FALL 2022: MATH 790 EXAM 2 SOLUTIONS

Throughout this exam, V will denote a finite dimensional vector space over the field F. When V is an inner product space, F will denote either one of the fields $\mathbb{R}$ or $\mathbb{C}$, unless stated otherwise. Each problem is worth 10 points. You may use your notes from class or your homework, **but you may not use your book, any other book, the internet, or consult with any students or professors, other than your Math 790 professor**. Please upload a **pdf file** containing your solutions to the exam to Canvas by 5pm on Monday, October 31.

1. Let V be a finite dimensional vector space $T \in \mathcal{L}(V,V)$. Assume T has all of its eigenvalues in F and $W \subseteq V$ is a T-invariant subspace.

(i) For any vector $v \in V \backslash W$, show that there exists a unique monic polynomial $p(x) \in F[x]$ of least degree such that $p(T)(v) \in W$ and if $g(x) \in F[x]$ satisfies $q(T)(v) \in W$, then $p(x)$ divides $g(x)$. In particular, $p(x)$ divides $\chi_T(x)$.

(ii) Conclude that we may write $p(x) = (x-\lambda)q(x)$, for some $\lambda \in F$ and $q(x) \in F[x]$.

(iii) Suppose $w_1, \ldots, w_k$ is a basis for W. Set $w_{k+1} := q(T)(v)$. Show that: $w_1, \ldots w_k, w_{k+1}$ are linearly independent and $W_{k+1} := \langle w_1, \ldots, w_{k+1}\rangle$ is T-invariant.

(iv) Show by induction, using (iii), that there exists a basis $B \subseteq V$ such that $[T]_B^B$ is upper triangular. In this case T is said to be *triangularizable.*

(v) Conclude that an operator T on a finite dimensional vector space is triangularizable if and only if T has all of its eigenvalues in F.

(vi) State a version of (v) for matrices in $\mathrm{M}_n(F)$ and use (v) to prove this statement.

Solution. For (i), it follows from the Well Ordering Principle in $\mathbb{Z}$ that there exists a monic polynomial $p(x)$ of least degree such that $p(T)(v) \in W$. Suppose $g(x) \in F[x]$ satisfies $g(T)(v) \in W$. Applying the division algorithm, we have $g(x) = p(x)h(x) + r(x)$, where $r(x) = 0$ or the degree of $r(x)$ is less than the degree of $p(x)$. In the latter case, we have $r(T)(v) = g(T)(v) - h(T)p(T)(v)$. We have $g(T)(v), p(T)(v) \in W$, by assumption. Moreover, $h(T)p(T)(v) \in W$, since W is T-invariant. Thus, $r(T)(v) \in W$, which contradicts the choice of $p(x)$. Thus, $r(x) = 0$, and therefore $p(x)$ divides $f(x)$. Uniqueness now follows from the division property. Moreover, $\chi_Y(T) = 0$, by the Cayley-Hamilton theorem, so that $\chi_T(T)(v) = 0 \in W$, and thus $p(x)$ divides $\chi_T(x)$.

For (ii), by the hypothesis on T, we have $\chi_T(x) = (x-\lambda_1)^{e_1} \cdots (x-\lambda_r)^{e_r}$, for distinct $\lambda_1, \ldots, \lambda_r \in F$. Since $\chi_T(x) = p(x)a(x)$, for some $a(x) \in F[x]$, we must have that some (possibly many) $x - \lambda_i$ divides $p(x)$. Setting $\lambda := \lambda_i$, we have $p(x) = (x-\lambda)q(x)$, for some $q(x) \in F[x]$.

For (iii), since $\deg(q(x)) < \deg(p(x))$, $w_{k+1} \notin W$. Thus, since $w_1, \ldots, w_k$ are linearly independent, $w_1, \ldots, w_k, w_{k+1}$ are also linearly independent. Moreover, since W is T-invariant, it suffices to show that $T(w_{k+1}) \in W_{k+1}$. For, this, on the one hand, $w := p(T)(v) = (T-\lambda)q(T)(v) = (T-\lambda)(w_{k+1})$ belongs to W. On the other hand, $w := T(w_{k+1}) - \lambda w_{k+1}$, so that $T(w_{k+1}) = w - \lambda w_{k+1}$ belongs to W_{k+1}, which is what we want.

