Solutions of selected JPE problems in Linear Algebra

Dr Nikolai Chernov

Note to students preparing for JPE in Linear Algebra: it is *highly* recommended that you honestly attempt to work on past JPE problems on your own and read these solutions only as the last resort. Just reading the solutions, without trying to solve the problems, will not help you prepare for the exam.

JPE, September 2013, #4.

Let us denote $A^{(1)} = (a_{ij})$ for $1 \le i, j \le n$ and $A^{(2)} = (b_{ij})$ for $2 \le i, j \le n$ (note that the indices of b_{ij} start with 2, hence the rows and columns of $A^{(2)}$ will have the same numbers as the rows and columns of $A^{(1)}$). The strict column dominance of $A^{(1)}$ means that

$$|a_{jj}| > \sum_{i \ne j} |a_{ij}| \qquad \forall j = 1, \dots, n. \tag{1}$$

The first step of Gauss elimination produces

$$b_{ij} = a_{ij} - m_i a_{1j} \qquad \forall i, j = 2, \dots, n,$$

where

$$m_i = a_{i1}/a_{11} \qquad \forall i = 2, \dots, n$$

are multipliers. It follows from (1) with $j = 1$ that

$$\sum_{i=2}^{n} |m_i| < 1 \tag{2}$$

Now we have for each $j = 2, \dots, n$

$$\begin{aligned}
|b_{jj}| &= |a_{jj} - m_j a_{1j}| \ge |a_{jj}| - |m_j||a_{1j}| \\
&\ge |a_{jj}| - |a_{1j}| + \sum_{i \ge 2,\ i \ne j} |m_i||a_{1j}| \qquad \text{by (2)} \\
&> \sum_{i \ne j} |a_{ij}| - |a_{1j}| + \sum_{i \ge 2,\ i \ne j} |m_i||a_{1j}| \qquad \text{by (1)} \\
&= \sum_{i \ge 2,\ i \ne j} |a_{ij}| + \sum_{i \ge 2,\ i \ne j} |m_i||a_{1j}| \ge \sum_{i \ge 2,\ i \ne j} |b_{ij}|
\end{aligned}$$

where in the first and last lines we used the triangle inequality. Thus $|b_{jj}| > \sum_{i \ge 2,\ i \ne j} |b_{ij}|$ for each $j = 2, \dots, n$, hence $A^{(2)}$ is strictly column dominant.

JPE, May 2013, #3 (and May 2000, #2).

This problem can be solved by means of Linear Algebra, but there is also an elegant solution using SVD. First, by Linear Algebra, we know a general formula

$$\dim\big(\operatorname{Ker} T\big) + \dim\big(\operatorname{Range} T\big) = \dim V.$$

Since $\dim\big(\operatorname{Range} T\big) = \operatorname{rank} T < \dim(V)$ we have $\dim\big(\operatorname{Ker} T\big) > 0$. Let $L \subset V$ be a subspace such that $L \oplus \operatorname{Ker} T = V$. Then

$$\dim(L) = \dim(V) - \dim\big(\operatorname{Ker} T\big) = \dim\big(\operatorname{Range} T\big)$$

Let T_L denote the restriction of T to the subspace L. It is easy to see that T_L is a linear transformation taking L to $\operatorname{Range} T$. We claim that T_L is a bijection. Indeed, if it was not a bijection, there would be two vectors $x, y \in L$ such that $T_L(x) = T_L(y)$, i.e., $T(x) = T(y)$. But then $T(x - y) = 0$, hence $x - y \in \operatorname{Ker} T$, implying that $x - y = 0$. Now we construct bases α and β. Let α_L be a basis in L and α_0 be a basis in $\operatorname{Ker} T$. Then $\alpha = \alpha_L \cup \alpha_0$ is a basis in V. Now $\beta_L = T(\alpha_L) = T_L(\alpha_L)$ is a basis in $\operatorname{Range} T$. Let β be an arbitrary extension of β_L to a basis in W. Then one can verify easily that the bases α and β solve the problem.

Here is an elegant solution via SVD. Choose any basis in V and any basis in W, represent T by a matrix A in the chosen bases. By the SVD we have $A = U_1 D V_1^*$, where U_1 defines a basis β' in W of left singular vectors and V_1 defines a basis α in V of right singular vectors. In the bases α and β', the transformation T is represented by the diagonal matrix D with exactly $r = \operatorname{rank} T$ nonzero diagonal components $\sigma_1, \dots, \sigma_r$. Now by stretching the first r basis vectors of β' by the scalar factors $\sigma_1, \dots, \sigma_r$ we get a basis β solving the problem.

JPE, May 2013, #7.

(a) If $A = A^*V$, then

$$AA^* = A^*VA^* = A^*(AV^*)^* = A^*(AV^{-1})^* = A^*(A^*)^* = A^*A,$$

hence A is normal. If A is normal, then it is unitary equivalent to a diagonal matrix, i.e., $A = Q^*DQ$, where $D = \operatorname{diag}\{d_1, \dots, d_n\}$. Now $A^* = Q^*D^*Q$, where $D = \operatorname{diag}\{\bar{d}_1, \dots, \bar{d}_n\}$. Now $D = D^*U$, where $U = \operatorname{diag}\{u_1, \dots, u_n\}$ with

$$u_i = \begin{cases} d_i/\bar{d}_i & \text{if } \ d_i \neq 0 \\ -1 & \text{if } \ d_i = 0. \end{cases}$$

It is clear that U is a unitary matrix, hence

$$A = Q^*DQ = Q^*D^*UQ = Q^*D^*QQ^*UQ = A^*V$$

where $V = Q^*UQ$ is a unitary matrix, as required.

(b) The eigenvalues of A are the diagonal entries $d_1, \dots, d_n$. If they are purely imaginary, then in the above construction $u_i = -1$ for all i, hence $U = -I$. Therefore $V = -Q^*Q = -I$ and $A = -A^*$.

JPE, May 2012, #2.

(a) Let (λ, u) be an eigenpair for B. Then

$$u^*Bu = \langle Bu, u\rangle = \langle \lambda u, u\rangle = \lambda\langle u, u\rangle = \lambda\|u\|^2.$$

It is given to us that $u^*Bu = 0$, hence $\lambda\|u\|^2 = 0$. Since $\|u\| \neq 0$, we conclude that $\lambda = 0$. So all the eigenvalues of B are zero. Since B is Hermitian, the Spectral Theorem applies, and it says that B is unitary equivalent to a diagonal matrix, D, whose diagonal entries are the eigenvalues of B. Since all the eigenvalues of B are zero, we conclude that $D = 0$. That implies $B = Q^*DQ = Q^*0Q = 0$.

