0 & 2^{n+1} \end{pmatrix}.$$

Thus the formular holds by induction. Then $A = PJP^{-1} \implies A^n = PJ^nP^{-1}$. So

$$\begin{aligned} A^n &= \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2^n & n2^{n-1} \\ 0 & 2^n \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 2^{n+1} & (n+1)2^n \\ 2^n & n2^{n-1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \\ &= \begin{pmatrix} (n+1)2^n & -n2^{n+1} \\ n2^{n-1} & -(n-1)2^n \end{pmatrix}. \end{aligned}$$

Then

$$A^{100} = \begin{pmatrix} 101 \cdot 2^{100} & -100 \cdot 2^{101} \\ 100 \cdot 2^{99} & -99 \cdot 2^{100} \end{pmatrix} \quad \text{and} \quad J^{100} = \begin{pmatrix} 2^{100} & 100 \cdot 2^{99} \\ 0 & 2^{100} \end{pmatrix}.$$

(iii) The sequence a_n satisfies

$$A^n \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$$

so $a_n = n2^{n-1}b - (n-1)2^n a$.

Problem F10.5. Prove or disprove the following two statements. For any two subsets U, W of a vector space V,

(a) $\text{Span}(U) \cap \text{Span}(W) = \text{Span}(U \cap W)$

(b) $\text{Span}(U) + \text{Span}(W) = \text{Span}(U \cup W)$

Solution.

(a) The statement is false. The sets

$$U = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}, \quad \text{and} \quad W = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

provide a counterexample since

$$\text{Span}(U) \cap \text{Span}(W) = \mathbb{R}^2 \cap \mathbb{R}^2 = \mathbb{R}^2 \neq \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} = \text{Span}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) = \text{Span}(U \cap W).$$

(b) The statement is true. Take $x + y \in \text{Span}(U) + \text{Span}(W)$ then $x \in \text{Span}(U) \implies x \in \text{Span}(U \cup W)$ and likewise for y. Hence, since $\text{Span}(U \cup W)$ is a vector space, we have $x + y \in \text{Span}(U \cup W)$ and so $\text{Span}(U) + \text{Span}(W) \subset \text{Span}(U \cup W)$.

Conversely, $\text{Span}(U) + \text{Span}(W)$ clearly contains $U \cup W$, but $\text{Span}(U \cup W)$ is the smallest vector space containing $U \cup W$, so it must be the case that $\text{Span}(U \cup W) \subset \text{Span}(U) + \text{Span}(W)$.

Problem F10.6. Let T be an invertible linear map on a finite dimensional vector space V over a field $\mathbb{F}$. Prove there is a polynomial $f \in \mathbb{F}[x]$ such that $T^{-1} = f(T)$.

Solution. By the Cayley-Hamilton Theorem, the linear operator satisfies $p_T(T) = 0$ where p_T is the characteristic polynomial. Put

$$p_T(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n.$$

Specifically, since T is invertible, $x = 0$ is not a root of $p_T(x)$ and thus $\alpha_0 \neq 0$. Then

$$p_T(T) = \alpha_0 I + \alpha_1 T + \cdots + \alpha_n T^n$$

where I is the identity map on V. Then

$$-\alpha_0 I = \alpha_1 T + \cdots + \alpha_n T^n \quad \Longrightarrow \quad T^{-1} = -\frac{\alpha_1}{\alpha_0} I - \cdots - \frac{\alpha_n}{\alpha_0} T^{n-1}.$$

Hence T^{-1} is a polynomial expression of T.

Problem F10.7. Let V, W be inner product spaces over $\mathbb{C}$ such that $\dim(V) \leq \dim(W) < \infty$. Prove that there is a linear transformation $T : V \to W$ satisfying $(T(x), T(y))_W = (x, y)_V$ for all $x, y \in V$.

Solution. . Suppose $n \in \mathbb{N}$ and $k \in \mathbb{N} \cup \{0\}$ are such that $\dim(V) = n$, $\dim(W) = n + k$. Let $\{v_1, \ldots, v_n\}$ be an orthonormal basis for V and $\{w_1, \ldots, w_n, w_{n+1}, \ldots, w_{n+k}\}$ be an orthonormal basis for W. Define $T : V \to W$ such that $T(v_j) = w_j$ for $j = 1, \ldots, n$ and T is linear (a linear map is completely determined by how it treats basis elements). For any $x, y \in V$, there are $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n \in \mathbb{F}$ such that

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n, \qquad y = \beta_1 v_1 + \cdots + \beta_n v_n.$$

Then

$$\begin{aligned} (T(x), T(y))_W &= (\alpha_1 w_1 + \cdots + \alpha_n w_n, \beta_1 w_1 + \cdots + \beta_n w_n)_W \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\beta}_j (w_i, w_j)_W = \sum_{i=1}^n \alpha_i \overline{\beta}_i \end{aligned}$$

by orthogonality. Also,

$$\begin{aligned} (x, y)_V &= (\alpha_1 v_1 + \cdots + \alpha_n v_n, \beta_1 v_1 + \cdots + \beta_n v_n)_V \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\beta}_j (v_i, v_j)_V = \sum_{i=1}^n \alpha_i \overline{\beta}_i. \end{aligned}$$

Thus $(T(x), T(y))_W = (x, y)_V$ for all $x, y \in V$.

Problem F10.9. Consider the following iterative method:

$$x_{k+1} = A^{-1}(B x_k + c)$$

where

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(a) Assume the iteration converges. To what vector x does the iteration converge?

(b) Does the iteration converge for arbitrary initial vectors?

Solution.

(a) Assuming the iteration converges to $x = (y, z)^t$, we must have

$$x = A^{-1}(Bx + c) \quad \Longrightarrow \quad \begin{aligned} y &= y + \tfrac{1}{2}z + \tfrac{1}{2}, \\ z &= \tfrac{1}{2}y + z + \tfrac{1}{2}. \end{aligned}$$

Thus $x = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$.

(b) Putting $x_0 = (a_0, a_0)^t$ for some $a_0 \in \mathbb{R}$, we see that $x_n = (a_n, a_n)^t$, $n \in \mathbb{N}$ where

$$a_n = \tfrac{1}{2}(3a_{n-1} + 1) = \tfrac{3}{2}a_{n-1} + \tfrac{1}{2}, n \in \mathbb{N}.$$

Then

$$2a_n = 3a_{n-1} + 1 \quad \Longrightarrow \quad 2a_n + 2 = 3a_{n-1} + 3, \quad n \in \mathbb{N}.$$

Putting $b_n = a_n + 1$, $n \in \mathbb{N} \cup \{0\}$, we get that

$$2b_n = 3b_{n-1}, \quad n \in \mathbb{N}$$

from which an easy induction yields,

$$b_n = \left(\tfrac{3}{2}\right)^n b_0, \quad n \in \mathbb{N}.$$

Then

$$a_n = \left(\tfrac{3}{2}\right)^n (a_0 + 1) - 1, \quad n \in \mathbb{N}.$$

From here it is clear that the solution blows up as $n \to \infty$ unless $a_0 = -1$. Thus the iteration does not converge for any initial vector of the form $x_0 = (a_0, a_0)^t$ for $a_0 \in \mathbb{R} \setminus \{-1\}$.

[Note: a classmate informs me that the iteration diverges for any initial vector other that $x_0 = (-1, -1)^t$ though I couldn't be bothered to prove this.]

Problem F10.8. Let U, W be subspaces of a finite dimensional inner product space V. Prove that $(U \cap W)^\perp = U^\perp + W^\perp$.

Solution. Instead we prove that $U \cap W = (U^\perp + W^\perp)^\perp$. This is sufficient because then

$$(U \cap W)^\perp = ((U^\perp + W^\perp)^\perp)^\perp = U^\perp + W^\perp$$

since the spaces are all finite-dimensional.

Let $x \in U \cap V$. Then $x \in U$ and $x \in V$. If $y \in U^\perp + W^\perp$, then $y = y_1 + y_2$ with $y_1 \in U^\perp, y_2 \in W^\perp$ and we see

$$(x, y) = (x, y_1) + (x, y_2) = 0 + 0 = 0.$$

Hence x is orthogonal to all vectors in $U^\perp + W^\perp$ and so $x \in (U^\perp + W^\perp)^\perp$. Thus $U \cap V \subset (U^\perp + W^\perp)^\perp$.

Let $x \in (U^\perp + W^\perp)^\perp$. Take $y \in U^\perp$. Then $y \in U^\perp + W^\perp$ and so $(x, y) = 0$. Thus x is orthogonal to every member or $U^\perp$ so $x \in (U^\perp)^\perp = U$. Likewise $x \in (W^\perp)^\perp = W$. Thus $x \in U \cap W$ so $(U^\perp + W^\perp)^\perp \subset U \cap W$

Hence $U \cap W = (U^\perp + W^\perp)^\perp$ and the result follows.

Problem S11.2. Show that a positive power of an invertible matrix is diagonalizable if and only if the matrix itself is diagonalizable.

Solution. Suppose that $A \in M_n(\mathbb{C})$ is invertible.