For (iv), take w_1 an eigenvector for T, so that $W_1 := \langle w_1 \rangle$ is T-invariant. Suppose that we have found linearly independent vectors $w_1, \ldots, w_k$ such that for $1 \leq i \leq k$, $W_i := \langle w_1, \ldots, w_i \rangle$ is T-invariant. Then part (iii) shows that there exist linearly independent vectors $w_1, \ldots, w_{k+1}$ such that for $W_{k+1} := \langle w_1, \ldots, w_{k+1}\rangle$, W_{k+1} is T-invariant. Thus, by induction on k, if $n := \dim(V)$, there exists a basis $B := \{w_1, \ldots, w_n\}$ for V such that for all $1 \leq i \leq n$, $W_i := \langle w_1, \ldots, w_i\rangle$ is T-invariant. It follows that $[T]_B^B$ is upper triangular.

For (v), Suppose T is triangularizable, and $B \subseteq V$ is a basis such that $A := [T]_B^B$ is upper triangular. Let $\lambda_1, \ldots, \lambda_n$ in F be the diagonal entries of A. Since $xI_n - A$ is upper triangular with diagonal entries $x-\lambda_1, \ldots, x-\lambda_n$, it follows that $\chi_T(x) = \chi_A(x) = (x-\lambda_1)\cdots(x-\lambda_n)$, which shows that T has all of its eigenvalues in F. The converse is just part (iv).

For (vi), let $A \in \mathrm{M}_n(F)$. The analogous statement for A is that there exists an invertible matrix $P \in \mathrm{M}_n(F)$ such that $P^{-1}AP$ is upper triangular if and only A has all of its eigenvalues in F. Let $T : F^n \to F^n$ be defined by $T(v) := Av$, for all $v \in F^n$. Then $A = [T]_E^E$, where E is the standard basis for F^n. Suppose A has its of of it eigenvalues in F. Then the same applies to T. Thus, there exists a basis $B \subseteq F^n$ such that $[T]_B^B = C$, an upper triangular. Therefore,

$$C = [T]_B^B = [I]_E^B \cdot [T]_E^E \cdot [I]_B^E = P^{-1}AP,$$

where $P = [I]_B^E$. Conversely, if $P^{-1}AP$ is upper triangular, then since A and $P^{-1}AP$ have the same characteristic polynomial, it follows that A has all of its eigenvalues in F.

2. Let T be a linear transformation from the inner product space V to the inner product space W and assume $B_V \subseteq V$ and $B_W \subseteq W$ be orthonormal bases. Set $A := [T]_{B_V}^{B_W}$. Let A^* denote the conjugate transpose of A. Define $T^* : W \to V$ by the equation $[T^*]_{B_W}^{B_V} = A^*$[1]. Show that $\langle T(v), w\rangle_W = \langle v, T^*(w)\rangle_V$, for all $v \in V$ and $w \in W$.

Solution. Let $B_V := \{v_1, \ldots, v_n\}$, $B_W := \{w_1, \ldots, w_m\}$, so that $A := [T]_{B_V}^{B_W}$. Write $A = (a_{ij})$ and $A^* = (a_{ij}^*)$. For all i, j on the one hand, we have

$$\langle v_i, T^*(w_j)\rangle_V = \langle v_i, a_{1j}^* v_1 + \cdots + a_{nj}^* v_n\rangle_V = a_{ji},$$

while on the other hand,

$$\langle T(v_i), w_j\rangle_W = \langle a_{1i}w_1 + \cdots + a_{mi}w_m, w_j\rangle_W = a_{ji},$$

so that $\langle T(v_i), w_j\rangle_V = \langle v_i, T^*(w_j)\rangle_W$, for all i, j. Suppose $v = \Sigma_{i=1}^n \alpha_i v_i$ and $w = \Sigma_{j=1}^m \beta_j w_j$. Then

$$\langle T(v), w\rangle_W = \Sigma_{i,j}\alpha_i\overline{\beta_j}\langle T(v_i), w_j\rangle_W = \Sigma_{i,j}\alpha_i\overline{\beta_j}\langle v_i, T^*(w_j)\rangle_V = \langle v, T^*(w)\rangle_V,$$

which gives what we want.

