NOTES FOR 17 OCT (THURSDAY)

1. Recap

(1) Proved that eigenvectors corresponding to distinct eigenvalues are linearly independent.
(2) Proved that $gM \leq aM$ with equality holding for all eigenvalues iff A is diagonalisable.
(3) Proved that every complex matrix can be made upper triangular and hence, diagonalisable complex matrices are dense and the Cayley-Hamilton theorem holds.

2. More on diagonalisation or the lack of thereof

Before proceeding further, let us do an example of diagonalisation to clear confusion. The main point is that if P consists of the eigenvectors of A as its columns, then $A = PDP^{-1}$ or alternatively, $P^{-1}AP = D$. Let $A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$. Its eigenvalues and eigenvectors are $(5, (1, 1))$ and $(2, (2, -1))$. So let $P = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$. Then $P^{-1} = -\frac{1}{3}\begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix}$. Now $P^{-1}AP$ can be easily computed to be the diagonal matrix $D = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$.

An interesting question is "When can two (or more) matrices be simultaneously diagonalised ?" Note that if $PA_1P^{-1} = D_1$ and $PA_2P^{-1} = D_2$, then $[A_1, A_2] = A_1A_2 - A_2A_1 = P^{-1}[D_1, D_2]P = 0$. (The symbol $[A, B] = AB - BA$ is called the commutator of A and B. It measures how far they are from commuting. It is the central theme of "Lie algebras".) In fact, this necessary condition applies to any number of such matrices. This condition is also sufficient :

Theorem 2.1. *If $A_i : V \to V$ (where V is finite-dimensional as usual) are diagonalisable over $\mathbb{F}$, and if $[A_i, A_j] = 0\ \forall\ i \neq j$, then there exists a basis such that all the A_i are simultaneously diagonal in this basis.*

Proof. Without loss of generality, assume that at least one of the A_i (call it A_{i_0}) is not a multiple of identity. We induct on $n = dim(V)$. For $n = 1$ it is trivial. Assume it is true for $1, 2 \ldots, n-1$. Let e_1 be an eigenvector of A_{i_0} for some i_0. Then $A_{i_0}A_je_1 = A_jA_{i_0}e_1 = \lambda_0 A_je_1$ for all j. So either $A_je_1 = 0$ (which means e_1 is an eigenvector of A_j) or A_je_1 is an eigenvector of A_{i_0} with eigenvalues $\lambda_{0,l}$.

Let $V_{0,l} \subset V$ be the eigenspace of $\lambda_{0,l}$ (for A_{i_0}). It is an invariant subspace of A_j for all j. In other words, every eigenspace of A_{i_0} is an invariant subspace for all A_j. Since A_{i_0} is diagonalisable, $V \equiv \oplus_k V_{0,l}$ (why ?) If we prove that all the $A_j : \oplus_{k\neq 1} V_{0,l} \to \oplus_{k\neq l} V_{0,l}$ are diagonalisable, by the induction hypothesis, there exists a basis of $\oplus_{k\neq l} V_{0,l}$ such that the basis vectors are eigenvectors for all A_j. Likewise, by the induction hypothesis, the same can be done (by assumption that A_{i_0} is not a multiple of the identity) for $V_{0,1}$. Hence we get a basis of V consisting of eigenvectors for all A_j.

The only thing remaining is to show the following lemma. Now we have the following lemma.

Lemma 2.2. *If $T : V \to V$ is diagonalisable and $V_0 \subset V$ (of dimension > 0) is an invariant subspace of V, then $T : V_0 \to V_0$ is also diagonalisable.*

The proof of this lemma will be deferred to a later class where we will have proved that a linear map is diagonalisable iff the minimal polynomial has no repeated roots. □

In the proof of the above result, we had to put in a little effort to prove something seemingly obvious about diagonalisability (when restricted to an invariant subspace). So we need some better ways to check whether something is diagonalisable or not.