If we suppose that A is diagonalizable, then there is diagonal D and invertible P in $M_n(\mathbb{C})$ such that $A = PDP^{-1}$. But then $A^m = PD^mP^{-1}$ for all $m \in \mathbb{N}$ and D^m is still diagonal so A^m is diagonalizable for all $m \in \mathbb{N}$ (the assumption that A is invertible is not necessary for this direction).

Now suppose there is $m \in \mathbb{N}$ such that A^m is diagonalizable. Then the minimial polynomial of A^m has the form

$$m_{A^m}(x) = (x - \lambda_1) \cdots (x - \lambda_k)$$

where $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ are distinct and $\lambda_i \neq 0$ for $i = 1, \ldots, k$ since A (and thus A^m) is invertible. Let

$$p(x) = m_{A^m}(x^m) = (x^m - \lambda_1) \cdots (x^m - \lambda_k).$$

Then $p(A) = m_{A^m}(A^m) = 0$. Hence, the minimal polynomial of A must divide $p(x)$. Let x_0 be a root of $p(x)$. Then $x_0^m = \lambda_i$ for some $i = 1, \ldots, k$. Suppose that $x_0^m = \lambda_1$ (the proof is identical in other cases, but the details become more tedious to write down). Then

$$\begin{aligned} p'(x_0) = {} & mx_0^{m-1}(x_0^m - \lambda_2) \cdots (x_0^m - \lambda_k) \\ & + mx_0^{m-1}(x_0^m - \lambda_1)(x_0^m - \lambda_3) \cdots (x_0^m - \lambda_k) + \\ & \vdots \\ & + mx_0^{m-1}(x_0^m - \lambda_1) \cdots (x_0^m - \lambda_{k-1}) \end{aligned}$$

where there are k summands and the i^{th} summand omits $(x_0^m - \lambda_i)$ [this is simply the product rule]. However, all of the summands except the first summand go to zero since $x_0^m - \lambda_1 = 0$. Thus

$$p'(x_0) = mx_0^{m-1}(x_0^m - \lambda_2) \cdots (x_0^m - \lambda_k).$$

However, $x_0^m = \lambda_1 \neq \lambda_2, \ldots, \lambda_k$ and $x_0 \neq 0$ since $\lambda_1 \neq 0$. Thus $p'(x_0) \neq 0$ and so $p(x)$ has only simple roots. Since $m_A(x)$ (the minimal polynomial of A) divides $p(x)$ it must have only simple roots as well. Hence

$$m_A(x) = (x - \mu_1) \cdots (x - \mu_\ell)$$

where $\mu_1, \ldots, \mu_\ell$ are distinct. This implies that A is diagonalizable.

Problem S11.5. Let A be an $n \times n$ matrix with real entries and let $b \in \mathbb{R}^n$. Prove that there exists $x \in \mathbb{R}^n$ such that $Ax = b$ if and only if b is in the orthocomplement of the kernel of the transpose of A.

Solution. "There exists $x \in \mathbb{R}^n$ such that $Ax = b$" is an equivalent statement to $b \in \text{Col}(A)$ and "b is in the orthocomplement of the kernel of the transpose of A" is the same as $b \in \text{Null}\,(A^t)^\perp$. So the questions is asking us to prove that

$$\text{Col}(A) = \text{Null}\left(A^t\right)^\perp.$$

We prove this a bit more generally. Let V be an n-dimensional vector space and let $T : V \to V$ be a linear transform with adjoint T^*. We prove that

$$\text{im}(T) = \ker(T^*)^\perp.$$

This clearly subsumes the above equality by letting $V = \mathbb{R}^n$ and $T(x) = Ax, x \in \mathbb{R}^n$.

Let $u \in \ker(T^*)$ and $v \in \text{im}(T)$. Then $T(x) = v$ for some $x \in V$ and

$$(v, u) = (T(x), u) = (x, T^*(u)) = (x, 0) = 0.$$

Hence u is orthogonal to every member of $\text{im}(T)$ so $u \in \text{im}(T)^\perp$. Thus $\ker(T^*) \subset \text{im}(T)^\perp$.

Let $u \in \text{im}(T)^\perp$. Then u is orthogonal to every member of $\text{im}(T)$. But $TT^*(u) \in \text{im}(T)$, so

$$0 = (u, TT^*(u)) = (T^*(u), T^*(u)) = ||T^*(u)|| \quad \Longrightarrow \quad T^*(u) = 0.$$

Thus $u \in \ker(T^*)$ and so $\text{im}(T)^\perp \subset \ker(T^*)$.

We conclude that

$$\text{im}(T)^\perp = \ker(T^*).$$

Since the two subspaces are equal, their orthocomplements are equal and thus

$$(\text{im}(T)^\perp)^\perp = \ker(T^*)^\perp.$$

But for any subspace U of a finite-dimensional inner product space, we have $(U^\perp)^\perp = U$. Thus

$$\text{im}(T) = \ker(T^*)^\perp,$$

which completes the proof.

Note, for a finite dimensional vector space V and a linear operator $T : V \to V$, we have

$$\begin{aligned} \ker(T) &= \text{im}(T^*)^\perp, \\ \ker(T^*) &= \text{im}(T)^\perp, \\ \text{im}(T) &= \ker(T^*)^\perp, \\ \text{im}(T^*) &= \ker(T)^\perp. \end{aligned}$$

If V is not finite dimensional, in general we can only say

$$\begin{aligned} \ker(T) &= \text{im}(T^*)^\perp, \\ \ker(T^*) &= \text{im}(T)^\perp, \\ \text{im}(T) &\subset \ker(T^*)^\perp, \\ \text{im}(T^*) &\subset \ker(T)^\perp. \end{aligned}$$

Problem S11.6. Let V, W be finite dimensional real inner product spaces and let $A : V \to W$ be a linear transform. Fix $w \in W$. Show that the elements $v \in V$ for which the norm $||Av - w||$ is minimal are exactly the solutions to the equation $A^*Av = A^*w$.

Solution. Let $w = w_1 + w_2$ be the orthogonal projection of w onto $\text{im}(A)$. That is, $w_1 \in \text{im}(A)$ and $w_2 \in \text{im}(A)^\perp$ and we know that w_1 is the vector in $\text{im}(A)$ closest to w. That is, the minimizers of $||Av - w||$ are those $v \in V$ such that $Av = w_1$. We must show that these satisfy the normal equation.

Suppose that $v \in V$ is such that $Av = w_1$. We know that $w_2 = w - w_1 = w - Av$ is orthoronal to the image of A. Thus for all $x \in V$,

$$(Ax, Av - w) = 0 \quad \Longrightarrow \quad (x, A^*(Av - w)) = 0).$$

Thus $A^*(Av - w)$ is orthogonal to the whole space V and thus $A^*(Av - w) = 0$ and so $A^*Av = A^*w$.

Conversely, suppose $A^*Av = A^*w$. It suffices to show that $Av = w_1$ because w_1 is the closest member of $\text{im}(A)$ to w. Consider

$$\begin{aligned}
(Av - w_1, Av - w_1) &= (Av - w_1, Av - (w - w_2)) \\
&= (Av - w_1, (Av - w) + w_2) \\
&= (Av, Av - w) + (Av, w_2) - (w_1, Av - w) - (w_1, w_2) \\
&= (v, A^*Av - A^*w) + (Av, w_2) - (w_1, Av - w) - (w_1, w_2).
\end{aligned}$$

Further, we know that w_2 is orthogonal to the image of A. Hence, $(Av, w_2) = (w_1, w_2) = 0$. Also $A^*Av - A^*w = 0$ so

$$(Av - w_1, Av - w_1) = -(w_1, Av - w).$$

However, w_1 is in $\text{im}(A)$. Thus $w_1 = Ax$, for some $x \in V$. Hence

$$(Av - w_1, Av - w_1) = -(Ax, Av - w) = -(x, A^*Av - A^*w) = 0.$$

Thus $||Av - w_1|| = 0$ so $Av = w_1$ and so v is a minimizer of $||Av - w||$.

Problem F11.8. Assume that a complex matrix A satisfies

$$\ker(A - \lambda I) = \ker((A - \lambda I)^2)$$

for all $\lambda \in \mathbb{C}$. Prove from first principles (i.e., without using canonical forms) that A is diagonalizable.

Solution. Let λ be an eigenvalue of A. Suppose λ has algebraic multiplicity more ≥ 2. Then λ is a root of the minimal polynomial more than once. That is

$$m_A(x) = (x - \lambda)^2 q(x)$$

for some polynomial q. Then forany vector v,

$$m_A(A)v = 0 \quad \Longrightarrow \quad (A - \lambda I)^2 q(A)v = 0.$$

But $\ker(A - \lambda I) = \ker((A - \lambda I)^2)$ so

$$(A - \lambda I)q(A)v = 0.$$

Since this holds for all vectors v, this shows that $(A - \lambda I)q(A) = 0$. But this contradicts the minimality of m_A. Thus all eigenvalues of A have algebraic multiplicity 1 so A is diagonalizable.

Problem F11.9. Let V be a finite dimensional complex inner product space and let $L : V \to V$ be a self-adjoint linear operator. Suppose $\mu \in \mathbb{C}$, $\varepsilon > 0$ are given and assume there is a *unit* vector $x \in V$ such that

$$||L(x) - \mu x|| \le \varepsilon.$$

Prove there is an eigenvalue λ of L such that $|\lambda - \mu| \le \varepsilon$.