3. For $T : V \to W$ as in problem 2, show that the definition of T^* is independent of the orthonormal bases chosen for V and W. (Hint: The formula $[T(v)]_{B_W} = [T]_{B_V}^{B_W} \cdot [v]_{B_V}$, and its variants, together with 6(iv) will be useful.)

Solution. Maintaining the notation from problem 3, let $C_V \subseteq V$ and $C_W \subseteq W$ be orthonormal bases, and set $D := [T]_{C_V}^{C_W}$. Thus, $D = [T]_{C_V}^{C_W} = [I]_{B_W}^{C_W} \cdot [T]_{B_V}^{B_W} \cdot [I]_{C_V}^{B_V}$. Let us define $T_1^* : W \to V$ by the equation $[T_1^*]_{C_W}^{C_V} = D^*$. We have to show that $T^* = T_1^*$. In other words, $T^*(w) = T_1^*(w)$, for all $w \in W$. We first present a solution based upon matrices. Since each of these vectors is determined by their representation as a linear combination of the basis elements in B_V, it suffices to show that $[T^*(w)]_{B_V} = [T_1^*(w)]_{B_V}$, for all $w \in W$. On the one hand, $[T^*(w)]_{B_V} = [T^*]_{B_W}^{B_V} \cdot [w]_{B_W}$. On the other hand,

$$\begin{aligned}
[T_1^*(w)]_{B_V} &= [T_1^*]_{B_W}^{B_V} \cdot [w]_{B_W}\\
&= ([I]_{C_V}^{B_V} \cdot [T_1^*]_{C_W}^{C_V} \cdot [I]_{B_W}^{C_W}) \cdot [w]_{B_W}\\
&= ([I]_{C_V}^{B_V} \cdot D^* \cdot [I]_{B_W}^{C_W}) \cdot [w]_{B_W}\\
&= ([I]_{C_V}^{B_V} \cdot \{[I]_{B_W}^{C_W} \cdot [T]_{B_V}^{B_W} \cdot [I]_{C_V}^{B_V}\}^* \cdot [I]_{B_W}^{C_W}) \cdot [w]_{B_W}\\
&= ([I]_{C_V}^{B_V} \cdot \{[I]_{B_V}^{C_V} A^* [I]_{C_W}^{B_W}\} \cdot [I]_{B_W}^{C_W}) \cdot [w]_{B_W}, \quad \text{since } [I]_{B_V}^{C_V} \text{ and } [I]_{B_W}^{C_W} \text{ are unitary, by 6(iv)}\\
&= A^* \cdot [w]_{B_W}\\
&= [T^*]_{B_W}^{B_V} \cdot [w]_{B_W},
\end{aligned}$$

which completes the proof.

For an proof in terms of operators, from the previous problem, we have $\langle v, T^*(w)\rangle_V = \langle T(v), w\rangle_V = \langle v, T_1^*(w)\rangle_V$, for all $v \in V$ and $w \in W$. Thus, $\langle v, T^*(w) - T_1^*(w)\rangle_V = 0$, for all $v \in V$ and $w \in W$, from which we have $T^*(w) = T_1^*(w)$, for all $w \in W$.

4. Maintaining the notation from problem 2, prove that:

(i) $\langle T^*(w), v\rangle_V = \langle w, T(v)\rangle_W$, for all $v \in V$ and $w \in W$.

(ii) $(T^*)^* = T$.

(ii) If $S : W \to U$ is a linear transformation of inner product spaces, then $(ST)^* = T^*S^*$.

