NOTES FOR 22 OCT (TUESDAY)

1. Recap

(1) Almost proved that a family of commuting diagonalisable matrices is diagonalisable, except for a lemma that if $T : V \to V$ is diagonalisable, and V_0 is an invariant subspace, then $T : V_0 \to V_0$ is also diagonalisable. Here is a simple proof of this lemma : Let $v_1, \ldots, v_k$ be a basis of V_0 and let $e_1, \ldots, e_n$ be an eigenvector basis of V. Then $v_i = \sum_j c_{ij} e_j$. Thus, $T^k v_i \in V_0$ and $T^k v_i = \sum_j c_{ij} \lambda_j^k e_j$. Using the Vandermonde determinant we see that a linear combination of eigenvectors in each eigenspace such that $c_{ij} \neq 0$ is in V_0. In other words, $V_0 = \oplus_k V_0 \cap V_{\lambda_k}$ where V_λ is the eigenspace corresponding to λ. Hence $T : V_0 \to V_0$ is diagonalisable.

(2) Defined the minimal polynomial and proved its properties.

2. Erratum

In one of the previous lectures (long ago) I claimed to have proven that for a characteristic 0 field, polynomial functions and abstract polynomials are one and the same thing. I used the notion of the derivative to prove this statement. Unfortunately, unless the field is $\mathbb{R}$ or $\mathbb{C}$, the definition of the derivative itself uses the fact we are trying to prove. So that proof only works for $\mathbb{R}$ and $\mathbb{C}$. Instead, here is a different proof (that works for all infinite fields). Indeed, let $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ be a polynomial function that is identically zero. We will prove that $a_i = 0 \ \forall\ i$, thus proving that polynomial functions uniquely determine the abstract polynomial from whence they came. Indeed, substitute $n+1$ distinct x_i into $f(x)$ (which we are allowed to do because the field is infinite). Solve for the a_i using the Vandermonde determinant. We are done.

3. The Jordan Canonical Form

When a matrix is not diagonalisable, we saw that it is similar to an upper-triangular matrix. However, this form is not unique (it is dependent on many choices). We ideally want a standard or "canonical" form that is more or less unique. The answer is in the form of the Jordan Canonical Form.

Theorem 3.1. *Let $T : V \to V$ be an operator between finite-dimensional vector spaces. Assume that all the eigenvalues of T are in $\mathbb{F}$. Then there exist unique invariant subspaces V_i such that $V = \oplus_i V_i$ and there is a basis of V obtained through bases of V_i such that T is of the Jordan Canonical Form in this basis, i.e., T is upper-triangular, with the only possible super-diagonal elements being* 1. *The V_i are direct sums of Jordan blocks where a Jordan block looks like* $\begin{bmatrix} \lambda & 1 & 0 & \ldots \\ 0 & \lambda & 1 & \ldots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda \end{bmatrix}$. *(The Jordan Blocks are unique up to permutation.)*

Before we prove the theorem, we define a useful notion : Let $T : V \to V$ be an operator. $v \in V$ is said to be a generalised eigenvector of T with generalised eigenvalue λ if there exists a positive integer k such that $(T - \lambda I)^k v = 0$. The smallest such k is called the order/rank of λ. Note that a generalised eigenvalue is a root of the characteristic polynomial. Moreover, generalised eigenvectors with different generalised eigenvalues are linearly independent : In fact, more generally, if $W_i = ker((T-\lambda_i)^{w_i})$ and likewise for j, where $w_i, w_j \geq 1$, then clearly W_i, W_j are invariant subspaces. Hence, $(T-\lambda_i)^{w_i} : W_i \to W_i$ is 0. Hence, the only eigenvalues of $T : W_i \to W_i$ is λ_i and likewise, $T : W_j \to W_j$ is λ_j. Now $W_i \cap W_j$ is also an invariant subspace. By comparing the eigenvalues, we see that it is trivial.

Also, given a generalised eigenvector v of generalised eigenvalue λ and order k, consider $v, (T-\lambda)v, \ldots, (T-\lambda)^{k-1}v$. These are all generalised eigenvectors of λ that are linearly independent. Indeed, let us induct on k. $k = 1$ is trivial. Assuming the induction hypothesis, let $\sum_i c_i(T-\lambda)^i v = 0$. Applying $(T-\lambda)^k$ to both sides we see that $c_0 = 0$. Using $v_1 = (T-\lambda)v$ and induction we are done. The set of vectors $v, (T-\lambda)v, \ldots$ is called a Jordan chain generated by v. Note that the subspace spanned by these vectors is an invariant subspace and in this basis, T is a Jordan block in this subspace with diagonal elements being λ. Now we prove the theorem above. Note that the theorem above actually asserts that there exist linearly independent vectors in V_i whose Jordan chains span V_i. The following proof is from Artin's book.

Proof. Choose an eigenvalue λ of T. Now $T - \lambda I$ will be in the JCF iff T is so in that basis. So we can assume wLog that 0 is an eigenvalue of T.

Define $K_i = ker(T^i)$ and $R_i = Ran(T^i)$. Then $K_1 \subset K_2 \ldots$ and $\ldots \subset R_2 \subset R_1$. Thus, there is an m (by finite-dimensionality) such that $K = K_m = K_{m+1} \ldots$ and $R = R_m = \ldots$. Clearly K, R are invariant subspaces. We claim that $K \cap R = \{0\}$ and hence by nullity-rank, $V = K \oplus U$. Indeed, if $z \in K \cap R$, then $T^m z = 0$ and $z = T^m v$. Thus, $T^{2m} v = 0$ which means that $v \in K$ and hence $T^m v = z = 0$. Therefore, by induction hypothesis, we can bring $T : R \to R$ to the JCF by a basis. However, we still have to prove the same for $T : K \to K$ because R can be zero. Def : A nilpotent operator is one such that $T^r = 0$ for some r. We have reduced our theorem's proof to nilpotent operators. So assume from now on that T is nilpotent.

To be cont'd.... □