(b) We define $B = \frac{1}{2}(A+A^*)$ and $C = \frac{1}{2i}(A-A^*)$. Then we have $B^* = \frac{1}{2}(A^*+A) = B$ and $C^* = -\frac{1}{2i}(A^* - A) = C$, so both B and C are Hermitian. Lastly,

$$B + iC = \frac{1}{2}\left(A + A^*\right) + \frac{1}{2}\left(A - A^*\right) = A,$$

as required.

(c) By part (b), $A = B + iC$ where B and C are Hermitian matrices. Now

$$x^*Ax = x^*Bx + x^*(iC)x = \langle Bx, x\rangle + \langle iCx, x\rangle = \langle Bx, x\rangle + i\langle Cx, x\rangle$$

We know that for any Hermitian matrix P and any vector $x \in \mathbb{C}^n$ the inner product $\langle Px, x\rangle \in \mathbb{R}$ (is a real number). Thus in the above formula $\langle Bx, x\rangle$ is a real part of x^*Ax and $\langle Cx, x\rangle$ is its imaginary part. It is given to us that x^*Ax is real, hence its imaginary part is zero: $\langle Cx, x\rangle = 0$. By part (a) we conclude that $C = 0$ (the zero matrix). Hence $A = B$ is Hermitian.

JPE, September 2012, #6.

For any generalized eigenpair (λ, x) we have

$$\langle Ax, x\rangle = \langle \lambda Bx, x\rangle = \lambda\langle Bx, x\rangle$$

On the other hand,

$$\langle Ax, x\rangle = \langle x, Ax\rangle = \langle x, \lambda Bx\rangle = \bar{\lambda}\langle x, Bx\rangle = \bar{\lambda}\langle Bx, x\rangle$$

(where we used the given condition that A and B are Hermitian). For any $x \neq 0$ we have $\langle Bx, x\rangle > 0$, because B is positive definite, hence $\lambda = \bar{\lambda}$, which implies that λ is real.

Now let $B = GG^*$ be the Cholesky factorization for B. Note that B is invertible, hence so are G and G^*, Now we have

$$Ax = \lambda Bx \quad \Longrightarrow \quad Ax = \lambda GG^* x \quad \Longrightarrow \quad G^{-1}Ax = \lambda G^* x$$

Note that $I = (G^*)^{-1}G^* = (G^{-1})^* G^*$, therefore

$$G^{-1}A(G^{-1})^* G^* x = \lambda G^* x$$

Let us denote $G^*x = y$ and $G^{-1}A(G^{-1})^* = C$. Then we have

$$Cy = \lambda y$$

hence (λ, y) is an eigenpair for C. It is easy to see that C is Hermitian. Thus there is a basis (actually, an ONB) of eigenvectors $y_1, \dots, y_n$ of C. Now the vectors

$$x_1 = (G^*)^{-1}y_1, \dots, x_n = (G^*)^{-1}y_n$$

make a basis, too, because G^* is an invertible matrix. And the above vectors are generalized eigenvectors for the pair of matrices A, B.

JPE, May 2012, #6.

(a) The characteristic polynomial of A can be written as

$$p_A(x) = (x - \lambda_1) \cdots (x - \lambda_m)$$

where $\lambda_1, \dots, \lambda_m$ are the eigenvalues of A. Therefore

$$p_A(B) = (B - \lambda_1 I) \cdots (B - \lambda_m I)$$

Since $\lambda_1, \dots, \lambda_m$ are *not* the eigenvalues of B, all the above matrices $B - \lambda_1 I$, ..., $B - \lambda_m I$ are nonsingular. The product of nonsingular matrices is nonsingular, hence $p_A(B)$ is nonsingular.

(b) For any $k \ge 1$ we have

$$A^k X = A^{k-1}AX = A^{k-1}XB = A^{k-2}AXB = A^{k-2}XB^2 = \cdots = XB^k$$

therefore for any polynomial $P(x) = a_n x^n + \cdots + a_1 x + a_0$ we have

$$\begin{aligned} P(A)X &= a_n A^n X + \cdots + a_1 AX + a_0 IX \\ &= a_n XB^n + \cdots + a_1 XB + a_0 XI = XP(B). \end{aligned}$$

Now let $p_A(x)$ be the characteristic polynomial of A. By the Cayley-Hamilton theorem, $p_A(A) = 0$. On the other hand, $p_A(A)X = Xp_A(B)$, therefore $Xp_A(B) = 0 \times X = 0$ (the zero matrix). By part (a) the matrix $p_A(B)$ is nonsingular, hence

$$X = 0 \times [p_A(B)]^{-1} = 0.$$

(c) Denote the components of X by (x_{ij}). The equation $AX - XB = C$ can be written, componentwise, as m^2 equations with unknowns x_{ij}:

$$\sum_{i=1}^{m} a_{pi}x_{iq} - \sum_{j=1}^{m} x_{pj}b_{jq} = c_{pq}$$

where $p, q = 1, \ldots, m$. Note that these are *linear* equations in x_{ij}. And the number of equations (m^2) is the same as the number of unknowns. A system of linear equations has either a unique solution for every right hand side (for every matrix C) or the number of solutions is zero or infinity depending on the right hand side. If we show that the system cannot have more than one solution, it would follow that there is always exactly one solution.

Suppose that there are two matrices, X and X', that satisfy $AX - XB = C$ and $AX' - X'B = C$. Subtractive one equation from the other gives

$$A(X - X') - (X - X')B = C - C = 0 \quad \Longrightarrow \quad A(X' - X) = (X - X')B$$

and by part (b) we have $X - X' = 0$, hence $X = X'$.

JPE, September 2011, #2.

(a) Since A is upper triangular, its eigenvalues are its diagonal entries, i.e., they are all equal to 1. Hence $\det A = 1$. This implies that A is nonsingular, i.e., its rank is n.

(b) A^{-1} is an upper triangular matrix with 1 on the main diagonal, -2 on the first superdiagonal, 4 on the second superdiagonal, …, $(-2)^k$ on the kth superdiagonal.