Solution. By the spectral theorem, there is a orthonormal basis $\{v_1, \ldots, v_n\}$ of V consisting of eigenvectors of L (here, $n = \dim(V)$). Let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues respectively. Then $x = \sum_i (x, e_i)e_i$ and

$$L(x) = \sum_{i=1}^n (x, e_i)L(e_i) = \sum_{i=1}^n (x, e_i)\lambda_i e_i.$$

This leads to

$$\begin{aligned}
||L(x) - \mu x||^2 &= \left|\left|\sum_{i=1}^n (x, e_i)(\lambda_i - \mu)e_i\right|\right|^2 \\
&= \left(\sum_{i=1}^n (x, e_i)(\lambda_i - \mu)e_i, \sum_{j=1}^n (x, e_j)(\lambda_j - \mu)e_j\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n (x, e_i)\overline{(x, e_j)}(\lambda_i - \mu)\overline{(\lambda_j - \mu)}(e_i, e_j) \\
&= \sum_{i=1}^n |(x, e_i)|^2 |\lambda_i - \mu|^2, \qquad \text{by orthonormality.}
\end{aligned}$$

If for every i, we had

$$|\lambda_i - \mu| > \varepsilon$$

then

$$||L(x) - \mu x||^2 > \varepsilon^2 \sum_{i=1}^n |(x, e_i)|^2 = \varepsilon^2 ||x|| = \varepsilon^2.$$

However, this contradicts our assumption that $||L(x) - \mu x|| \le \varepsilon$. Hence there is some $i = 1, \ldots, n$ such that $|\lambda_i - \mu| \le \varepsilon$.

Problem F11.10. Let A be a 3×3 real matrix with $A^3 = I$. Show that A is similar to a matrix of the form
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$
What values of θ are possible?

Solution. From $A^3 = I$, we get $\det(A)^3 = 1$ so $\det(A) = 1$. Also since the characteristic polynomial of A has degree 3, it has a real root. Thus A has a real eigenvalue, λ. We see that λ^3 is an eigenvalue of $A^3 = I$ and thus $\lambda^3 = 1$ and so $\lambda = 1$.

If all eigenvalues of A are real then they are all 1 by the reasoning above. Since $A^3 = I$, we know A is diagonalizable and thus A is similar to I. But this implies $A = I$. Then A is of the above form with $\theta = 0$ so the claim clearly holds.

Otherwise, since complex eigenvalues of a real matrix come in conjugate pairs, besides 1, A has eigenvalues of the form $\mu, \overline{\mu}$ where $\mu \in \mathbb{C}$. Then $1 = \det(A) = \lambda\mu\overline{\mu} = |\mu|^2$ and so $\mu = e^{i\theta} = \cos\theta + i\sin\theta$ for some $\theta \in [0, 2\pi)$. The eigenvectors corresponding to $\mu, \overline{\mu}$ will also be a conjugate pair. Let the eigenvector of μ be $w = w_1 + iw_2$. where $w_1, w_2 \in \mathbb{R}^3$. Then
$$Aw = \mu w \quad \Longrightarrow \quad Aw_1 = \cos\theta w_1 - \sin\theta w_2 \quad \text{and} \quad Aw_2 = \sin\theta w_1 + \cos\theta w_2.$$

Put $P = [v\ w_1\ w_2]$. Then
$$AP = (Av \,|\, Aw_1 \,|\, Aw_2) = (v \,|\, \cos\theta w_1 - \sin\theta w_2 \,|\, \sin\theta w_1 + \cos\theta w_2) = P\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$
which shows that A is similar to a matrix of the desired form. There are no restrictions put on θ throughout the derivations, so it could be any $\theta \in \mathbb{R}$. Of course, $\theta \in [0, 2\pi)$ will suffice.

Problem F11.11.

(a) State and prove the rank-nullity theorem.

(b) Suppose U, V, W are finite dimensional vector spaces over $\mathbb{R}$ and that $T : U \to V$ and $S : V \to W$ are linear operators. Suppose that T is injective, S is surjective and $S \circ T = 0$. Prove that $\operatorname{im}(T) \subset \ker(S)$ and that $\dim(V) - \dim(U) - \dim(W) = \dim(\ker(S)/\operatorname{im}(T))$.

(a) **Rank-Nullity Theorem.** Let U, V be finite dimensional vector spaces and $T : U \to V$ be a linear map. Then
$$\dim(\operatorname{im}(T)) + \dim(\ker(T)) = \dim(U).$$

Proof. Assume $\dim(\ker(T)) = n$ and $\dim(U) = n + k$. Let $\{x_1, \ldots, x_n\}$ be a basis for $\ker(T)$. We can extend this to a basis of U: $\{x_1, \ldots, x_n, u_1, \ldots u_k\}$. Take $v \in \operatorname{im}(T)$.

Then $v = T(u)$ for some $u \in U$. But we have a basis for U, so there are scalars $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_k$ such that

$$u = \alpha_1 x_1 + \cdots + \alpha_n x_n + \beta_1 u_1 + \cdots + \beta_k u_k.$$

Then

$$\begin{aligned} v = T(u) &= \alpha_1 T(x_1) + \cdots + \alpha_n T(x_n) + \beta_1 T(u_1) + \cdots + \beta_k T(u_k) \\ &= \beta_1 T(u_1) + \cdots + \beta_k T(u_k) \quad (\text{since } x_i \in \ker(T), i = 1, \dots, n). \end{aligned}$$

Thus $\{T(u_1), \dots, T(u_k)\}$ is a spanning set for $\mathrm{im}(T)$. Now assume that $\gamma_1, \dots, \gamma_k$ are scalars such that

$$\gamma_1 T(u_1) + \cdots + \gamma_k T(u_k) = 0.$$

Then

$$T(\gamma_1 u_1 + \cdots + \gamma_k u_k) = 0 \quad \Longrightarrow \quad \gamma_1 u_1 + \cdots + \gamma_k u_k \in \ker(T).$$

Then there are scalars $\delta_1, \dots, \delta_n$,

$$\gamma_1 u_1 + \cdots + \gamma_k u_k = \delta_1 x_1 + \cdots + \delta_n x_n \quad \Longrightarrow \quad \gamma_1 u_1 + \cdots + \gamma_k u_k - \delta_1 x_1 - \cdots - \delta_n x_n = 0.$$

But $\{x_1, \dots, x_n, u_1, \dots u_k\}$ form a basis and are thus linearly independent. Hence, $\gamma_1 = \cdots = \gamma_k = 0$ so $\{T(u_1), \dots, T(u_k)\}$ is a linearly independent set.

Thus $\{T(u_1), \dots, T(u_k)\}$ is a basis for $\mathrm{im}(T)$ so $\dim(\mathrm{im}(T)) = k$ and the result is proven.

(b) Take $x \in \mathrm{im}(T)$. Then $x = T(u)$ for some $u \in U$. But then $S(x) = (S \circ T)(u) = 0$ since $S \circ T$ is the zero transformation. Thus $x \in \ker(S)$ so $\mathrm{im}(T) \subset \ker(S)$.

Apply the rank-nullity theorem to T and S to see

$$\begin{aligned} \dim(\mathrm{im}(T)) + \dim(\ker(T)) &= \dim(U), \\ \dim(\mathrm{im}(S)) + \dim(\ker(S)) &= \dim(V). \end{aligned}$$

But $\dim(\ker(T)) = 0$ since T is injective and $\dim(\mathrm{im}(S)) = \dim(W)$ because S is surjective. Hence

$$\begin{aligned} \dim(\mathrm{im}(T)) &= \dim(U), \\ \dim(\ker(S)) &= \dim(V) - \dim(W). \end{aligned}$$

Then

$$\dim(\ker(S)/\mathrm{im}(T)) = \dim(\ker(S)) - \dim(\mathrm{im}(T)) = \dim(V) - \dim(W) - \dim(U)$$

which is exactly the conclusion we needed to prove.

Problem S12.7. Let F be a finite field of p elements and V be an n-dimensional vector space over F. Compute the number of invertible linear maps from $V \to V$.

Solution. The matrix form of any invertible map must be invertible and thus must have linearly independent columns. The choice for elements in the first column is arbitrary except they can't all be zero. Thus there are $p^n - 1$ choices. The choice for elements in the next column simply cannot be a scalar multiple of the first. Thus there are $p^n - p$ choices. The next column cannot be a linear combination of the first two so there are $p^n - p \cdot p = p^n - p^2$ choices. Similarly, there are $p^n - p^{k-1}$ choices for the k^{th} column. Multiplying these, we see there are

$$(p^n - 1)(p^n - p)\cdots(p^n - p^{n-1}) = \prod_{k=0}^{n-1}(p^n - p^k)$$

invertible linear maps on V.