Solution. For (i), $\langle T^*(w), v\rangle_V = \overline{\langle v, T^*(w)\rangle_V} = \overline{\langle T(v), w\rangle_W} = \langle w, T(v)\rangle_W$. For (ii), for all $v \in V$ and $w \in W$, we have $\langle T(v), w\rangle_V = \langle v, T^*(w)\rangle_V = \langle T^{**}(v), w\rangle_V$, the latter equality following from part (i). Thus, $T^{**}(v) = T(v)$, for all $v \in T$, i.e., $T = T^{**}$. For (iii), for all $v \in V$ and $u \in U$, $\langle ST(v), u\rangle_U = \langle v, (ST)^*(u)\rangle_V$, while on the other hand, $\langle ST(v), u\rangle_U = \langle T(v), S^*(u)\rangle_W = \langle v, T^*(S^*(u))\rangle_V$. Thus, $\langle v, (ST)^*(u)\rangle_V = \langle v, T^*(S^*(u)\rangle_V$ for all $v \in V$, from which it follows that $(ST)^*(u) = T^*(S^*(u))$, for all $u \in U$, and therefore $(ST)^* = T^*S^*$.

5. Assume that V is an inner product space and $T \in \mathcal{L}(V, V)$. Give a detailed proof showing $\text{rank}(T) = \text{rank}(T^*)$.

Solution. Let B denote an orthonormal basis for V and set $A := [T]_B^B$ and $A^* := [T^*]_B^B$. Then $\text{rank}(T) = \text{rank}(A)$ and $\text{rank}(T^*) = \text{rank}(A^*)$, so it suffices to show that A and A^* have the same rank. Ultimately, this is just the fact that the row rank and column rank of A are the same. To see this, if we use elementary row operations to to put A into its reduced row echelon form $\tilde{A}$, then A and $\tilde{A}$ have the same row space, and hence the same row rank. On the

[1]Recall that to define a linear transformation from W to V, it suffices to specify how that transformation acts on a basis for W.

other hand, the null space of A and $\tilde{A}$ are the same, so A and $\tilde{A}$ have the same nullity, and therefore the same rank, by the Rank plus Nullity theorem. But the row rank and column rank of a matrix in reduced row echelon form are clearly the same, by counting the number of pivots.

Now, the rank of A is the dimension of the column space of A which equals the the dimension of the row space of A^t, which in turn equals the rank of A^t. Thus, A and A^t have the same rank. Thus, if $F = \mathbb{R}$, we are done. Suppose $F = \mathbb{C}$. Then for any matrix B, columns $C_1, \dots C_r$ from B are linearly independent over $\mathbb{C}$ if and only if the conjugate of those columns are linearly independent over $\mathbb{C}$. Thus, B and its conjugate $\overline{B}$ have the same rank. Applying this to the case $B = A^t$ shows that $\text{rank}(A) = \text{rank}(A^t) = \text{rank}(\overline{A}^t)$, which gives what we want.

6. Let $T \in \mathcal{L}(V, V)$ be a linear operator. Show that the conditions (i)-(iv) below are equivalent. Any $T \in \mathcal{L}(V, V)$ satisfying these conditions is called an *isometry*. This is the operator analogue of a unitary or orthogonal matrix.

(i) $||T(v)|| = ||v||$, for all $v \in V$.
(ii) $\langle T(v), T(w)\rangle = \langle v, w\rangle$, for all $v, w \in V$.
(iii) T takes an orthonormal basis of V to an orthonormal basis.
(iv) $[T]_{B_1}^{B_2}$ is a unitary matrix for all orthonormal bases $B_1, B_2 \subseteq R$.