(c) Recall that σ_1^2 is the largest eigenvalue of A^*A. By direct computation, A^*A is a symmetric matrix with $1, 5, 5, \ldots, 5$ on the main diagonal and $2, 2, \ldots, 2$ on the first subdiagonal and first superdiagonal (zeros elsewhere). By Gershgorin theorem, all its eigenvalues lie in the union of two disks:

$$|z - 1| \le 2 \qquad \text{and} \qquad |z - 5| \le 4.$$

In fact, all the eigenvalues must be real and non-negative, hence they are confined to the union of two intervals: $[0, 3]$ and $[1, 9]$, which is one interval $[0, 9]$. Hence the largest eigenvalue is ≤ 9, as desired.

JPE, September 2011, #3.

Let $A = QR$ be a QR decomposition of A. Then $|\det A| = |\det Q|\,|\det R|$, and since Q is a unitary matrix, $|\det Q| = 1$, and we get $|\det A| = |\det R| = \prod_{j=1}^{n} |r_{jj}|$.

On the other hand, $a_j = Qr_j$ for all $j = 1, \ldots, n$, where a_j and r_j denote the jth columns of A and R, respectively. Now $\|a_j\|_2 = \|r_j\|_2$ because Q is a unitary matrix. Therefore we really need to prove that $\prod_{j=1}^{n} |r_{jj}| \le \prod_{j=1}^{n} \|r_j\|_2$. In fact, we can easily prove more than that: for each j we have $|r_{jj}| \le \|r_j\|_2$ because r_{jj} is just one components of the column r_j of the matrix R.

JPE, September 2011, #5.

(a) Similar matrices have the same determinant, and obviously $\det B = 0$, therefore our first task is to find x such that $\det A = 0$. We easily compute $\det A = -x^2$, hence we get equation $-x^2 = 0$ with the only solution $x = 0$.

Next, the matrix A with $x = 0$ has three eigenvalues: 0, 1, and 2. Since they are distinct, A is diagonalizable, and it is equivalent to a diagonal matrix $\operatorname{diag}\{0, 1, 2\}$, which is exactly B. Thus $A \sim B$ if and only if $x = 0$.

(b) Note that A is real symmetric, i.e., Hermitian. By the Spectral Theorem, A is unitary equivalent to a diagonal matrix whose diagonal components are the eigenvalues of A. In the case $x = 0$ those eigenvalues are 0, 1, and 2, hence A is unitary equivalent to B.

JPE, September 2011, #6.

Let us decompose the matrix A as follows:

$$A = \begin{bmatrix} a_{11} & a_{21} & \cdots & 0 & 0 \\ a_{21} & a_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & a_{n,n-1} \\ 0 & 0 & \cdots & a_{n,n-1} & a_{nn} \end{bmatrix} = B + C$$

with

$$B = \begin{bmatrix} a_{11} & a_{21} & \cdots & 0 & 0 \\ a_{21} & a_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & 0 \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n,n-1} \\ 0 & 0 & \cdots & a_{n,n-1} & 0 \end{bmatrix}$$

Note that C only has two non-zero components, both in its trailing 2×2 block.

The matrix B is block-diagonal, with one big block of size $(n-1) \times (n-1)$ and one tiny block of size 1×1. By a general rule, its eigenvalues are those of its blocks. The trailing 1×1 block $[a_{nn}]$ obviously has eigenvalue a_{nn}.

On the other hand, $\|C\|_2 = |a_{n,n-1}|$ as one can verify directly (we omit that verification). Also note that both B and C are real symmetric, hence Hermitian. By a general theorem in the course, the eigenvalues of A and those of B can be paired so that the difference between the corresponding eigenvalues of A and B is $\le \|C\|_2$. Thus there is an eigenvalue λ of A such that

$$|\lambda - a_{nn}| \le \|C\|_2 = |a_{n,n-1}|$$

Note: it is tempting to use the Gershgorin theorem, but it will not work because we do not have enough control over the locations and sizes of Gershgorin disks, other than the last one. Therefore the Gershgorin disks can overlap and we will not be able to prove that the last one contains at least one eigenvalue of A.

JPE, May 2011, #1.

We will use the following fact (proven in the course[1]): if $\langle Sz, z\rangle = 0$ for every $z \in \mathbb{C}^n$, then $S = 0$. So in order to show that $T^* = -T$, or $T + T^* = 0$, it is enough to check that $\langle (T + T^*)z, z\rangle = 0$ for every $z \in \mathbb{C}^n$. Let $z = x + iy$, where $x, y \in \mathbb{R}^n$ and $i = \sqrt{-1}$. Now we have

$$\begin{aligned}\langle (T + T^*)z, z\rangle &= \langle (T + T^*)(x + iy), x + iy\rangle \\ &= \langle Tx, x\rangle + \langle T^*x, x\rangle + i\langle Ty, x\rangle + i\langle T^*y, x\rangle \\ &\quad - i\langle Tx, y\rangle - i\langle T^*x, y\rangle + \langle Ty, y\rangle + \langle T^*y, y\rangle\end{aligned}$$

We are given that $\langle Tx, x\rangle = 0$ for every $x \in \mathbb{R}^n$, which also implies that $\langle T^*x, x\rangle = \langle x, Tx\rangle = \langle Tx, x\rangle = 0$. So the first two and the last two terms in the above expression vanish. The middle four terms cancel one another because

$$\langle T^*x, y\rangle = \langle x, Ty\rangle = \langle Ty, x\rangle \quad \text{and} \quad \langle T^*y, x\rangle = \langle y, Tx\rangle = \langle Tx, y\rangle$$

Thus indeed $\langle (T + T^*)z, z\rangle = 0$ for every $z \in \mathbb{C}^n$, hence $T^* = -T$.

JPE, May 2011, #5.

Let us begin with $n = 2$. A counterclockwise rotation by angle θ is represented by matrix

$$G_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

One can find, by examining the rotation of the xy plane by angle θ in geometric terms, that it is a composition of two reflectors: one across the horizontal line $L_1 = \operatorname{span}\{e_1\}$ and the other across the line bisecting the angle θ, i.e., the line

$$L_2 = \operatorname{span}\bigl\{[\cos\theta/2,\ \sin\theta/2]^T\bigr\}.$$

We will prove this algebraically. The first reflector takes $e_1 \mapsto e_1$ and $e_2 \mapsto -e_2$, so it is defined by matrix

$$P_1 = \begin{bmatrix} e_1 & -e_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The matrix of the second reflector can be found by the Householder formula

$$P_2 = 1 - 2xx^T$$

[1] This fact also follows from the solution of JPE, May 2012, #2.

where x is a unit vector orthogonal to the line L_2. Taking $x = [-\sin\theta/2, \cos\theta/2]^T$ gives

$$P_2 = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

Now one can verify directly that

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

hence $G_\theta = P_2 P_1$. Of course, the reflectors P_1 and P_2 are not unique, one can choose them in many ways.