To this end, let $T : V \to V$ be a linear map (V is finite-dimensional) such that all its eigenvalues lie in $\mathbb{F}$. Let S be the set of all abstract polynomials $p : \mathbb{F} \to \mathbb{F}$ such that $p(T) = 0$ (when thought of as a polynomial function). This set is non-empty by the Cayley-Hamilton theorem. Actually, it is much easier to see it is non-empty : Indeed, since $L(V,V)$ has dimension n^2, the set $1, T, T^2, \ldots, T^{n^2}$ is linearly dependent and hence S is non-empty. Here is an important theorem.

Theorem 2.3. *Let k be the smallest degree of non-zero polynomials in S. There exists a unique monic (that is, the highest order coefficient is 1) polynomial $m(x)$of degree k in S. (This polynomial is called the minimal polynomial of T.) The minimal polynomial satisfies the following properties.*

1. *Every element of S is divisible by the minimal polynomial.*
2. *Every eigenvalue of T is a root of $m(x)$.*
3. *If T is diagonalisable then $m(x)$ has no repeated roots.*

Proof. If there are two monic minimal degree polynomials m_1, m_2, then $m_1(x) = m_2(x)q(x) + r(x)$ where $deg(r(x)) < deg(m_2(x)) = deg(m_1(x))$. (This fact is the Euclidean algorithm for polynomials over a field. It can be proven by induction.) Therefore, $m_1(T) = 0 = m_2(T)q(T) + r(T) = 0 + r(T)$ and hence $r(x) \in S$ contradicting the minimality of the degrees of m_1, m_2. This means that $r(x) = 0$. Comparing the highest order coefficients, $m_1(x) = m_2(x)$.
Now

1. The same Euclidean algorithm shows that p is divisible by m.
2. The Cayley-Hamilton theorem implies that the roots of m form a subset of the roots of the characteristic polynomial p_T. If any eigenvalue λ does not occur in $m(x)$, then $m(T)e = m(\lambda)e \neq 0$ where e is an eigenvector corresponding to λ. This is a contradiction.
3. In the basis of eigenvectors, T is diagonal and hence the polynomial $(T-\lambda_1)(T-\lambda_2)\ldots(T-\lambda_i) = 0$ is the minimal polynomial. Hence diagonalisability implies non-repeated roots of $m(x)$.

$\square$

Here are a few examples :

1. If $A \in Mat_{5\times 5}(\mathbb{C})$ such that $A^{100} = 0$, then $A^5 = 0$. Indeed, the minimal polynomial divides x^{100} and hence has only one root, namely, 0. So the characteristic polynomial is x^5. By Cayley-Hamilton, $A^5 = 0$.
2. Let $A = \begin{bmatrix} \lambda & 1 & 0 & 0 \ldots 0 \\ 0 & \lambda & 1 & 0 \ldots 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda \end{bmatrix}$, i.e., A is upper-triangular with λ on the diagonal and 1s on the off-diagonal. Clearly the only root of the minimal polynomial is λ. So $m(x) = (x-\lambda)^k$. Let $p_r(x) = (x-\lambda)^r = p_1^r$. Now $p_1(A)(e_i) = e_{i-1}$ if $i \geq 2$ and $p_1(A)(e_1) = 0$. Therefore, $p_r(A)(e_i) = e_{i-r}$ if $i \geq r+1$ and 0 otherwise. Thus, $r = n$ is the smallest number such that $p_r(A) = 0$. Thus $m(x) = (x-\lambda)^n$. So A is not diagonalisable.
3. Let $A \in Mat_{2\times 2}(\mathbb{C})$ satisfy $A^2 = A$. Then $A(A-1) = 0$ and hence $m(x)$ is either $x, x-1$ or $x(x-1)$. Thus $A = 0, I$ or is similar to $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

(4) $A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$. $p_A(t) = t^2(t^2 - 4)$. The minimal polynomial has to be $t^k(t^2 - 4)$ for $k \geq 1$. Upon calculation, $A(A^2 - 4) = 0$. Actually, it is easy to see that the rank of A is 2. So $dim(ker(A)) = 2$. Hence A is diagonalisable over $\mathbb{Q}$.