Solution. That (i)implies (ii) proceeds just as the proof given in class (and Homework 16) that $||(Tv)|| = ||T^*(v)||$ for all $v \in V$ is equivalent to $\langle T(v), T(w)\rangle = \langle T^*(v), T^*(w)\rangle$, for all v, w. One just replaces T^* by the identity map. Now suppose (ii) holds. If $\{v_1, \dots, v_n\}$ is an orthonormal basis, then $\langle T(v_i), v_j)\rangle = \langle v_i, v_j\rangle = \delta_{ij}$, for all i, j showing that $\{T(v_1), \dots, T(v_j)\}$ is an orthonormal basis. Suppose (iii) holds and $B_1, B_2 \subseteq V$ are orthonormal bases. We have $[T]_{B_1}^{B_2} = [T]_{B_2}^{B_2} \cdot [I]_{B_1}^{B_2}$, so it suffices to check that $[T]_B^B$ is unitary whenever B is an orthonormal basis and $[I]_{B_1}^{B_2}$ is unitary, since if $A = BC$, and B, C are unitary then

$$A^* = (BC)^* = C^*B^* = C^{-1}B^{-1} = (BC)^{-1} = A^{-1}.$$

Suppose $B = \{v_1, \dots, v_n\}$ is an orthonormal basis for V and $A = (a_{ij})$ is $[T]_B^B$. Let us write $A^* = (a_{ij}^*)$, so that $a_{ij}^* = \overline{a_{ji}}$, for all i, j. Then, $\langle T(v_i), T(v_j)\rangle = \langle v_i, v_j\rangle = \delta_{ij}$, the Kroenecker delta. On the other hand,

$$\begin{aligned}\langle T(v_i), T(v_j)\rangle &= \langle a_{1i}v_1 + \cdots + a_{ni}v_n, a_{1j}v_1 + \cdots + a_{nj}v_n\rangle \\ &= \Sigma_{r=1}^n a_{ri}\overline{a_{rj}} \\ &= \Sigma_{r=1}^n a_{ri}a_{jr}^* \\ &= \Sigma_{r=1}^n a_{jr}^* a_{ri} \\ &= \text{ the } (j, i) \text{ entry of } A^*A.\end{aligned}$$

Thus, $A^*A = I_n$ which shows that A is unitary. Now suppose $B = B_1$ and $B_2 = \{w_1, \dots, w_n\}$ is another orthonormal basis. Let us again write $A = [I]_{B_1}^{B_2}$. Then,

$$\begin{aligned}\delta_{ij} = \langle v_i, v_j\rangle &= \langle \Sigma_{r=1}^n a_{ri}v_r, \Sigma_{s=1}^n a_{sj}w_s\rangle \\ &= \Sigma_{r,s} a_{ri}\overline{a_{sj}}\langle w_r, w_s\rangle \\ &= \Sigma_{r=1}^n a_{jr}^* a_{ri} \\ &= \text{ the } (j, i) \text{ entry of } A^*A.\end{aligned}$$

Thus, $A^*A = I_n$ which shows that A is unitary. Therefore, (iii) implies (iv). To see that (iv) implies (ii), note that in particular, $[T]_B^B$ is unitary, for B as above. Reversing the first sequence of equations above will show that $\langle T(v_i), T(v_j)\rangle = \langle v_i, v_j\rangle$, from which it readily follows that $\langle T(v), T(w)\rangle = \langle v, w\rangle$, for all $v, w \in V$.

7. Let $T : \mathbb{R}^4 \to \mathbb{R}^4$ have the property that its minimal polynomial is a product of two distinct irreducible polynomials of degree two. Prove that T is an isometry if and only if there is an orthonormal basis of $\mathbb{R}^4$ such that

$$[T]_B^B = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 & 0 \\ -\sin(\theta) & \cos(\theta) & 0 & 0 \\ 0 & 0 & \cos(\gamma) & \sin(\gamma) \\ 0 & 0 & -\sin(\gamma) & \cos(\gamma) \end{pmatrix},$$

for $0 \leq \theta, \gamma \leq \pi$.

Solution. Suppose first that T is an isometry. We first note that $T^* = T^{-1}$. To see this, for all $v, w \in V$, we have $\langle v, w\rangle = \langle T(v), T(w)\rangle = \langle v, T^*T(w)\rangle$. Since we also have $\langle v, w\rangle = \langle v, T^{-1}(T(w)\rangle$, for all $v, w \in V$, it follows that $T^* = T^{-1}$.