Now for arbitrary $n > 2$ every Givens rotator $G_{i,j,\theta}$ acts as a rotator in the 2D subspace V spanned by two canonical basis vectors, e_i and e_j. In the orthogonal complement to V, it is an identity, i.e., $G_{i,j,\theta}(e_k) = e_k$ for all $k \notin \{i, j\}$. We first construct two reflectors, P_1 and P_2, in the space V, as above, and then extend them to the whole space by requiring that $P_1(e_k) = e_k$ and $P_2(e_k) = e_k$ for all $k \notin \{i, j\}$. This makes the identity $G_{i,j,\theta} = P_2 P_1$ valid not only in V, but in the whole space.

Lastly, a reflector cannot be a product of two (or any other number) of rotators. One can easily check that reflectors have determinant -1 and rotators have determinant $+1$. Thus we cannot have the identity $-1 = 1 \cdot 1 \cdots 1$.

JPE, May 2010, #2.

(a) Since the space V is finite-dimensional, we can represent S and T by matrices, which we will denote by the same letters, S and T.

Now $\det(ST) = \det S \cdot \det T = \det T \cdot \det S = \det(TS)$. Thus either both $\det(ST)$ and $\det(TS)$ are zero or both are not. This implies that zero either is an eigenvalue for both ST and TS or is not an eigenvalue for either. It remains to consider non-zero eigenvalues.

If $\lambda \neq 0$ is an eigenvalue of ST, then $STx = \lambda x$ for some $x \neq 0$. Premultiplying by T gives $TSTx = \lambda Tx$. Denote $y = Tx$; then we can write $TSy = \lambda y$. If $y \neq 0$, we conclude that λ is an eigenvalue of TS, as required. If $y = 0$, then $Tx = 0$, hence $STx = S0 = 0$, which implies $\lambda x = 0$, therefore $\lambda = 0$, which contradicts our assumption that $\lambda \neq 0$. So the case $y = 0$ is impossible.

(b) If the matrix has distinct eigenvalues, it is similar to a diagonal matrix, i.e., $T = X^{-1}DX$ for a diagonal matrix D. Since S has the same eigenvectors, we have $S = P^{-1}D'P$, where D' is another diagonal matrix (whose diagonal entries are the eigenvalues of S. Now

$$ST = P^{-1}D'PP^{-1}DP = P^{-1}D'DP = P^{-1}DD'P = P^{-1}DPP^{-1}D'P = TS$$

where we used the simple fact that diagonal matrices commute.

JPE, May 2010, #6.

(a) Since $f(A) = 0$ and f is a polynomial with two roots, $x = 2$ and $x = -5$, we conclude that the matrix A has just two distinct eigenvalues: $\lambda = 2$ and $\lambda = -5$. Let their algebraic multiplicities be p and q, respectively. Then $\operatorname{tr} A = 2p - 5q$. It is given to us that $\operatorname{tr} A = 0$, therefore $2p = 5q$. Hence p is a multiple of 5 and q is a multiple of 2, i.e., $p = 5k$ and $q = 2k$ for some $k \geq 1$. This implies $n = p + q = 7k$.

(b) Since A is symmetric, it is diagonalizable, hence all Jordan blocks have minimal size 1×1. Therefore the minimal polynomial is $m(x) = (x-2)(x+5)$.

(c) The characteristic polynomial is $(x-2)^p(x+5)^q = (x-2)^{5k}(x+5)^{2k}$.

(d) The eigenvalues of A^2 are $2^2 = 4$ with multiplicity $p = 5k$ and $(-5)^2 = 25$ with multiplicity $q = 2k$. Hence $\operatorname{tr} A^2 = 4 \cdot 5k + 25 \cdot 2k = 70k$.

JPE, May 2010, #7.

By a general formula we learned in the course,

$$\dim\bigl(\operatorname{Ker} T\bigr) + \dim\bigl(\operatorname{Range} T\bigr) = \dim V.$$

Also we know that

$$\dim\bigl(\operatorname{Range} T\bigr) = \operatorname{rank}(T) = \operatorname{rank}(T^*) = \dim\bigl(\operatorname{Range} T^*\bigr)$$

Thus the subspaces $\operatorname{Ker} T$ and $\operatorname{Range} T^*$ have "complimentary" dimensions: their dimensions add up to $\dim V$. Now it is enough to show that $\operatorname{Ker} T \cap \operatorname{Range} T^* = \{0\}$; this would imply

$$V = \operatorname{Ker} T \oplus \operatorname{Range} T^*$$

hence the union of bases of these two subspaces would be a basis in V.

Suppose, by way of contradiction, that there is $x \neq 0$ such that $x \in \operatorname{Ker} T \cap \operatorname{Range} T^*$. Then there is a $y \in W$ such that $T^*y = x$. Now we have

$$\langle Tx, y\rangle = \langle x, T^*y\rangle = \langle x, x\rangle > 0.$$

On the other hand, $Tx = 0$, hence $\langle Tx, y\rangle = 0$, a contradiction.

JPE, September 2009, #4 (and Sept. 2006, #1).

(a) The characteristic polynomial of A is $(\lambda - 1)^4$, so it has one eigenvalue $\lambda = 1$ of algebraic multiplicity four. If it has one Jordan block, then $A - \lambda I = A - I$ must have rank three. A direct inspection shows that this happens whenever $(a+b)(c-d) \neq 0$, i.e., the conditions are $a \neq -b$ and $c \neq d$.

(b) Routine calculations.

JPE, September 2008, #6.

