Math 225- HW 11 Due: Dec 9 by Midnight

Submit the first two problems, along with any three additional problems of your choice.

1. - Prove that if U and T simultaneously diagonalizable then U and T commute. i.e. $UT = TU$

     If U and T are simultaneously diagonalizable, this means that $[T]_\beta$ and $[U]_\beta$ are diagonal matrices. The product of two diagonal matrices is obviously commutative (i.e. if matrices X and Y are diagonal, $XY = YX$). Thus, if U and T are simultaneously diagonalizable then:

$$[T]_\beta[U]_\beta = [U]_\beta[T]_\beta$$
$$[TU]_\beta = [UT]_\beta$$
$$TU = UT$$

   - Conclude that if matrices A,B are simultaneously diagonalizable then A,B commute

     If A and B are simultaneously diagonalizable then we know that $\exists\ Q$ invertible s.t. $Q^{-1}AQ$ and $Q^{-1}BQ$ are diagonal and so, using the logic that the product of two diagonal matrices is commutative, we have:

$$Q^{-1}AQQ^{-1}BQ = Q^{-1}BQQ^{-1}AQ$$
$$Q^{-1}ABQ = Q^{-1}BAQ$$
$$AB = BA$$

   - Let T be diagonalizable linear operator on a finite dimensional vector space, then T and T^m are simultaneously diagonalizable for any m positive integer.

     Because T is diagonalizable, $[T]_\beta$ is a diagonal matrix. Thus, $[T^m]_\beta = [\underbrace{T \dots T}_{m \text{ times}}]_\beta = \underbrace{[T]_\beta \dots [T]_\beta}_{m \text{ times}} = \prod_{i=1}^{m}[T]_\beta$. Because $[T]_\beta$ is diagonal, $[T^m]_\beta = \prod_{i=1}^{m}[T]_\beta$ is also diagonal $\Rightarrow T$ and T^m are simultaneously diagonalizable.

2. a) For any vector $w \in E_\lambda$, $T(w) = \lambda w \in E_\lambda$. Let us define $u = \lambda w$. Because $T(u) = T(\lambda w) = \lambda T(w) = \lambda^2 w = \lambda u$, $u = \lambda w \in E_\lambda$. Thus we have shown $\forall w \in E_\lambda, T(w) \in E_\lambda \Rightarrow E_\lambda$ is a T-invariant subspace of V.

   b) Let us define this T-cyclic subspace generated by v as $W \leq V$. W can be expressed as $\text{Span}(\{v, T(v), \dots T^n(v)\})$. We now show that $\forall w \in W, T(w) \in W$. By definition, $\forall w \in W$ can be expressed as $w = \Sigma_{i=0}^{n} c_i T^i(v)$ and so $T(w) = c_n T^{n+1}(v) + \Sigma_{i=1}^{n} c_i T^i(v)$. Because $T^{n+1}(v)$ can be expressed as a linear combination of $\{v, T(v), \dots T^n(v)\}$, this means that $T(w)$ can be expressed as a linear combination of $\{v, T(v), \dots T^n(v)\} \Rightarrow T(w) \in \text{Span}(\{v, T(v), \dots T^n(v)\}) \Rightarrow T(w) \in W$. Thus, we have shown $\forall w \in W, T(w) \in W \Rightarrow W$ is a T-invariant subspace of V.

   c) The T-cyclilc subspace W can be given by $W = \text{Span}\{v, T(v), \dots T^n(v)\}$. We now prove both directions of this statement:

      1. If $w \in W$, $w = g(T)v$

         If $w \in W$, w can be expressed as $\Sigma_{i=0}^{n} c_i T^i(v) = U(v)$, where $U = \Sigma_{i=0}^{n} c_i T^i$ is an operator. Defining $g(x) = \Sigma_{i=0}^{n} c_i x^i$, $U = g(T)$ and so we have that $w = g(T)v$.