The proof that the required basis B exists reduces to the case of a two-dimensional vector space over $\mathbb{R}$ if one has a T-invariant subspace W_1 and W_2 of dimension such that $V = W_1 \oplus W_2$ and each W_i is T-invariant. For then

T restricted to each subspace is still an isometry, by condition (i) in the previous problem. Now, by assumption, $\mu_T(x) = p_1(x)p_2(x)$ where each $p_i(x)$ is irreducible of degree two over $\mathbb{R}$. Thus, $V = W_1 \oplus W_2$, where $W_i := \ker(p_i(x))$, and each W_i is T-invariant. By hypothesis, T has no eigenvalues, so each W_i must have dimension two, so we have found the required W_i.

Now, changing notation, consider $T : W \to W$ an isometry of the two dimensional vector space W. Let $B = \{u_1, u_2\}$ be an orthonormal basis. Thus, by property (iii) in problem 6, $\{T(u), T(v)\}$ is an orthonormal basis. Therefore, $[T]_B^B$ is a 2×2 orthogonal matrix over R. Suppose $[T]_B^B = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Since the columns of this matrix form an orthonormal basis for $\mathbb{R}^2$, we can write $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$ and $\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$ or $\begin{pmatrix} \sin(\theta) \\ -\cos(\theta) \end{pmatrix}$. In the first case, if we set $\theta := -\alpha$, then $[T]_B^B = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$ and in the second case, writing $B' := \{-u_2, u_1\}$, $[T]_{B'}^{B'} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$.

On the other hand, if $B \subseteq \mathbb{R}^4$ is an orthonormal basis such that $[T]_B^B$ has the form given in the statement of the problem, then the columns of $[T]_B^B$ form an orthonormal basis for $\mathbb{R}^4$. Thus, T takes an orthonormal basis to an orthonormal basis, so that T is an isometry by part (iii) of the previous problem.

8. Given $C = \begin{pmatrix} -4 & 2 & -2 \\ 2 & -7 & 4 \\ -2 & 4 & -7 \end{pmatrix}$, find an orthogonal matrix Q such that $Q^{-1}CQ$ has the eigenvalues of C down its diagonal. Now find a 2×2 matrix over $\mathbb{R}$ that is diagonalizable, but not orthogonally diagonalizable.

Solution Outline. $\chi_A(x) = (x+3)^2(x+12)$. E_{-3} is the null space of the matrix $\begin{pmatrix} -1 & 2 & -2 \\ 2 & -4 & 4 \\ -2 & 4 & -4 \end{pmatrix}$ which simplifies to $\begin{pmatrix} 1 & -2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ via elementary row operations. From this, one obtains a basis $\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$ for E_2. Applying Gram-Schmidt gives the following orthonormal basis for E_{-3}: $u_1 := \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix}$, $u_2 = \begin{pmatrix} \frac{-2}{3\sqrt{5}} \\ \frac{4}{3\sqrt{5}} \\ \frac{5}{3\sqrt{5}} \end{pmatrix}$.

E_{-12} is the nullspace of the matrix $\begin{pmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{pmatrix}$ which has the unit vector $u_3 = \begin{pmatrix} \frac{1}{3} \\ \frac{-2}{3} \\ \frac{2}{3} \end{pmatrix}$ for a basis. Note that u_1, u_2, u_3 is an orthonormal basis for $\mathbb{R}^3$ and if Q is the matrix with columns u_1, u_2, u_3, Q is an orthogonal matrix and satisfies $Q^{-1}AQ = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 12 \end{pmatrix}$.

The matrix $B := \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ over $\mathbb{R}$ is diagonalizable because it has distinct eigenvalues; it is not orthogonally diagonalizable because it is not a symmetric matrix.