This problem can be solved by using the standard analysis of rank one matrices (see JPE, May 2007, #2, and May 2005, #1, below). But we also give an independent solution:

First, we show that A is a projector:

$$\begin{aligned} A^2 &= \big(I - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T\big)\big(I - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T\big) \\ &= I - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T + \tfrac{1}{n^2}(\mathbf{1}\mathbf{1}^T)(\mathbf{1}\mathbf{1}^T) \end{aligned}$$

We note that

$$(\mathbf{1}\mathbf{1}^T)(\mathbf{1}\mathbf{1}^T) = \mathbf{1}\,\underbrace{(\mathbf{1}^T\mathbf{1})}_{=n}\,\mathbf{1}^T = n\,\mathbf{1}\mathbf{1}^T$$

because $\mathbf{1}^T\mathbf{1} = \langle \mathbf{1}, \mathbf{1}\rangle = 1 + \cdots + 1 = n$ is the inner product of two vectors. Then

$$A^2 = I - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T + \tfrac{1}{n}\mathbf{1}\mathbf{1}^T = I - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T = A$$

and because $A^2 = A$, we see that A is a projector.

Second, we verify that A is an orthogonal projector. A projector A is orthogonal iff A is Hermitian, $A^* = A$. Since A is real, we must verify that $A^T = A$. Here goes:

$$A^T = I^T - \tfrac{1}{n}\big(\mathbf{1}\mathbf{1}^T\big)^T = I - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T = A$$

hence A is an orthogonal projector.

Next we identify its range. For a projector, the range consists of vectors x such that $Ax = x$. This means

$$x = Ax = \big(I - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T\big)x = x - \tfrac{1}{n}\mathbf{1}\mathbf{1}^T x = x - \tfrac{1}{n}(\mathbf{1}^T x)\,\mathbf{1}$$

where $\mathbf{1}^T x = \langle x, \mathbf{1}\rangle = x_1 + \cdots + x_n$ is again a scalar product of two vectors. Thus we have

$$x \in \text{Range}\,A \quad\Longleftrightarrow\quad \tfrac{1}{n}\langle x, \mathbf{1}\rangle\,\mathbf{1} = 0 \quad\Longleftrightarrow\quad \langle x, \mathbf{1}\rangle = 0$$

Thus the range is the orthogonal complement to the vector $\mathbf{1}$. It is obviously an $(n-1)$-dimensional vector space.

The kernel is the orthogonal complement to the range, hence $\text{Ker}\,A = \text{span}\{\mathbf{1}\}$.

Next we find the singular values of A. Since A is Hermitian, its singular values are the absolute values of its eigenvalues. So we are looking for the eigenvalues of A.

For each nonzero vector $x \in \text{Range}\,A$ we have $Ax = x$, hence x is an eigenvector with eigenvalue $\lambda = 1$. Thus $\lambda = 1$ is an eigenvalue of geometric multiplicity $n-1$ (the dimensionality of the range of A).

For each nonzero vector $x \in \text{Ker}\,A$ we have $Ax = 0$, hence x is an eigenvector with eigenvalue $\lambda = 0$. Thus $\lambda = 0$ is an eigenvalue of geometric multiplicity 1 (the dimensionality of the kernel of A).

So the eigenvalues of A are $1, 1, \ldots, 1, 0$. And so are its singular values.

JPE, May 2008, #3.

It was shown in class that for Hermitian matrices the maximum value of the Rayleigh quotient $r(x)$ is $\lambda_{\max}$ and its minimum value is $\lambda_{\min}$. Now since $r(x)$ is a continuous function on its domain $\mathbb{C}^n \setminus \{0\}$ and the domain is connected, it follows that the range is connected, too, hence the range is the interval $[\lambda_{\min}, \lambda_{\max}]$.

JPE, September 2007, #6.

(a) The eigenvalues of M are *distinct*, because $b \neq 0$. Hence M is similar to the diagonal matrix $D = \begin{bmatrix} a+ib & 0 \\ 0 & a-ib \end{bmatrix}$. By direct calculation, the matrix $N = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ has the same eigenvalues, $a \pm ib$. Thus N is also similar to D. Hence M and N are similar, i.e., $Q^{-1}MQ = N$ for some nonsingular matrix Q.

The problem also specifies that Q must be a real matrix, i.e., $Q \in \mathbb{R}^{2\times 2}$. The existence of a *real* matrix Q that establishes the similarity between two real matrices, M and N, follows from a general, though little known, fact. Here we prove it, for the sake of completeness:

Fact: If $A, B \in \mathbb{R}^{n\times n}$ are two similar real matrices, then there exists a real nonsingular matrix $Q \in \mathbb{R}^{n\times n}$ such that $Q^{-1}AQ = B$.

Proof: By similarity, there exists a complex matrix $P \in \mathbb{C}^{n\times n}$ such that $P^{-1}AP = B$. We can rewrite this equation as $AP = PB$. Let $P = U + iV$, where $U, V \in \mathbb{R}^{n\times n}$ are real matrices (the "real" and "imaginary" parts of P). Then we have

$$A(U+iV) = (U+iV)B \quad \Rightarrow \quad AU + iAV = UB + iVB \quad \Rightarrow \quad AU = UB \ \ \& \ \ AV = VB$$

(because A and B are *real* matrices). Now if at least one of U and V is nonsingular, we are done. If both are singular, we need to sweat a little more. The above equations imply

$$A(U+zV) = (U+zV)B \qquad \forall z \in \mathbb{C}$$

It is enough for us to show that $\exists z \in \mathbb{R}$ such that $U + zV$ is invertible. Now $\det(U + zV)$ is a polynomial in z of degree n, and its coefficients are real numbers (they are algebraic expressions involving only the entries of the matrices U and V). We know that $\det(U + iV) \neq 0$, because $P = U + iV$ is invertible, hence our polynomial is not identically zero (i.e., not all of its coefficients are zero). This implies that $\exists z \in \mathbb{R}$ such that $\det(U + zV) \neq 0$. Therefore $U + zV$ is invertible, and we set $Q = U + zV$. $\square$

(b) Real Schur decomposition that we learned in class gives this result.

Next three problems involve rank one matrices, so we put them together:

JPE, May 2007, #2.

(a) Suppose $\operatorname{rank} A = 1$; then the range of A is a one-dimensional space. Pick a non-zero vector $x \in \operatorname{Range} A$. Then $\operatorname{Range} A = \operatorname{span}(x)$. By a general formula,

$$\dim\bigl(\operatorname{Ker} A\bigr) = n - \dim\bigl(\operatorname{Range} A\bigr) = n - 1,$$

so $\operatorname{Ker} A$ is a hyperplane. Pick a unit vector u orthogonal to $\operatorname{Ker} A$, then $\operatorname{Ker} A = u^{\perp}$. Since $u \notin \operatorname{Ker} A$, we conclude that $Au \neq 0$, and since $Au \in \operatorname{Range} A$ we conclude that $Au = cx$ for some scalar c. Let $y = \bar{c}u$. Now we will verify that $A = xy^*$. Until then we denote $B = xy^*$.