Problem S12.9. Let $a_1 = 1, a_2 = 4, a_{n+2} = 4a_{n+1} - 3a_n$ for all $n \geq 1$. Find a 2×2 matrix A such that

$$A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}.$$

Use the eigenvalues of A to determine the limit

$$\lim_{n\to\infty} (a_n)^{1/n}.$$

Solution. From $n = 1$, we see that A has the form

$$A = \begin{pmatrix} 4 & b \\ 1 & d \end{pmatrix}.$$

Then

$$A^2 = \begin{pmatrix} 16 + b & 4b + bd \\ 4 + d & b + d^2 \end{pmatrix}.$$

From $n = 2$,

$$A^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 13 \\ 4 \end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix} 16 + b \\ 4 + d \end{pmatrix} = \begin{pmatrix} 13 \\ 4 \end{pmatrix},$$

so

$$A = \begin{pmatrix} 4 & -3 \\ 1 & 0 \end{pmatrix}.$$

We see that $\det(A - \lambda I) = \lambda(\lambda - 4) + 3 = (\lambda - 1)(\lambda - 3)$. Thus $\lambda_1 = 1$ and $\lambda_2 = 3$ are eigenvalues of A with corresponding eigenvectors

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Notice that

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 3 \\ 1 \end{pmatrix} := x_1 + x_2,$$

where x_1, x_2 are still eigenvectors corresponding to λ_1, λ_2 respectively. Finally, for any $n \geq 1$,

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = A^n(x_1 + x_2) = \lambda_1^n x_1 + \lambda_2^n x_2,$$

which gives $a_n = \frac{1}{2}3^n - \frac{1}{2}$. Thus

$$\lim_{n\to\infty} (a_n)^{1/n} = 3.$$

Problem S12.11.

(a) Find a polynomial $P(x)$ of degree 2 such that $P(A) = 0$ for

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}.$$

(b) Prove that $P(x)$ from part (a) is unique up to scalar multiplication.

Solution.

(a) The characteristic polynomial of A is $P(x) = (x-1)(x-2) - 12 = x^2 - 3x - 10$. We check

$$P(A) = A^2 - 3A - 10I = \begin{pmatrix} 13 & 9 \\ 12 & 16 \end{pmatrix} - \begin{pmatrix} 3 & 9 \\ 12 & 6 \end{pmatrix} - \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

(We knew to check the characteristic polynomial because the Cayley-Hamilton Theorem tells us that the characteristic polynomial of a matrix always annihilates the matrix.)

(b) It is clear that there is no first degree polynomial Q such that $Q(A) = 0$ because A is not a scalar multiple of the identity.

Suppose $Q(A) = 0$. By the above, this implies that Q is a second degree polynomial. Then there is $\alpha \in \mathbb{R}$ such that αQ is a monic polynomial. Then $(P - \alpha Q)(A) = 0$ and $P - \alpha Q$ is a first degree polynomial or a constant polynomial. However, as stated above, there is no first degree polynomial which annihilates A. Hence $P - \alpha Q$ is constant so $(P - \alpha Q)(A) = 0 \implies P - \alpha Q = 0$. Thus $P = \alpha Q$ so the polynomial is unique up to multiplication by a constant.

Problem F12.7. Let A be an invertible $m \times m$ matrix over $\mathbb{C}$ and suppose the set of powers A^n of A is bounded for $n \in \mathbb{Z}$. Prove that A is diagonalizable.

Solution. Let $\lambda \in \mathbb{C}$ be an eigenvalue of A. Then $\lambda \neq 0$ since A is invertible. Further, there is a unit vector $v \in \mathbb{C}$ such that $Av = \lambda v$. Then $A^n v = \lambda^n v$ for all $n \in \mathbb{Z}$. Then we see

$$||A^n|| \geq ||A^n v|| = ||\lambda^n v|| = |\lambda|^n.$$

If $|\lambda| \neq 1$, then $|\lambda|^n$ is unbounded for $n \in Z$ and thus A^n would be unbounded for $n \in \mathbb{Z}$; a contradiction. Thus all eigenvalues λ of A satisfy $||\lambda|| = 1$.

Assume A is not diagonalizable. Then the Jordan canonical form of A has a block of at least size 2. Then there is an eigenvalue λ of A and two vectors v, w such that $Av = \lambda v$ and $Aw = v + \lambda w$. This implies

$$A^2 w = Av + \lambda Aw = \lambda v + \lambda(v + \lambda w) = 2\lambda v + \lambda^2 w.$$

Further

$$A^3 w = 2\lambda Av + \lambda^2 Aw = 2\lambda^2 v + \lambda^2(v + \lambda w) = 3\lambda^2 v + \lambda^3 w.$$

Indeed, proceeding by induction, we have

$$A^n w = n\lambda^{n-1} + \lambda^n w.$$

Then

$$||A^n|| \geq ||A^n w|| = \left|\left|n\lambda^{n-1} v + \lambda^n w\right|\right| \geq \left|n\lambda^{n-1}\right| ||v|| - |\lambda^n| \, ||w|| = n \, ||v|| - ||w|| \, .$$

Since this holds for all $n \in \mathbb{N}$, letting $n \to \infty$ shows that the powers of A are not bounded, a contradiction. Thus A is diagonalizable.

Problem F12.10. Let A be a linear operator on a four dimensional complex vector space that satisfies the polynomial equation $P(A) = A^4 + 2A^3 - 2A - I = 0$, there I is the identity operator on V. Suppose that $|\text{tr}(A)| = 2$ and that $\dim(\text{range}(A + I)) = 2$. Give a Jordan canonical form of A.

Solution. Let's factor $P(x)$. By inspection, we see that 1 is a root. So $P(x) = (x-1)(x^3 + \alpha x^2 + \beta x + 1)$. Expanding, we see $\alpha - 1 = 2$ and $\beta - \alpha = 0$ (from the x^3 and x^2 term respectively). Then $\alpha = \beta = 3$ so $P(x) = (x-1)(x+1)^3$. Thus possible eigenvalues of A are $1, -1$. By the rank-nullity theorem, $\dim(\ker(A + I)) = 4 - \dim(\text{range}(A + I)) = 2$. Thus the algebraic multiplicity of -1 is at least 2. Then $\text{tr}(A) = \pm 2$ forces -1 to have algebraic multiplicity 3. Thus A is not diagonalizable. A canonical form has 3 Jordan blocks; two of these are 1×1 blocks containing 1 and -1. The third block is 2×2 with -1 as the two diagonal elements and 1 on the superdiagonal. That is a canonical form J of A is

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Problem F12.12. Let M be an $n \times m$ matrix. Prove that the row rank of M equals the column rank of M. Interpret this result as an equality of the dimensions of two vector spaces naturally attached to the map defined by M.

Solution. Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be the associated linear map: $T(x) = Mx$, $x \in \mathbb{R}^m$. Also define $T^* : \mathbb{R}^n \to \mathbb{R}^m$ by $T^*(y) = M^t y$, $y \in \mathbb{R}^n$.

Then $\text{im}(T) = \text{col}(M)$ and $\text{im}(T^*) = \text{col}(M^t) = \text{row}(M)$. Thus the problem boils down to showing that $\dim(\text{im}(T)) = \dim(\text{im}(T^*))$.

Let $y_1, \ldots, y_k \in \mathbb{R}^n$ be a basis for $\text{im}(T)$. Then $T^*(y_1), \ldots, T^*(y_k) \in \text{im}(T^*)$. Take $\alpha_1, \ldots \alpha_k \in \mathbb{R}$ such that

$$\alpha_1 T^*(y_1) + \cdots + \alpha_k T^*(y_k) = 0.$$

Then

$$T^*(\alpha_1 y_1 + \cdots + \alpha_k y_k) = 0$$

so $\alpha_1 y_1 + \cdots + \alpha_k y_k \in \ker(T^*)$. Recall that $\ker(T^*) = \text{im}(T)^\perp$, so $\alpha_1 y_1 + \cdots + \alpha_k y_k \in \text{im}(T)^\perp$. However since $y_1, \ldots, y_k \in \text{im}(T)$, we know that $\alpha_1 y_1 + \cdots + \alpha_k y_k \in \text{im}(T)$. Hence, then

$$\alpha_1 y_1 + \cdots \alpha_k y_k \in \text{im}(T) \cap \text{im}(T)^\perp$$

so

$$\alpha_1 y_1 + \cdots \alpha_k y_k = 0.$$

But $y_1, \ldots, y_k$ form a basis, so $\alpha_1 = \cdots = \alpha_k = 0$. Hence $T^*(y_1), \ldots, T^*(y_k)$ are linearly independent so $\dim(\text{im}(T^*)) \geq k = \dim(\text{im}(T))$.

Using the same argument but starting with a basis of $\text{im}(T^*)$ shows that $\dim(\text{im}(T^*)) \leq \dim(\text{im}(T))$, thus the values are equal and so the row rank of M is equal to its column rank.

The interpretation as an equality of dimensions of two vector spaces naturally attached to M would be

$$\dim(\text{im}(T)) = \dim\left(\ker(T)^\perp\right) \quad \text{or} \quad \dim(\text{col}(M)) = \dim\left(\text{null}(M)^\perp\right)$$

since T is the left multiplication operator for M.

Problem S13.8. Let V, W be finite dimensional inner product space and $T : V \to W$ be a linear map.

(a) Define the adjoint map $T^* : W \to V$.

(b) Show that if the matrices are written relative to orthonormal bases of V and W then the matrix of T^* is the transpose of the matris of T.

(c) Show that the kernel of T^* is the orthogonal complement of the range of T.

(d) Use (b) and (c) to prove that the row rank of a matrix is the same as the column rank of the matrix.

Solution.