9. Let $A \in M_n(\mathbb{R})$ and suppose that the columns $v_1, \ldots, v_n$ of A are mutually orthogonal with lengths $\sigma_1, \ldots, \sigma_n$, respectively. Find the Singular Value Decomposition of A.

Solution. The hypothesis on A implies that A^tA is the $n \times n$ diagonal matrix with $\sigma_1^2, \ldots, \sigma_n^2$ down its diagonal. Without loss of generality, we assume $\sigma_1 \geq \cdots \geq \sigma_n$. Thus, Σ is the $n \times n$ diagonal matrix with $\sigma_1, \ldots, \sigma_n$ down its diagonal. If we let $e_1, \ldots, e_n$ denote the standard ordered basis for $\mathbb{R}^n$, then this is clearly an orthonormal basis consisting of eigenvectors for A^tA, so we take I_n, the $n \times n$ identity matrix, for P. On the other hand, $Ae_1 = v_1, \ldots, Ae_n = v_n$ are the columns of A. By the hypothesis on A, it follows that $\frac{1}{\sigma_1}v_1, \ldots, \frac{1}{\sigma_n}v_n$ is an orthonormal basis for $\mathbb{R}^n$, so we take for Q the matrix having these vectors as columns. A direct calculation shows that $Q^tAP = \Sigma$, thus, $A = Q\Sigma P^t = Q\Sigma I_n = Q\Sigma$ is the singular value decomposition of A.

10. Let $F[x]$ denote the ring of polynomials with coefficients in the field F.

(i) Let $p(x) \in F[x]$ be a non-constant irreducible polynomial. Prove that for any non-constant $f(x)$ in $F[x]$, the GCD of $p(x), f(x)$ is either $p(x)$ or 1.

(ii) Show that if $p(x)$ is irreducible over F and $p(x)$ divides $f(x) \cdot g(x)$, then $p(x)$ divides $f(x)$ or $p(x)$ divides $g(x)$. (Hint: Use (i) and Bezout's Principle.)

(iii) Prove that if $p_1(x)\cdots p_r(x) = q_1(x)\cdots q_s(x)$, and each $p_i(x), q_j(x)$ is irreducible over F, then $r = s$, and after re-indexing, $q_i(x) = \alpha_i \cdot p_i(x)$, for some $\alpha_i \in F$. In other words, the factorization property for polynomials in $F[x]$ is in fact a *unique factorization* property.

Solution. Part (i) follows since, any GCD of $p(x)$ and $f(x)$ must divide $p(x)$, while at the same time, only 1 and $p(x)$ divide $p(x)$, since $p(x)$ is irreducible. For part (ii) suppose $p(x)$ divides $f(x)\cdot g(x)$. If $p(x)$ divides $f(x)$, we're done. Otherwise, the GCD of $p(x)$ and $f(x)$ is 1. By Bezout's Principle, there exist $a(x), b(x) \in F[x]$ such that $1 = a(x)f(x) + b(x)p(x)$. Multiplying by $g(x)$ gives $g(x) = a(x)g(x)f(x) + g(x)b(x)p(x)$. Since $p(x)$ divides $f(x)g(x)$, we may write $f(x)g(x) = c(x)p(x)$. Therefore,

$$g(x) = a(x)c(x)p(x) + g(x)b(x)p(x) = (a(x)c(x) + g(x)b(x))p(x),$$

showing that $p(x)$ divides $g(x)$, as required. Keeping the notation in the statement of (iii), we assume $r \leq s$ and induct on r. When $r = 1$, we have $p_1(x) = q_1(x)\cdots q_s(x)$. Since $p_1(x)$ is irreducible, we must have $s = 1$ and $p_1(x) = q_1(x)$.