2. If $w = g(T)v$, $w \in W$
We can express polynomial g as $g(x) = \Sigma_{i=0}^{n} c_i x^i$. Thus, we have that $w = g(T)v = \Sigma_{i=0}^{n} c_i T^i(v) \Rightarrow w \in \text{Span}\{v, T(v), \dots T^n(v)\} \Rightarrow w \in W$.

d) Because V is a T-cyclic subspace of itself, we can express $V = \text{Span}(\{v, T(v), T^2(v), \dots, T^n(v)\})$. Thus, this means $\forall z \in V$, $z = \Sigma_{i=0}^{n} c_i T^i(v) \Rightarrow$ because $U(v) \in V, U(v)$ can be expressed as a linear combination of $T^i(v)$. Note that if U commutes with T, that means $UT^2 = UTT = TUT = TTU = T^2U$, or more generally $UT^\alpha = T^\alpha U$ for $\alpha \geq 0$. Thus, we have that for $i \geq 0$:

$$
\begin{aligned}
UT^i &= T^iU \\
UT^i(v) &= T^iU(v) \\
U(T^i(v)) &= T^i(\Sigma_{k=0}^{n} c_k T^k(v)) \\
U(T^i(v)) &= \Sigma_{k=0}^{n} c_k T^{i+k}(v)
\end{aligned}
$$

Setting $a = T^i(v)$, we have:

$$U(a) = \Sigma_{k=0}^{n} c_k T^k(a)$$

Thus, we can clearly see that $U = g(T)$, where polynomial g is given by $g(x) = \Sigma_{k=0}^{n} c_k x^k$.

e) There are two cases in this scenario: (1) all vectors in V are eigenvectors or (2) not all vectors in V are eigenvectors. We address both cases below:

1. All vectors in V are eigenvectors
This means that $\forall v \in V, T(v) = \lambda v$ where $\lambda \in \mathbb{F}$. Let us define two vectors $a, b \in V$ and compute $T(a+b)$:

$$T(a+b) = \lambda_{a+b}(a+b)$$

However, by linearity, we also have that $T(a+b) = T(a) + T(b) = \lambda_a a + \lambda_b b$. Thus, we have that:

$$
\begin{aligned}
T(a+b) &= T(a+b) \\
\lambda_{a+b}(a+b) &= \lambda_a a + \lambda_b b
\end{aligned}
$$

This means that $\lambda_{a+b} = \lambda_a = \lambda_b \Rightarrow \forall v \in V, T(v) = \lambda_a v \Rightarrow T = cI$ where $c \in \mathbb{F}$.

2. Not all vectors in V are eigenvectors
This means that $\exists\ v \neq 0 \in V$ s.t. $T(v) \neq \lambda v$, $\forall \lambda \in \mathbb{F}$. Consider the set $\{v, T(v)\}$. In order for the set of vectors $\{a, b\}$ to be linearly independent, neither a nor b can be expressed as a scalar multiple of the either vector. Because we know that $\forall \lambda \in \mathbb{F}$, $T(v) \neq \lambda v$, $\{v, T(v)\}$ are a linearly independent set of two vectors $\Rightarrow$ because $\dim(V) = 2$, $\{v, T(v)\}$ serve as a basis for $V \Rightarrow V = \text{Span}(\{v, T(v)\}) \Rightarrow V$ is a T-cyclic subspace of itself.

3. I didn't do this question.

4. (a) We use induction to prove this statement.

1. Base Case: Single element v_1
If $n = 1$, then given $\Sigma_{i=1}^{n} v_i \in W \Rightarrow v_1 \in W$.

2. Inductive Step: Given $v_1, \ldots, v_{k-1} \in W$, prove that $v_k \in W$
For proof by contrapositive, let us assume that $v_k \notin W$. Let us define $v = v_1 + \cdots + v_n$. We start with our given:

$$v = v_1 + \cdots + v_n \in W$$
$$v = (v_1 + \cdots + v_{k-1}) + v_k + (v_k + \cdots + v_n)$$
$$v_1 + \cdots + v_{k-1} = v - v_k - (v_k + \cdots + v_n)$$

Because W is a subspace, it is closed under addition. Thus, because $v_k \notin W \Rightarrow v - v_k - (v_k + \cdots + v_n) \notin W \Rightarrow v_1 + \cdots + v_{k-1} \notin W$. By proof by contrapositive, we have proven if $v_1, \ldots, v_{k-1}$, then $\in W, v_k \in W$.