Indeed, for any $z \in \operatorname{Ker} A$ we have $\langle z, u\rangle = 0$, hence

$$Bz = xy^*z = \langle z, y\rangle x = c\langle z, u\rangle x = 0 \cdot x = 0$$

hence $Bz = Az$. Also,

$$Bu = xy^*u = \langle u, y\rangle x = c\langle u, u\rangle x = cx$$

hence $Bu = Au$. Therefore $B = A$, as claimed.

Note that x and y defining $A = xy^*$ are not unique. We can replace x with sx and y with $\bar{s}^{-1}y$ for any non-zero scalar $s \in \mathbb{C}$.

(b) We have

$$A^2 = xy^*xy^* = x(y^*x)y^* = \langle x, y\rangle xy^* = \langle x, y\rangle A$$

In order to find a Jordan canonical form for $M = I + A$ we need to describe the action of A. First we note that for any $z \in \operatorname{Ker} A$

$$Mz = (I + A)z = z + Az = z$$

hence z is an eigenvector of M with eigenvalue 1. Thus $\lambda = 1$ is an eigenvalue of M with geometric multiplicity at least $n - 1$. The further analysis involves two cases:

(i) (Main case) $x \notin \operatorname{Ker} A$. Then $\langle x, y\rangle \neq 0$. Note that

$$Mx = (I + A)x = x + xy^*x = x + \langle x, y\rangle x = (1 + \langle x, y\rangle)x$$

hence x is an eigenvector of M with eigenvalue $1+\langle x, y\rangle$ (which is *different* from 1). We see that M has two distinct eigenvalues: $\lambda_1 = 1$ with algebraic and geometric multiplicity $n - 1$ and $\lambda_2 = 1 + \langle x, y\rangle$ with algebraic and geometric multiplicity 1. Therefore, M is diagonalizable, and its Jordan canonical form is a diagonal matrix with one entry $1 + \langle x, y\rangle$ and $n - 1$ entries equal to 1. Its minimal polynomial is

$$m(\lambda) = (\lambda - 1)(\lambda - 1 - \langle x, y\rangle)$$

(ii) (Special case) $x \in \operatorname{Ker} A$. Then $\langle x, y\rangle = 0$. Note that $M - I = xy^*$, hence

$$(M - I)y = xy^*y = \langle y, y\rangle x \neq 0$$

and

$$(M - I)^2y = \langle y, y\rangle xy^*x = \langle y, y\rangle\langle x, y\rangle x = 0.$$

Therefore y is a generalized eigenvector for M corresponding to $\lambda = 1$, but not an eigenvector. Thus M has a unique eigenvalue $\lambda = 1$ of algebraic multiplicity n and geometric multiplicity $n-1$. Its Jordan canonical form has all 1's on the main diagonal and a single 1 above the main diagonal. There is one Jordan block of size 2×2 and $n-1$ Jordan blocks of size 1×1. The minimal polynomial of M is

$$m(\lambda) = (\lambda - 1)^2$$

JPE, May 2007, #8.

(b) After k steps of the Gaussian elimination we obtain a *partial* LU decomposition:

$$A = L^{(k+1)} A^{(k+1)},$$

where $L^{(k+1)}$ is of the form

$$L^{(k+1)} = \begin{bmatrix} L_{11}^{(k+1)} & 0 \\ L_{12}^{(k+1)} & I \end{bmatrix}$$

where $L_{11}^{(k+1)} \in \mathbb{C}^{k \times k}$ is a unit lower triangular matrix, as is expected. Note that the L factor in the LU decomposition, below its main diagonal, is filled with multipliers. Since those have only been constructed for the first k columns, the last $n-k$ columns of $L^{(k+1)}$, below the main diagonal, are filled with zeros. In other words, the bottom right $(n-k) \times (n-k)$ block of $L^{(k+1)}$ is the identity matrix, I.

Now multiplying blocks of the matrix $L^{(k+1)}$ and those of $A^{(k+1)}$ gives us

$$\begin{aligned} A_{11} &= L_{11}^{(k+1)} A_{11}^{(k+1)} \\ A_{12} &= L_{11}^{(k+1)} A_{12}^{(k+1)} \\ A_{21} &= L_{21}^{(k+1)} A_{11}^{(k+1)} \\ A_{22} &= L_{21}^{(k+1)} A_{12}^{(k+1)} + A_{22}^{(k+1)} \end{aligned}$$

From these equations one easily gets

$$\begin{aligned} A_{22}^{(k+1)} &= A_{22} - L_{21}^{(k+1)} A_{12}^{(k+1)} \\ A_{22}^{(k+1)} &= A_{22} - A_{21} [A_{11}^{(k+1)}]^{-1} [L_{11}^{(k+1)}]^{-1} A_{12} \\ A_{22}^{(k+1)} &= A_{22} - A_{21} A_{11}^{-1} A_{12} \end{aligned}$$

Note: the Gaussian elimination operations cannot alter the determinant of every principal minor, therefore $\det A_{11} = \det A_{11}^{(k+1)}$. It is given to us that A_{11} is invertible, hence $\det A_{11} \neq 0$, so the matrix $A_{11}^{(k+1)}$ is invertible, too.

Note: it is easy to guess the right formula for $A_{22}^{(k+1)}$ as follows: assume that $k = 1$ and $n = 2$; then by elementary calculation

$$a_{22}^{(2)} = a_{22} - a_{21}a_{12}/a_{11}$$

Replacing the individual components a_{ij} with the blocks A_{ij} and arranging the last term so that the multiplication is possible for any $n > k$ we get the right formula.

(b) From part (a), using $k = 1$, we have

$$A_{22}^{(2)} = A_{22} - \tfrac{1}{a_{11}} A_{21}A_{12}$$

Since A is Hermitian, we have $A_{12} = A_{21}^*$, $A_{22}^* = A_{22}$, and $a_{11} \in \mathbb{R}$. Therefore

$$[A_{22}^{(2)}]^* = A_{22}^* - \tfrac{1}{\bar{a}_{11}} A_{12}^*A_{21}^* = A_{22} - \tfrac{1}{a_{11}} A_{21}A_{12},$$

which proves that $A_{22}^{(2)}$ is Hermitian, too.