(a) The adjoint $T^* : W \to V$ is the unique linear map such that

$$(Tv, w)_W = (v, T^* w)_V$$

for all $v \in V$, $w \in W$, where $(\cdot, \cdot)_V, (\cdot, \cdot)_W$ are the inner products on V and W respectively.

(b) Let $\{v_1, \ldots, v_n\}$ be an orthonormal basis for V and $\{w_1, \ldots, w_m\}$ be an orthonormal basis for W. For each v_j, there are scalars $\alpha_{1j}, \ldots, \alpha_{mj}$ such that

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Then the matrix of T is given by $[T] = (a_{ij})$ where $i = 1, \ldots, m, j = 1, \ldots, n$. Similarly, for each w_i there are scalars $\beta_{1i}, \ldots, \beta_{ni}$ such that

$$T^*(w_i) = \sum_{j=1}^{n} \beta_{ji} v_j.$$

Then $[T^*] = (\beta_{ji})$. We need to show that $\beta_{ji} = \alpha_{ij}$ and this will imply that $[T] = [T^*]^t$. We see by orthogonality that, $\alpha_{ij} = (Tv_j, w_i)$. Then $\alpha_{ij} = (v_j, T^* w_i)$ by definition of the adjoint. But then by orthogonality, $\alpha_{ij} = \beta_{ji}$. Hence $[T] = [T^*]^t$.

(c) Take $w \in \ker(T^*)$. Let $z \in \operatorname{im}(T)$. Then $z = Tv$ for some $v \in V$. Then

$$(z, w) = (Tv, w) = (v, T^* w) = (v, 0) = 0.$$

Thus w is orthogonal to z. Since $z \in \operatorname{im}(T)$ was arbitrary, this shows that $w \in \operatorname{im}(T)^\perp$.

Take $w \in \operatorname{im}(T)^\perp$. Then w is orthogonal to every member of the image of T. In particular $T(T^* w) \in \operatorname{im}(T)$. Thus

$$0 = (T(T^* w), w) = (T^* w, T^* w) = ||T^* w||^2 \quad \Longrightarrow \quad T^* w = 0.$$

Thus $w \in \ker(T^*)$.

We conclude that $\ker(T^*) = \operatorname{im}(T)^\perp$.

(d) Translating (c) into matrix form using (b), we see that for $A \in M_{n,m}(\mathbb{C})$, we have

$$\operatorname{null}(A^t) = \operatorname{col}(A)^\perp.$$

We see that $\operatorname{col}(A^t) = \operatorname{row}(A)$. So we must show that $\operatorname{col}(A) = \operatorname{col}(A^t)$.

Let $\{x_1, \ldots, x_k\}$ be a basis for the column space of A. Then $A^t x_1, \ldots, A^t x_k$ are in the column space of A^t. Let $\alpha_1, \ldots, \alpha_k \in \mathbb{C}$ be such that

$$\alpha_1 A^t x_1 + \cdots + \alpha_k A^t x_k = 0.$$

Then

$$A^t(\alpha_1 x_1 + \cdots + \alpha_k x_k) = 0$$

so $\alpha_1 x_1 + \cdots + \alpha_k x_k \in \operatorname{null}(A^t)$. But by (c), this implies that $\alpha_1 x_1 + \cdots + \alpha_k x_k \in \operatorname{col}(A)^\perp$. However $\alpha_1 x_1 + \ldots + \alpha_k x_k$ is also in $\operatorname{col}(A)$. Thus it is orthogonal to itself and so it is zero. But, $x_1, \ldots, x_k$ are linearly independent so $\alpha_1 = \cdots = \alpha_k = 0$. Hence $A^t x_1 + \cdots + A^t x_k$ are linearly independent in $\operatorname{col}(A^t) = \operatorname{row}(A)$. Thus $\dim(\operatorname{row}(A)) \geq k = \dim(\operatorname{col}(A))$.

Making the same argument but beginning with a basis for the column space of A^t shows that $\dim(\operatorname{row}(A)) \leq \dim(\operatorname{col}(A))$. Thus $\dim(\operatorname{row}(A)) = \dim(\operatorname{col}(A))$ so the row rank and column rank of A are the same.

Problem S14.1.

(a) Find a real matrix A whose minimal polynomial is equal to

$$t^4 + 1.$$

(b) Show that the real linear map determined by A has no non-trivial invariant subspace.

Solution.

(a) Put

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \end{pmatrix}.$$

Then

$$A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad A^4 = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Thus $A^4 + I = 0$. Further, A^3, A^2, A, I are linearly independent, so there is no third degree polynomial which annihilates A. Thus $t^4 + 1$ is the minimal polynomial of A.

(b) This isn't true. For example, if

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix}$$

then A still has minimal polynomial $t^4 + 1$ but $\text{Span}\,(e_1, e_2, e_3, e_4)$ is invariant under A. Even if A has to be 4×4, we can only prove that A has no 1 or 3 dimensional invariant subspaces. There are cases where A has a 2 dimensional invariant subspace.

Problem S14.2. Suppose that $S, T : V \to V$ are linear where V is a finite dimensional vector space over $\mathbb{R}$. Show that

$$\dim(\text{im}(S)) + \dim(\text{im}(T)) \le \dim(\text{im}(S \circ T)) + \dim(V).$$

Solution. Adding $\dim(\ker(S))$ and $\dim(\ker(T))$ to both sides and using the rank nullity theorem, we see that the given inequality is equivalent to

$$\dim(V) \le \dim(\text{im}(S \circ T)) + \dim(\ker(S)) + \dim(\ker(T)).$$

Next adding $\dim(\ker(S \circ T))$ to both sides, we see the original inequality is equivalent to

$$\dim(\ker(S \circ T)) \leq \dim(\ker(S)) + \dim(\ker(T)).$$

If we can prove this last inequality, we will have proven the first.

Consider, if $x \in \ker(T)$ then $(S \circ T)(x) = S(T(x)) = S(0) = 0$ so $x \in \ker(S \circ T)$. Thus $\ker(T)$ is a subspace of $\ker(S \circ T)$. Let $\{x_1, \ldots, x_k\}$ be a basis of $\ker(T)$. Extend this to a basis $\{x_1, \ldots, x_k, y_1, \ldots, y_\ell\}$ of $\ker(S \circ T)$. Then

$$(S \circ T)(y_i) = S(T(y_i)) = 0$$

for each i, so $T(y_i) \in \ker(S)$. Suppose $\alpha_1, \ldots, \alpha_\ell \in \mathbb{R}$ are such that

$$\alpha_1 T(y_1) + \cdots + \alpha_\ell T(y_\ell) = 0.$$

Then

$$T(\alpha_1 y_1 + \cdots \alpha_\ell y_\ell) = 0$$

so $\alpha_1 y_1 + \cdots \alpha_\ell y_\ell \in \ker(T)$. Then there are $\beta_1, \ldots, \beta_k \in \mathbb{R}$ such that

$$\alpha_1 y_1 + \cdots + \alpha_\ell y_\ell = \beta_1 x_1 + \cdots + \beta_k x_k \quad \Longrightarrow \quad \alpha_1 y_1 + \cdots + \alpha_\ell y_\ell - \beta_1 x_1 - \cdots - \beta_k x_k = 0.$$

But these vectors form a basis for $\ker(S \circ T)$ so they are linearly independent. Hence all coefficients are zero. Thus $\{T(y_1), \ldots, T(y_\ell)\}$ is a lineary independent set in $\ker(S)$. Hence $\dim(\ker(S)) \geq \ell$. Then

$$\dim(\ker(S \circ T)) = \ell + k = \ell + \dim(\ker(T)) \leq \dim(\ker(S)) + \dim(\ker(T)).$$

The result follows.

Problem S14.3. Suppose that $A, B \in M_n(\mathbb{C})$ satisfy $AB - BA = A$. Show that A is not invertible.

Solution. Suppose that A is invertible. Let $\lambda_1, \ldots, \lambda_\ell \in \mathbb{C}$ be the distinct eigenvalues of B ordered such that

$$\mathrm{Re}(\lambda_1) \leq \mathrm{Re}(\lambda_2) \leq \cdots \leq \mathrm{Re}(\lambda_\ell).$$

Multiplying by A^{-1} on the right, we see

$$ABA^{-1} - B = I \qquad \Longrightarrow \qquad ABA^{-1} = B + I.$$

Thus B is similar to $B+I$. However, $B+I$ has $\lambda_\ell+1$ as an eigenvalue and $\mathrm{Re}(\lambda_\ell+1) > \mathrm{Re}(\lambda_j)$ for all $j = 1, \ldots, \ell$. Thus B and $B + I$ do not have the same eigenvalues and thus are not similar; a contradiction. Hence A is not invertible.

Problem S14.4. Suppose $A, B \in M_n(\mathbb{C})$. Show that the characteristic polynomials of AB and BA are equal.

Solution. Suppose B is invertible. Then

$$AB = (B^{-1}B)AB = B^{-1}(BA)B.$$

Thus AB and BA are similar so they have the same characteristic polynomial.

If B is not invertible, let λ be the non-zero eigenvalue of B with least real part. Then for $0 < t < |\mathrm{Re}(\lambda)|$, $B_t = B + tI$ is invertible [note: if all eigenvalues of B have zero real part, then this holds for all $t > 0$]. Then AB_t has the same characteristic polynomial as B_tA. However, the characteristic polynomials of a matrix is a continuous function of the matrix itself [this is because the determinant map is smooth]. Thus taking the limit as $t \to 0$, we see that AB and BA have the same characteristic polynomial.