(b) Let us define U as a non-trivial T-invariant subspace of V. If T is a diagonalizable linear operator, that means its eigenvectors $v_1, v_2, \ldots, v_n$ form a basis for V. Because U is a non-trivial subspace, $\exists\ v \neq 0 \in U$. Furthermore, given that $\forall v \neq 0 \in U \leq V$, v can be written as a linear combination of $\{v_1, \ldots, v_n\}$, we can define the nonempty set of eigenvectors which all elements of U are a linear combination of as $\{u_1, \ldots, u_k\} \Rightarrow \text{Span}(\{u_1, \ldots, u_k\}) = U$. Note that $\{u_1, \ldots, u_k\}$ are all part of the basis $\{v_1, \ldots, v_n\}$ for V and so they are are all linearly independent. Thus, we can conclude the linearly independent and generating set of eigenvectors $\{u_1, \ldots, u_k\}$ forms a basis for U and so $T|_U$ is diagonalizable.

(c) Because $v_1, v_2, \ldots, v_n \in V$ all correspond to distinct eigenvalues, they are all linearly independent. Given these n linearly independent vectors and that $\dim(V) = n$, we can conclude that the eigenvectors $v_1, v_2, \ldots, v_n$ form a basis for V. This means that $V = \text{Span}(\{v_1, v_2, \ldots, v_n\})$. Let us define vector $v = v_1 + v_2 + \cdots + v_n$. Note that $\text{Span}(\{v, T(v), \ldots, T^n(v)\}) = \text{Span}(\{\Sigma_{i=1}^n v_i, \Sigma_{i=1}^n \lambda_i v_i, \ldots, \Sigma_{i=1}^n \lambda_i^n v_i\})$. We can write out this transformation from the eigenvectors $v_1, \ldots, v_n$ to $\Sigma_{i=1}^n v_i, \Sigma_{i=1}^n \lambda_i v_i, \ldots \Sigma_{i=1}^n \lambda_i^n v_i$ as such:

$$\begin{bmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \vdots & \ddots & \ddots & \vdots \\ \lambda_1^n & \lambda_2^n & \ldots & \lambda_n^n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \Sigma_{i=1}^n v_i \\ \Sigma_{i=1}^n \lambda_i v_i \\ \vdots \\ \Sigma_{i=1}^n \lambda_i^n v_i \end{bmatrix}$$

Note that the leftmost matrix above, which I refer to as V, is the Vandermonde matrix (pg 230.) Because all $\forall 0 \leq i < j \leq, \lambda_i \neq \lambda_j$, $\det(V) = \prod_{0 \leq i < j \leq n} (\lambda_j - \lambda_i) \neq 0 \Rightarrow V$ is invertible $\Rightarrow$ because $\{v_1, v_2, \ldots, v_n\}$ serve as a basis for V, so does $\{\Sigma_{i=1}^n v_i, \Sigma_{i=1}^n \lambda_i v_i, \ldots, \Sigma_{i=1}^n \lambda_i^n v_i\} \Rightarrow V = \text{Span}(\{v, T(v), \ldots, T^n(v)\}) \Rightarrow V$ is a T-cyclic subspace of itself.

5. (a) We prove both directions of this statement below:

1. If T is diagonalizable, V is the direct sum of one-dimensional T-invariant subspaces
If T is diagonalizable, that means that eigenvectors $v_1, v_2, \ldots v_n \in V$ serve as a basis for V. This means $V = \text{Span}(\{v_1, \ldots, v_n\}) = \{\Sigma_{i=1}^n c_i v_i : c_i \in \mathbb{R}\} = \{\Sigma_{i=1}^n T(\frac{c_i}{\lambda_i} v_i)\}$. Let us define the set $W_j = \{cv_j : c \in \mathbb{F}\}$ for a given eigenvector v_j. Note that W_j is a one-dimensional subspace as W_j is composed of scalar multiples of one unique vector, v_j. Furthermore, W_j is a T-invariant subspace

as $\forall w \in W_j, T(w)$ is equal to a scalar multiple of $v_j \Rightarrow w \in W_j$. Furthermore, because $\{v_1, \ldots, v_n\}$, serve as a basis, that means that all the eigenvectors are linearly independent $\Rightarrow$ for $0 \leq i < j \leq n$, $W_i \cap W_j = \emptyset$ and so we have that $V = \{\Sigma_{i=1}^{n} c_i v_i : c_i \in \mathbb{R}\} = W_1 \bigoplus W_2 \cdots \bigoplus W_n$.