Now for every $y \in \mathbb{C}^{n-1}$ we have

$$\begin{aligned}\langle A_{22}^{(2)}y, y\rangle = y^*A_{22}^{(2)}y &= y^*A_{22}y - \tfrac{1}{a_{11}}\, y^*A_{21}A_{21}^*y \\ &= y^*A_{22}y - \tfrac{1}{a_{11}}\, |\langle A_{21}, y\rangle|^2\end{aligned}$$

We need to prove that $\langle A_{22}^{(2)}y, y\rangle > 0$ for every $y \neq 0$.

Let $x = \begin{bmatrix} c \\ y \end{bmatrix}$, where $c \in \mathbb{C}$ and y is as above. Then multiplying the blocks of x and those of A gives

$$\begin{aligned}\langle Ax, x\rangle &= a_{11}|c|^2 + cy^*A_{21} + \bar{c}A_{21}^*y + y^*A_{22}y \\ &= a_{11}|c|^2 + c\langle A_{21}, y\rangle + \overline{c\langle A_{21}, y\rangle} + y^*A_{22}y\end{aligned}$$

We now choose $c = -\frac{1}{a_{11}}\, \overline{\langle A_{21}, y\rangle}$ and obtain

$$\langle Ax, x\rangle = -\tfrac{1}{a_{11}}\, |\langle A_{21}, y\rangle|^2 + y^*A_{22}y = \langle A_{22}^{(2)}y, y\rangle.$$

Since A is positive definite, and $y \neq 0$ implies $x \neq 0$, we have

$$\langle A_{22}^{(2)}y, y\rangle = \langle Ax, x\rangle > 0.$$

JPE, May 2005, #1.

(a) The matrix A admits an eigenvalue decomposition only in the Main case above, i.e., under the condition $\langle x, y\rangle \neq 0$.

(b) The matrix A admits a unitary diagonalization if and only if there exists an ONB of eigenvectors. This happens when the eigenspaces corresponding to $\lambda_1 = 1$ and $\lambda_2 = 1 + \langle x, y\rangle$ are orthogonal, i.e., under the condition $x \perp \operatorname{Ker}(xy^*)$. This condition is equivalent to x and y being collinear, i.e., $x = cy$ for some scalar c.

JPE, September 2004, #7.

The result follows from the analysis above.

JPE, September 2006, #4 (and Sept. 1999, #6)

Let us denote the given $(n+1) \times (n+1)$ matrix by B. If $\det B > 0$, then by the Sylvester theorem B is positive definite. If $\det B = 0$, then B has an eigenvalue zero. We will show that neither is possible.

Let us take an arbitrary vector $v \in \mathbb{R}^{n+1}$ and represent it as $v = \begin{bmatrix} y \\ z \end{bmatrix}$, where $y \in \mathbb{R}^n$ and $z \in \mathbb{R}$. Then

$$Bv = \begin{bmatrix} A & x \\ x^T & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} Ay + zx \\ x^T y \end{bmatrix}$$

and

$$\langle Bv, v\rangle = \begin{bmatrix} y^T & z \end{bmatrix} \begin{bmatrix} Ay + zx \\ x^T y \end{bmatrix} = y^T Ay + zy^T x + zx^T y = \langle Ay, y\rangle + 2z\langle x, y\rangle$$

Choosing $y = x$ and $z = -\langle Ax, x\rangle / \langle x, x\rangle$ gives

$$\langle Bv, v\rangle = -\langle Ax, x\rangle < 0$$

which shows that B is *not* positive definite.

Now if B had an eigenvalue zero, there would be a nonzero vector v such that $Bv = 0$. This implies $Ay = -zx$ and $\langle x, y\rangle = 0$, therefore

$$\langle Ay, y\rangle = \langle -zx, y\rangle = -z\langle x, y\rangle = 0$$

Since A is positive definite, the above relation can only happen if $y = 0$. In that case $zx = -Ay = -A0 = 0$, so $z = 0$ as well, thus $v = 0$.

JPE, May 2006, #7.

Let $zI - A = UDV^*$ be an SVD for the matrix $zI - A$. Then $(zI - A)^{-1} = VD^{-1}U^*$ is an SVD for $(zI - A)^{-1}$. Thus the singular values of $(zI - A)^{-1}$ are the reciprocals of those of $zI - A$. In particular, σ_n^{-1} is the largest singular value of $(zI - A)^{-1}$. This implies $\|(zI - A)^{-1}\| = \sigma_n^{-1}$. Now the equivalence of (c) and (d) is obvious.

Now for any unit vector u we have

$$\|(A - zI)u\|_2 = \|(zI - A)u\|_2 = \|UDV^*u\|_2 = \|Dv\|_2$$

where $v = V^*u$ is a unit vector, too. Clearly,

$$\min_{\|v\|_2=1} \|Dv\|_2 = \sigma_n.$$

Thus the condition (b) simply says that $\sigma_n \le \varepsilon$, hence it is equivalent to (c).

Lastly, if z is an eigenvalue of $A+B$, then $S = A+B-zI$ is a singular matrix. We can rewrite it as $B = S-(zI-A)$, and the condition $\|B\|_2 \le \varepsilon$ becomes $\|S-(zI-A)\|_2 \le \varepsilon$. Now the condition (a) simply says that the "distance" (in the 2-norm) from $zI - A$ to the nearest singular matrix is $\le \varepsilon$. We know from the course that this distance is equal to the smallest singular value, σ_n. Thus (a) is equivalent to (c).

JPE, September 2005, #8 (and Sept. 2009, #8b, and May 2000, #7).

Taking the limit $k \to \infty$ gives $AQ_\infty = Q_\infty R_\infty$. Since R_k is upper triangular for every k, so is its limit R_∞. Since Q_k is unitary for every k, we have $Q_k^* Q_k = I$. Taking the limit $k \to \infty$ gives $Q_\infty^* Q_\infty = I$, hence Q_∞ is unitary, too. Now $A = Q_\infty R_\infty Q_\infty^*$, hence A is unitary equivalent to the upper triangular matrix R_∞. Thus the eigenvalues of A are the diagonal components of R_∞.

JPE, September 2004, #1.