Problem S14.6. Show that if $A \in M_n(\mathbb{C})$ is normal then $A^* = P(A)$ for some $P \in \mathbb{C}[x]$.

Solution. Since A is normal, by the spectral theorem, we can unitarily diagonalize it: $A = UDU^*$ where U is unitary, D is diagonal. Then for any polynomial Q, we have $Q(A) = UQ(D)U^*$. Thus we reduce the problem to finding $P \in \mathbb{C}[x]$ such that

$$U\overline{D}U^* = UP(D)U^* \quad \Longleftrightarrow \quad \overline{D} = P(D).$$

However, a polynomial acting on a diagonal matrix acts individually on each diagonal element. Let $\lambda_1, \ldots, \lambda_\ell \in \mathbb{C}$ be the distinct eigenvalues of A. Then these are the diagonal elements of D and the $\overline{\lambda}_1, \ldots, \overline{\lambda}_\ell$ are the diagonal elements of $\overline{D}$. Thus all we need a polynomial which satisfies $P(\lambda_j) = \overline{\lambda}_j$ for all $1, \ldots, \ell$. Such a polynomial certainly exists and can be constructed using Lagrange interpolants. Hence A^* can be expressed as a polynomial in A.

Problem F14.7. Among all solutions to the system

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 5 & 7 \\ -2 & -1 & 1 & 3 \end{pmatrix} x = \begin{pmatrix} 2 \\ 7 \\ -1 \end{pmatrix}$$

find the solution with minimal length.

Solution. We notice

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 2 & 3 & 5 & 7 & 7 \\ -2 & -1 & 1 & 3 & -1 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 3 & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Thus all solutions of the given system are of the form

$$x = \begin{pmatrix} -1 \\ 3 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 2 \\ -3 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 4 \\ -5 \\ 0 \\ 1 \end{pmatrix}$$

for some $s, t \in \mathbb{R}$ (in fact we could replace $(-1, 3, 0, 0)^t$ with any particular solution). Thus we need to minimize

$$f(s,t) = (-1 + 2s + 4t)^2 + (3 - 3s - 5t)^2 + s^2 + t^2$$

among all $(s, t) \in \mathbb{R}^2$. We know the minima must annihilate the first derivatives, so

$$4(-1 + 2s + 4t) - 6(3 - 3s - 5t) + 2s = 0 \quad \text{and} \quad 8(-1 + 2s + 4t) - 10(3 - 3s - 5t) + 2t = 0.$$

Simplifying, this gives

$$\begin{pmatrix} 1 & \frac{23}{14} \\ 1 & \frac{42}{23} \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} \frac{11}{14} \\ \frac{19}{23} \end{pmatrix}.$$

After copious amounts of infuriating and mind-numbing algebra, this yields

$$s = \frac{25}{59}, \quad t = \frac{13}{59}$$

so the solution of minimal length is

$$x = \frac{1}{59} \begin{pmatrix} 43 \\ 37 \\ 25 \\ 13 \end{pmatrix}.$$

Problem F14.8. Compute the eigenvalues of the $n \times n$ matrix

$$M = \begin{pmatrix} k & 1 & 1 & \cdots & 1 \\ 1 & k & 1 & \cdots & 1 \\ 1 & 1 & k & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & k \end{pmatrix}.$$

Use the eigenvalues to compute $\det(M)$.

Solution. We notice that $\lambda = k - 1$ is an eigenvalue and $M - \lambda I$ has rank 1. Thus the algebraic multiplicity of λ is $n - 1$. Hence we only need one more eigenvalue. By inspection if $x = (1, 1, \ldots, 1)^t$ then $Mx = (k + (n - 1))x$. Thus $k + (n - 1) = (k - 1) + n$ is another eigenvalue.

The determinant is the product of the eigenvalues so $\det(M) = (k - 1)^n + n(k - 1)^{n-1}$.

[Note: a less "inspective" approach might use induction, but I couldn't figure out how to do that.]

Problem F14.9. Suppose $A \neq 0$ is an $n \times n$ complex matrix. Prove that there is a matrix B such that B and $A + B$ have no eigenvalues in common.

Solution. Not sure.

Problem F14.10. What is the largest possible number of 1's an invertible $n \times n$ matrix with entries in $\{0, 1\}$ can have? You must show this number is possible and that no larger number is possible.

Solution. The answer is $n^2 - n + 1$.

First we show that no larger number is possible. Suppose an $n \times n$ matrix with entries in $\{0, 1\}$ has at least $n^2 - n + 2$ entries that are 1. Then there are less than $n - 2$ zeros in the matrix. But this means that at least two columns do not have a zero. These columns will be the linearly dependent and thus the matrix is not invertible.

Consider the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 & 1 \end{pmatrix}.$$

That is, A is a matrix full of 1's but with zeros on the first subdiagonal. Then there are $n^2 - n + 1$ entries that are 1. Consider solving $Ax = 0$. By subtracting the second row from the first we would find $x_1 = 0$. Then by subtracting the third row from the second, we would find $x_2 = 0$. Likewise we would find $x_k = 0, k = 1, \ldots, n$. Thus $Ax = 0$ has only the trivial solution so A is invertible. Thus we can find an invertible matrix with $n^2 - n + 1$ entries which are 1.

Problem F14.11. Suppose a 4×4 integer matrix has four distinct real eigenvalues $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$. Prove that $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 \in \mathbb{Z}$.

Solution. Recall that the trace of a matrix is the sum of the eigenvalues. Since A has integer entries so does A^2. Thus the trace of A^2 is an integer since the diagonal elements of A^2 are all integers. However, the eigenvalues of A^2 are $\lambda_1^2, \lambda_2^2, \lambda_3^2, \lambda_4^2$ so their sum must equal the trace of A^2 and hence $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 \in \mathbb{Z}$.

[Apparently the assumptions that the eigenvalues are real and distinct aren't necessary.]

Problem F14.12. Prove that the matrix $A = (a_{ij})$ given by $a_{ij} = \frac{1}{i+j-1}, \quad i, j = 1, \ldots, n$ is positive definite.

Solution. The matrix is the Gram matrix for the basis $\{1, x, \ldots, x^{n-1}\}$ of $\mathbb{P}^{n-1}[0, 1]$ with the inner product

$$(p, q) = \int_0^1 p(x)q(x)dx, \quad p, \in \mathbb{P}^{n-1}[0, 1].$$

All Gram matrices are positive definite. To see this, let $\{v_1, \ldots, v_n\}$ be a linearly independent set in an inner product space and let A be the matrix given by

$$A_{ij} = (v_i, v_j).$$

Then for any $x \in \mathbb{R}^n, x = (x_1, \dots, x_n)^t$, we see

$$x^t A x = \sum_{i,j=1}^n x_i x_j (v_i, v_j) = \sum_{i,j=1}^n (x_i v_i, x_j v_j) = \left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n x_j v_j \right) = (v, v) \geq 0$$

where $v = \sum_{i=1}^n x_i v_i$. Further, there is equality iff $v = 0$ which happens iff $x = 0$ since $\{v_1, \dots, v_n\}$ is a linearly independent set. Thus A is positive definite.

Problem S15.7. Let

$$f(x, y, z) = 9x^2 + 6y^2 + 6z^2 + 12xy - 10xz - 2yz.$$

Does there exist a point (x, y, z) such that $f(x, y, z) < 0$?

Solution. I assume they mean $(x, y, z) \in \mathbb{R}^3$. In this case the answer is no. We see that any critical point (x, y, z) of f must satisfy

$$\begin{aligned} 0 &= \frac{\partial f}{\partial x}(x, y, z) = 18x + 12y - 10z, \\ 0 &= \frac{\partial f}{\partial y}(x, y, z) = 12x + 12y - 2z, \\ 0 &= \frac{\partial f}{\partial z}(x, y, z) = -10x - 2y + 12z. \end{aligned}$$

However, the matrix

$$A = \begin{bmatrix} 9 & 6 & -5 \\ 6 & 6 & -1 \\ -5 & -1 & 6 \end{bmatrix}$$

is invertible since

$$\det(A) = 9(37) - 6(41) - 5(24) = 333 - 246 - 120 = 333 - 366 = -33.$$

Thus the only solution to the above system is $(0, 0, 0)$. Thus $(0, 0, 0)$ is either a global maximum or a global minimum. It is easy to see that $f(1, 1, 1) > 0$ whereas $f(0, 0, 0) = 0$ so $(0, 0, 0)$ must be a global minimum. Hence there is no point (x, y, z) such that $f(x, y, z) < 0$.

[I'm not sure what the indented solution was here but I'm fairly certain the correct approach was to factor the polynomial into a sum of squares. I couldn't figure out how to do this. The function f is the quadratic form which is induced by A as defined above so that problem is actually to prove that this matrix is positive definite.]

Problem S15.8. Prove or disprove the following claims:

(a) Matrices with determinant 1 are dense in the set of all 3×3 real matrices.

(b) Matrices with distinct eigenvalues are dense in the set of 3×3 complex matrices.