2. If V is the direct sum of one-dimensional T-invariant subspaces, T is diagonalizable
Let us define $V = W_1 \bigoplus W_2 \cdots \bigoplus W_k$ and the one basis vector for the jth subspace W_j as v_j. Because W_j is one-dimensional and a subspace (thus closed under addition and scalar multiplication), $W_j = \{cv_j : c \in \mathbb{F}\}$. Furthermore, because W_j is T-invariant and $v_j \in W$, $T(v_j) \in W_j \Rightarrow T(v_j) \in \{cv_j : c \in \mathbb{F}\} \Rightarrow v_j$ is an eigenvector of T_{W_j} as $T(v_j) = kv_j$ where $k \in \mathbb{F}$. Because V is a *direct* sum of $W_1, \ldots W_k$, the individual basis vector v_j for each subspace is linearly independent from all of the vectors in $\{v_1, \ldots, v_{j-1}\} \cup \{v_{j+1}, \ldots, v_k\} \Rightarrow \{v_1, \ldots, v_k\}$ are linearly independent[1]. Furthermore, $V = W_1 \bigoplus W_2 \cdots \bigoplus W_k$ means that V contains all possible linear combinations of $\{v_1, \ldots v_k\} \Rightarrow V = \text{Span}(\{v_1, \ldots, v_k\}$. Thus we can conclude that the linearly independent and generating eigenvectors $\{v_1, \ldots, v_k\}$ forms a basis for V and so T is diagonalizable.

b) Let us define the unordered basis for the T-invariant subspace W_j as β_j. This means that the ordered basis γ for vector space V can be given as $\gamma = \beta_1 \cup \beta_2 \cdots \cup \beta_k$. We now try to understand what the matrix $[T]_\beta$ looks like. Note that $\forall v \in \beta_j, v \in W_j$ and so $T(v) = T_{W_j}(v) \in W_j \Rightarrow T(v)$ can be expressed as a linear combination of β_j. Thus, $[T]_\beta$ will be given as a collection of block matrices $[T_{W_j}]_{\beta_j}$ along the diagonal:

$$[T]_\beta = \begin{bmatrix} [T_{W_1}]_{\beta_1} & 0 & \cdots & 0 \\ 0 & [T_{W_2}]_{\beta_2} & \cdots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & [T_{W_k}]_{\beta_k} \end{bmatrix}$$

From this matrix, it is obvious that:

$$\det(T) = \det([T]_\beta) = \prod_{i=1}^{k} \det([T_{W_i}]_{\beta_i}) = \prod_{i=1}^{k} \det(T_{W_i})$$

$$\det(T) = \prod_{i=1}^{k} \det(T_{W_i})$$

6. To prove this law, we compare the LHS with the RHS. The LHS can be given as:

$$\|x+y\|^2 + \|x-y\|^2 = \langle x+y, x+y\rangle + \langle x-y, x-y\rangle = \langle x, x+y\rangle + \langle y, x+y\rangle + \langle x, x-y\rangle - \langle y, x-y\rangle$$
$$= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle + \langle x,x\rangle - \langle x,y\rangle - \langle y,x\rangle + \langle y,y\rangle = 2\langle x,x\rangle + 2\langle y,y\rangle$$

We now compare this with the RHS:

[1] This can be trivially proven by induction by the following proof: $\{v_1\}$ is a linearly independent set, $\{v_1, v_2\}$ is a linearly independent set, $\{v_1, v_2, v_3\}$ is a linearly independent set, and so on until $\{v_1, \ldots, v_n\}$ is a linearly independent set. I believe we did this proof in class.

$$2\|x\|^2 + 2\|y\|^2 = 2\langle x, x\rangle + 2\langle y, y\rangle$$

Thus, we can clearly see that the LHS = RHS and so we have proven this law.