(a) If $\lambda \neq 0$ is an eigenvalue of TS, then $TSx = \lambda x$ for some $x \neq 0$. Premultiplying by S gives $STSx = \lambda Sx$. Denote $y = Sx$; then we can write $STy = \lambda y$. If $y \neq 0$, we conclude that λ is an eigenvalue of ST, as required. If $y = 0$, then $Sx = 0$, hence $TSx = T0 = 0$, which implies $\lambda x = 0$, therefore $\lambda = 0$, which contradicts our assumption that $\lambda \neq 0$.

(b) Indeed, we can take for example $T = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $S = \begin{bmatrix} 1 & 0 \end{bmatrix}$. Then $TS = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ has two eigenvalues: 1 and 0. On the other hand, $ST = [1]$ has one eigenvalue: 1.

JPE, September 2002, #5 (and Jan. 1989, #7, Jan. 1988, #5).

We have

$$\begin{aligned} x^{(k+1)} &= M^{-1}(b + Nx^{(k)}) = M^{-1}(Ax + Nx^{(k)}) \\ &= M^{-1}(M-N)x + M^{-1}Nx^{(k)} = x + M^{-1}N(x^{(k)} - x) \end{aligned}$$

therefore

$$e^{(k+1)} = M^{-1}Ne^{(k)}$$

hence $G = M^{-1}N$. Similarly,

$$\begin{aligned} r^{(k+1)} &= b - Ax^{(k+1)} = b - (M-N)M^{-1}(b + Nx^{(k)}) \\ &= NM^{-1}b - Nx^{(k)} + NM^{-1}Nx^{(k)} \\ &= NM^{-1}b - NM^{-1}(Mx^{(k)} - Nx^{(k)}) \\ &= NM^{-1}(b - Ax^{(k)}) = NM^{-1}r^{(k)} \end{aligned}$$

hence $H = NM^{-1}$. Note that

$$H = NM^{-1} = MM^{-1}NM^{1} = MGM^{-1},$$

hence H and G are similar. As a result, they have the same spectrum (and the same spectral radius).

JPE, May 2001, #3.

Since A has rank n, we have $m \geq n$. We have proved in class that for a full rank matrix A the smaller of A^*A and AA^* is nonsingular. Since $m \geq n$, the smaller is A^*A.

Next, we easily see that

$$\begin{bmatrix} I_{m\times m} & A \\ A^* & 0_{n\times n} \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} r + Ax \\ A^*r \end{bmatrix}$$

The given system of equations has a unique solution if and only if the matrix $\begin{bmatrix} I_{m\times m} & A \\ A^* & 0_{n\times n} \end{bmatrix}$ is non-singular, i.e., its kernel consists of a single vector (zero). Indeed, if a vector $\begin{bmatrix} r \\ x \end{bmatrix}$ is in the kernel, then $\begin{bmatrix} r + Ax \\ A^*r \end{bmatrix} = 0$, hence

$$r + Ax = 0 \qquad \text{and} \qquad A^*r = 0$$

Premultiplying the first equation by A^* gives

$$A^*r + A^*Ax = A^*Ax = 0$$

Since A^*A is nonsingular, we have $x = 0$, and then $r = -Ax = 0$, too. Hence the kernel of the given matrix is trivial, thus it is nonsingular.

Now the solution of the given system satisfies

$$\begin{bmatrix} r + Ax \\ A^*r \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}$$

hence

$$r + Ax = b \qquad \text{and} \qquad A^*r = 0$$

Premultiplying the first equation by A^* gives

$$A^*r + A^*Ax = A^*Ax = A^*b$$

hence x is a solution of the system of normal equations. Thus it minimizes $\|Ax - b\|_2$, as we learned in class. Finally, $r = b - Ax$ is the residual, as required.

JPE, September 1999, #4.

If x is an eigenvector corresponding to $\lambda = -10$, then

$$\frac{\|Ax\|}{\|x\|} = \frac{10\,\|x\|}{\|x\|} = 10,$$

hence $M(A) \geq 10$. On the other hand, if A is Hermitian, then

$$M(A) = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \|A\| = \max\{|\lambda| \colon \lambda \text{ is an e-value of } A\} = 10.$$

Hence 10 is the best lower bound for $M(A)$.

If x is an eigenvector corresponding to $\lambda = 0.01$, then

$$\frac{\|Ax\|}{\|x\|} = \frac{0.01\,\|x\|}{\|x\|} = 0.01,$$

hence $m(A) \leq 0.01$. On the other hand, if A is Hermitian, then

$$\begin{aligned} m(A) &= \min_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \min_{y \neq 0} \frac{\|y\|}{\|A^{-1}y\|} = \frac{1}{\max_{y \neq 0} \frac{\|A^{-1}y\|}{\|y\|}} = \frac{1}{\|A^{-1}\|} \\ &= \frac{1}{\max\{|\lambda^{-1}| \colon \lambda \text{ is an e-value of } A\}} = \frac{1}{0.01^{-1}} = 0.01. \end{aligned}$$

Hence 0.01 is the best upper bound for $m(A)$.

JPE, September 1998, #5 (and Sept. 1996, #3).

(a) If we choose an ONB in W and an ONB in $W^{\perp}$, then their union will be an ONB in V. In that basis, U is given by a diagonal matrix whose entries are $+1$ (coming from W) and -1 (coming from $W^{\perp}$). Thus U is unitary equivalent to a diagonal matrix, with eigenvalues being ± 1. Since the eigenvalues are real, the matrix U is self-adjoint. Since the eigenvalues have absolute value one, the matrix U is unitary.

(b) Note that

$$-U\alpha = -U(\gamma + \beta) = \gamma - \beta$$

for every $\gamma \in W^{\perp}$ and $\beta \in W$. Hence $-U$ is a reflector across the hyperplane $W^{\perp}$. By the Householder formula,

$$-U = I - 2\frac{xx^*}{\|x\|^2}$$

where $x = [1, 0, 1]^T$.

JPE, September 1997, #3.

(a) Routine calculation gives

$$J = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad P = \begin{bmatrix} i & 1 \\ 1 & 0 \end{bmatrix}$$

(note that P is not uniquely defined, you can choose it differently).

(b) For any 2×2 matrix B we have two options. If it is diagonalizable, then it is similar to a diagonal matrix, that is symmetric. If it is not diagonalizable, then it is similar to a Jordan matrix

$$B \sim \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = \lambda I + J = \lambda P^{-1} P + P^{-1} A P = P^{-1}(\lambda I + A) P$$

hence B is similar to the matrix $\lambda I + A$, which is symmetric. Here A and J are the same as in part (a).