Solution.

(a) Matrices with determinant 1 <u>are not</u> dense in the set of 3×3 real matrices. The determinant map is a C^∞-smooth map. Thus if the determinant was 1 on a dense set, then every matrix would have determinant 1.

(b) Matrices with distinct eigenvalues <u>are</u> dense in the set of 3×3 complex matrices. We prove this for upper triangular matrices first. Let $\varepsilon > 0$ and let T be an upper triangular 3×3 matrix. We show there is a matrix with distinct eigenvalues within ε of T. If T already has distinct eigenvalues, we're done. Otherwise, there are two cases we consider:

Case 1: All diagonal entries are T are the same value t. In this case, let $D_\varepsilon = \text{diag}(\varepsilon/\sqrt{4}, \varepsilon/\sqrt{6}, \varepsilon/\sqrt{12})$. Then $T + D_\varepsilon$ has distinct eigenvalues and

$$||T - (T + D_\varepsilon)|| = ||D_\varepsilon|| = \sqrt{\varepsilon^2/4 + \varepsilon^2/6 + \varepsilon^2/12} = \sqrt{\varepsilon^2/2} = \varepsilon/\sqrt{2} < \varepsilon.$$

Case 2: Two of the diagonal entries of T are the same while the other is different. Suppose without loss of generality that the first two are the same while the third is different. Call the eigenvalues a and b. Let $\delta = |b - a|$. Set $D_\varepsilon = \text{diag}(t, -t, t)$ where $t < \min\{\delta/3, \varepsilon/\sqrt{5}\}$. Then $T + D_\varepsilon$ has distinct eigenvalues and

$$||T - (T + D_\varepsilon)|| = ||D_\varepsilon|| < \sqrt{3\varepsilon^2/5} = \varepsilon\sqrt{3/5} < \varepsilon.$$

Thus the claim holds for upper triangular matrices. However, every matrix is similar to an upper triangular matrix by the Schur decomposition (see **W02.11**). Thus adding to the diagonal of the original matrix changes the eigenvalues the in the same way as adding to the diagonal of upper triangular matrix to which the original is similar. Hence the claim holds for all matrices.

Problem S15.9. Let $V = \mathbb{R}^n$ and let U_1, U_2, W_1, W_2 be subspaces of V of dimension d such that $\dim(U_1 \cap W_1) = \dim(U_2 \cap W_2) = \ell$, $\ell \le d \le n$. Prove that there is a linear operator $T : V \to V$ such that $T(U_1) = U_2$ and $T(W_1) = W_2$.

Solution. Not sure.

Problem 15.10. Let

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \text{and} \quad M^{-1} = \begin{pmatrix} P & Q \\ R & S \end{pmatrix}$$

where A, B, C, D, P, Q, R, S are all $k \times k$ matrices. Show that

$$\det(M) \cdot \det(S) = \det(A).$$

Solution. Not sure.

Problem S15.11. Two matrices A, B are called commuting if $AB = BA$. The order of a matrix A is defined to be the smallest non-negative integer k such that $A^k = I$; if no such k

exists, the matrix is said to have infinite order. Prove that there exist ten distinct real 2×2 matrices which are pairwise commuting and have the same finite order.

Solution. Matrices of the form
$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$
always commute and have the property that
$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^n = \begin{pmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{pmatrix}.$$
Choosing $\theta = 2\pi k/11$, $k = 1, 2, \ldots, 11$ gives ten different matrices, all of order 11 which pairwise commute.

Note: the reason you should think of these matrices is because, together with the identity, they form a group which is isomorphic to the subgroup of D_{11} (the symmetries of the regular 11-gon) consisting of rotations.

Problem S15.12. Let
$$M = \begin{pmatrix} 3 & 5 \\ 1 & -1 \end{pmatrix}.$$

(a) Compute $\exp(M)$.

(b) Is there a real matrix A such that $M = \exp(A)$?

Solution.

(a) To compute $\exp(M)$ we first find a Jordan form for M. The characteristic polynomial of M is
$$p_M(t) = (3-t)(-1-t) - 5 = t^2 - 2t - 8 = (t-4)(t+2).$$
Thus M has distinct eigenvalues and so it is diagonalizable. An eigenvector corresponding to $\lambda_1 = 4$ is given by $v_1 = (5, 1)^t$ and an eigenvector corresponding to $\lambda_2 = -2$ is $v_2 = (1, -1)^t$. Putting $P = [v_1 \;\; v_2]$, we see
$$\begin{aligned} P^{-1}MP &= -\frac{1}{6}\begin{pmatrix} -1 & -1 \\ -1 & 5 \end{pmatrix}\begin{pmatrix} 3 & 5 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix} \\ &= -\frac{1}{6}\begin{pmatrix} -1 & -1 \\ -1 & 5 \end{pmatrix}\begin{pmatrix} 20 & -2 \\ 4 & 2 \end{pmatrix} \\ &= -\frac{1}{6}\begin{pmatrix} -24 & 0 \\ 0 & 12 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -2. \end{pmatrix}. \end{aligned}$$
So $M = PDP^{-1}$ where $D = \mathrm{diag}(4, -2)$. Then
$$\begin{aligned} \exp(M) = P\exp(D)P^{-1} &= -\frac{1}{6}\begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} e^4 & 0 \\ 0 & e^{-2} \end{pmatrix}\begin{pmatrix} -1 & -1 \\ -1 & 5 \end{pmatrix} \\ &= -\frac{1}{6}\begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} -e^4 & -e^4 \\ -e^{-2} & 5e^{-2} \end{pmatrix} \\ &= -\frac{1}{6}\begin{pmatrix} -5e^4 - e^{-2} & -5e^4 + 5e^{-2} \\ -e^4 + e^{-2} & -e^4 - 5e^{-2} \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 5e^4 + e^{-2} & 5e^4 - 5e^{-2} \\ e^4 - e^{-2} & e^4 + 5e^{-2} \end{pmatrix}. \end{aligned}$$

(b) No. The eigenvalues of $\exp(A)$ are $e^{\lambda_1}, e^{\lambda_2}$ where λ_1, λ_2 are the eigenvalues of A.

Suppose $\exp(A) = M$. If A has real eigenvalues λ_1, λ_2, then (wlog) $e^{\lambda_1} = 4, e^{\lambda_2} = -2$, but this is impossible for $\lambda_2 \in \mathbb{R}$.

If A has complex eigenvalues then they must be a conjugate pair: $\lambda, \overline{\lambda}$. But then e^{λ} and $e^{\overline{\lambda}}$ form a conjugate pair which is impossible since $e^{\lambda} = 4, e^{\overline{\lambda}} = -2$ (or vice versa).

Problem F15.7. Let A, B be two 4×5 matrices of rank 3. Find all possible values for the rank of $C = A^t B$. Specifically, you must find examples for any values possible and prove that no other values are possible.

Solution. The possible values for $\text{rank}(C)$ are 2 and 3. Putting

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

gives $\text{rank}(C) = 2$. Putting

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

gives $\text{rank}(C) = 3$.

Let $\{v_1, v_2, v_3\} \subset \mathbb{C}^4$ be a basis for $\text{im}(B)$ (since $\text{rank}(B) = 3$). Take $y \in \text{im}(C)$. Then there is $v \in \mathbb{C}^5$ such that $Cv = y$. Then $A^t(Bv) = y$. But $Bv \in \text{im}(B)$ so there are $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{C}$ such that $Bv = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$. Then

$$y = \alpha_1 A^t v_1 + \alpha_2 A^t v_2 + \alpha_3 A^t v_3.$$

Since $y \in \text{im}(C)$ was arbitrary, this shows that $\{A^t v_1, A^t v_2, A^t v_3\}$ is a spanning set for $\text{im}(C)$. Thus $\text{rank}(C) \leq 3$ since the rank is less than or equal to the number of elements in any spanning set.

Consider, since $\text{rank}(A) = 3$, we have $\text{rank}(A^t) = 3$. Then by the rank-nullity theorem, $\dim(\ker(A^t)) = 1$ since A^t acts on $\mathbb{C}^4$. Then since v_1, v_2, v_3 are linearly independent, there is at most one $i = 1, 2, 3$ such that $A^t v_i = 0$. Suppose there is one; wlog, $A^t v_3 = 0$. Then $\{v_3\}$ must form a basis for $\ker(A^t)$ since the kernel has dimension 1. Let $\beta_1, \beta_2 \in \mathbb{C}$ be such that

$$\beta_1 A^t v_1 + \beta_2 A^t v_2 = 0.$$

Then $A^t(\beta_1 v_1 + \beta_2 v_2) = 0$ so $\beta_1 v_1 + \beta_2 v_2 \in \ker(A^t)$. Then there is $\beta_3 \in \mathbb{C}$ such that

$$\beta_1 v_1 + \beta_2 v_2 = \beta_3 v_3.$$

But these vectors form a basis for $\text{im}(B)$ so this implies (in particular) that $\beta_1 = \beta_2 = 0$. Thus $A^t v_1$ and $A^t v_2$ are linearly independent. However, they are in $\text{im}(C)$ so this implies

$\text{rank}(C) \geq 2$ since the rank is greater than or equal to the number of elements in any linearly independent subset.

The case where there is no $i = 1, 2, 3$ such that $Av_i = 0$ is similar.