Now suppose $r > 1$. Then $p_1(x)|q_1(x)\cdots q_s(x)$, so by part (ii), and an easy induction, $p_1(x)$ divides some $q_i(x)$. After indexing, we may assume $p_1(x)$ divides $q_1(x)$. Since $q_1(x)$ is irreducible, this can only happen if $q_1(x) = \alpha_1 \cdot p_1(x)$, for some $\alpha_1 \in F$. Thus,

$$p_1(x)p_2(x)\cdots p_r(x) = (\alpha_1 p_1(x))q_2(x)\cdots q_s(x),$$

so that $p_2(x)\cdots p_2(x) = q_2'(x)\cdots q_s(x)$, where $q_2'(x) = \alpha_1 q_2(x)$ is irreducible over F. By induction on r, $r-1 = s-1$, and thus, $r = s$, and moreover, there exist $\alpha_i \in F$ such that $q_2'(x) = \alpha_2 p_2(x), \ldots, q_r(x) = \alpha_r p_r(x)$, as required.

Bonus Problem. Let V be an inner product space over $\mathbb{C}$. Prove that the set of self-adjoint operators on V is a vector space over $\mathbb{R}$ and find the dimension of this vector space over $\mathbb{R}$.

Solution. We first recall that any vector space over $\mathbb{C}$ is automatically a vector space over $\mathbb{R}$. Thus, we may regard $\mathcal{L}(V,V)$ as a real vector space. Let $\mathcal{A}$ denote the self-adjoint operators in $\mathcal{L}(V,V)$. We need to show that $\mathcal{A}$ is a subspace of $\mathcal{L}(V,V)$, when regarded as a vector space over $\mathbb{R}$. If $S, T \in \mathcal{A}$, then $(S+T)^* = S^* + T^* = S + T$ and if $\lambda \in \mathbb{R}$, then $(\lambda T)^* = \overline{\lambda}T^* = \lambda T$, so $\mathcal{A}$ is a subspace of $\mathcal{L}(V,V)$.

For the second part, we note that it suffices to calculate the dimension of the space of self-adjoint matrices in $M_n(\mathbb{C})$ as a vector space over the real numbers[2], and for this, we begin by noting that if $C \in M_n(\mathbb{C})$, then $C = A + iB$, where $A, B \in M_n(\mathbb{R})$, and A, B are unique. Thus, $C^* = A^t - iB^t$, so that $C = C^*$ if and only if $A = A^t$ and $B = -B^t$. Note that B must have zeros down its diagonal. For $i = 1, \ldots, n$, let $D_i \in M_n(\mathbb{R})$ denote the matrix with 1 as its (i,i) entry and zeros elsewhere, and for $1 \leq i < j \leq n$, let $S_{ij} \in M_n(\mathbb{R})$ denote the matrix with 1 in the (i,j) and (j,i) entries and zeros elsewhere. Note that the collections $\{D_i\}$ and $\{S_{ij}\}$ together form a basis for the space of real symmetric matrices over $\mathbb{R}$, i.e., any matrix $A \in M_n(\mathbb{R})$ satisfying $A^t = A$ can be written uniquely as an $\mathbb{R}$-linear combination of these matrices. Now, for $1 \leq u < v \leq n$, let $T_{uv} \in M_n(\mathbb{R})$ denote the matrix with 1 in the (u,v) position, -1 in the (v,u) position and zeros elsewhere, so that any $B \in M_n(\mathbb{R})$ satisfying $B^t = -B$ can be written uniquely as a linear combination of the T_{uv}. It now follows that since any self-adjoint matrix $C \in M_n(\mathbb{C})$ can be written as $A + iB$, with $A, B \in M_n(\mathbb{R})$, $A = A^t$ and $B^t = -B$, the set $\{D_i\} \cup \{S_{ij}\} \cup \{iT_{uv}\}$ forms a basis for the real vector space of complex self-adjoint matrices. Thus, the dimension of this space is $n + (1 + \cdots + (n-1)) + (1 + \cdots + (n-1)) = n^2$.

[2] Letting B denote an orthonormal basis for V, the map from $\mathcal{L}(V,V)$ to $\mathrm{M}_n(\mathbb{C})$ taking T to $[T]_B^B$ is an isomorphism of vector spaces satisfying $T^* \to ([T]_B^B)^*$, giving the required isomorphism.