Problem F15.8. Find M^{-2} where

$$M = \begin{pmatrix} 2 & 3 & 2 & 1 \\ 3 & 6 & 4 & 2 \\ 4 & 8 & 6 & 3 \\ 2 & 4 & 3 & 1 \end{pmatrix}.$$

Solution. I'm not sure there is a "clever" way to do this. Just perform elementary row operations on M until you have the identity and then perform those same row operations to the identity to find M^{-1}. Doing this, we find

$$M^{-1} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix}.$$

Then squaring, we get

$$M^{-2} = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 5 & -4 & 1 & 0 \\ -4 & 7 & -3 & -1 \\ 2 & -6 & 4 & -1 \\ 0 & -2 & -1 & 5 \end{pmatrix}.$$

Problem F15.9. Let A be an $n \times n$ real matrix such that $A^t = -A$. Prove that $\det(A) \geq 0$.

Solution. If A is not invertible, then $\det(A) = 0$ so the claim is trivially satisfied.

Suppose A is invertible. Then A does not have zero as an eigenvalue. Let $\lambda \in \mathbb{C} - \{0\}$ be an eigenvalue of A with corresponsing eigenvector $0 \neq v \in \mathbb{C}^n$. Then

$$-\lambda(v, v) = (-\lambda v, v) = (-Av, v) = (A^t v, v) = (v, Av) = (v, \lambda v) = \overline{\lambda}(v, v).$$

Then since $(v, v) > 0$, we have $-\lambda = \overline{\lambda}$. But this yields $\text{Re}(\lambda) = 0$. Thus all eigenvalues of A are purely imaginary. Also the non-real eigenvalues of A come in conjugate pairs since A is real. Thus the eigenvalues of A can be listed: $i\mu_1, -i\mu_1, \ldots, i\mu_\ell, -i\mu_\ell$ for some $\ell \in \mathbb{N}$, $\mu_1, \ldots, \mu_\ell \in \mathbb{R} - \{0\}$. The determinant of A is the product of the eigenvalues so

$$\det(A) = \mu_1^2 \cdots \mu_\ell^2 \geq 0.$$

[Note: interestingly enough, this actually shows that if n is odd, then 0 must be an eigenvalue of A. Thus if A is an $n \times n$ skew-symmetric invertible matrix, then n is even.]

Problem F15.10. Let $F, G : \mathbb{R}^n \to \mathbb{R}^n$ be linear operators. Recall, we define the operator exponential by

$$\exp(F) = \sum_{k=0}^{\infty} \frac{1}{k!} F^k.$$

(a) Prove that when F and G commute, we have

$$\exp(F+G) = \exp(F)\exp(G).$$

(b) Find two non-commuting linear operators such that this equality fails.

Solution.

(a) If F, G commute, then the binomial theorem holds for F, G. That is,

$$(F+G)^k = \sum_{\ell=0}^{k} \binom{k}{\ell} F^\ell G^{k-\ell}, \quad k = 0, 1, \ldots$$

where by definition, $F^0 = I = G^0$ where I is the identity operator on $\mathbb{R}^n$. Recall the Cauchy product of two infinite series:

$$\left(\sum_{k=0}^{\infty} a_k\right)\left(\sum_{\ell=0}^{\infty} b_\ell\right) = \sum_{k=0}^{\infty}\sum_{\ell=0}^{k} a_\ell b_{k-\ell}$$

when all series converge absolutely. Using these we have

$$\begin{aligned}
\exp(F)\exp(G) &= \left(\sum_{k=0}^{\infty} \frac{1}{k!}F^k\right)\left(\sum_{\ell=0}^{\infty} \frac{1}{\ell!}G^\ell\right) \\
&= \sum_{k=0}^{\infty}\sum_{\ell=0}^{k} \left(\frac{1}{\ell!}F^\ell\right)\left(\frac{1}{(k-\ell)!}G^{k-\ell}\right) \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}\sum_{\ell=0}^{k} \binom{k}{\ell} F^\ell G^{k-\ell} \\
&= \sum_{k=0}^{\infty} \frac{1}{k!}(F+G)^k = \exp(F+G).
\end{aligned}$$

(b) Let

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

We see

$$\exp(A) = \exp(I)\exp\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) = \begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e & e \\ 0 & e \end{pmatrix}$$

and likewise

$$\exp(B) = \begin{pmatrix} e & 0 \\ e & e \end{pmatrix}$$

so

$$\exp(A)\exp(B) = \begin{pmatrix} e & e \\ 0 & e \end{pmatrix}\begin{pmatrix} e & 0 \\ e & e \end{pmatrix} = \begin{pmatrix} 2e^2 & e^2 \\ e^2 & e^2 \end{pmatrix}.$$

However,

$$\exp(A+B) = \exp\left(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\right) = \exp(2I)\exp(J),$$

where $J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ so $J^2 = I$. Then

$$\exp(J) = \sum_{k=0}^{\infty} \frac{1}{(2k)!} I + \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} J = \begin{pmatrix} \cosh(1) & \sinh(1) \\ \sinh(1) & \cosh(1) \end{pmatrix}$$

so

$$\exp(A+B) = \begin{pmatrix} e^2\cosh(1) & e^2\sinh(1) \\ e^2\sinh(1) & e^2\cosh(1) \end{pmatrix} \neq \exp(A)\exp(B).$$

Problem F15.11. Let $T : V \to V$ be a linear operator such that $T^6 = 0$ and $T^5 \neq 0$. Suppose $V \simeq \mathbb{R}^6$. Prove there is no linear operator $S : V \to V$ such that $S^2 = T$. Does the answer change if $V \simeq \mathbb{R}^{12}$?

Solution. Suppose that $V \simeq \mathbb{R}^6$ and that there is a linear operator $S : V \to V$ such that $T = S^2$. Then $0 = T^6 = S^{12}$ and $0 \neq T^5 = S^{10}$. Let $\lambda \in \mathbb{C}$ be an eigenvalue of S. Then λ^{12} is an eigenvalue of $S^{12} = 0$ so $\lambda^{12} = 0$ and so $\lambda = 0$. Thus all eigenvalues of S are zero. Then the characteristic polynomial of S is $p_S(x) = x^6$ since S acts on a 6 dimensional space. However, the Cayley-Hamilton theorem states that $p_S(S) = 0$ so $S^6 = 0 \implies S^{10} = 0$; a contradiction. Thus no such S exists.

Yes, the answer does change if V is 12 dimensional. Let $V = \mathbb{R}^{12}$ and T, S be the matrices

$$T = \begin{pmatrix} 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & & & \vdots \\ \vdots & \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix};$$

that is T has all zeroes except ones on the second superdiagonal and S is all zeroes except ones on the first superdiagonal. Then $T^5 \neq 0, T^6 = 0$ and $S^2 = T$.

Problem F15.12. Prove that the $n \times n$ matrix M is positive definite:

$$M = \begin{pmatrix} 2 & 1 & 1 & \cdots & 1 \\ 1 & 3 & 1 & \cdots & 1 \\ 1 & 1 & 4 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & n+1 \end{pmatrix}.$$

Solution. Let $x = (x_1, \dots, x_n) \in \mathbb{C}^n$ and let $M = (m_{ij})_{1 \le i,j \le n}$. Then

$$\begin{aligned} x^* M x &= \sum_{i,j=1}^{n} m_{ij} x_i \overline{x}_j \\ &= 2\,|x_1|^2 + x_1\overline{x}_2 + x_1\overline{x}_3 + \cdots + x_1\overline{x}_{n-1} + x_1\overline{x}_n \\ &\qquad + x_2\overline{x}_1 + 3\,|x_2|^2 + x_2\overline{x}_3 + \cdots + x_2\overline{x}_{n-1} + x_2\overline{x}_n \\ &\qquad\quad \vdots \\ &\qquad + x_n\overline{x}_1 + x_n\overline{x}_2 + x_n\overline{x}_3 + \cdots + x_n\overline{x}_{n-1} + (n+1)\,|x_n|^2 . \end{aligned}$$

From here, we see

$$|x_i|^2 + x_i\overline{x}_{i+1} + x_{i+1}\overline{x}_i + |x_{i+1}|^2 = (x_i + x_{i+1})(\overline{x}_i + \overline{x}_{i+1}) = (x_i + x_{i+1})\overline{(x_i + x_{i+1})} = |x_i + x_{i+1}|^2$$

for each $i = 1, \dots, n-1$, so

$$\begin{aligned} x^* M x = |x_1|^2 + |x_2|^2 + 2\,|x_3|^2 + \cdots + (n-2)\,|x_{n-1}|^2 + n\,|x_n|^2 \\ + |x_1 + x_2|^2 + |x_2 + x_3|^2 + \cdots + |x_{n-1} + x_n|^2 \end{aligned}$$

from which it is clear that $x^* M x \ge 0$ and so M is positive definite.

[Note: since M is real, it actually suffices to check $x \in \mathbb{R}^n$. If $v \in \mathbb{C}^n$, then $v = x + iy,\ x, y \in \mathbb{R}^n$ and $v^* M v = x^t M x + y^t M y$. Thus proving that $v^* M v \ge 0$ is reduced to showing that $x^t M x, y^t M y \ge 0$.]