6. Elementary Canonical Forms

How to characterize a transformation?

6.1. Introduction

- Diagonal transformations are easiest to understand.

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

- This involves the studying "dynamical properties" of the operators.

Elementary canonical forms

- T in L(V,V). Classify up to conjugations.
  - What is the behavior of T? (dynamical)
    - Invariant subspaces
    - Direct sum decompositions
    - Primary decompositions
  - Diagonalizable
  - Jordan canonical form

Characteristic values

- Definition: V a vector space over F. T:V->V. A characteristic (eigen-) value of T is a scalar c in F s.t. there is a nonzero vector *a* in V with *Ta = ca.*
- This measures how much T stretches or contracts objects in certain directions.
- *a* is said to be the characteristic (eigen-) vector of T.

- Characteristic space {a in V| Ta = ca} for a fixed c in F.
- This is a solution space of equation (T-cI)a=0. Equals null(T-cI).
- Theorem 1. V finite dim over F. TFAE:
  - (i) c is a characteristic value of T.
  - (ii) T-cI is singular
  - (iii) det(T-cI) = 0.
- We now consider matrix of T:

- B a basis of V. A the nxn-matrix $A=[T]_B$.
  - T-cI is invertible <-> A-cI is invertible.
- Definition: A nxn matrix over F. A characteristic value of A in F is c in F s.t. A-cI is singular.
- Define f(x) = det (xI - A) characteristic polynomial.
- c s.t. f(c)=0 (zeros of f) <-> (one-to-one) characteristic value of f.

- Lemma: Similar (conjugate) matrices have the same characteristic values.
- Proof: $B=P^{-1}AP$.
  - $\det(xI-B)=\det(xI-P^{-1}AP)$
    $= \det(P^{-1}(xI-A)P)=\det P^{-1}\det(xI-A)\det P$.
    $= \det(xI-A)$
- Remark: Thus given T, we can use any basis B and obtain the same characteristic polynomial and values.
-

- Diagonalizable operators:
- T is diagonalizable <-> There exist a basis of V where each vector is a characteristic vector of T.

$$\begin{aligned} \mathcal{B} &= \{\alpha_1, \ldots, \alpha_n\} \\ T\alpha_i &= \lambda_i \alpha_i \end{aligned}$$

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

- Fact: If T is diagonalizable, then $f_T(x)$ factors completely.
- Proof: T=

$$T=\begin{bmatrix} cI_{fxf} & & & & \\ & dI_{gxg} & & & \\ & & eI_{hxh} & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}$$

$$xI-T=\begin{bmatrix} xI-cI_{fxf} & & & & \\ & xI-dI_{gxg} & & & \\ & & xI-eI_{hxh} & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix}$$

- det(xI-T) =det(x-c)I_{fxf} det(x-d)I_{gxg}…

  = $(x-c)^f(x-d)^g$….

- Nondiagonalizable matrices exist:
  - $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, f(x)=(x-1)^2.
    - 1 is the only characteristic value.
    - If A is diagonalizable, then A can be written as I in some coordinate. Thus A=I. Contradiction.
  - There are many examples like this. In fact, all nondiagonalizable matrices are similar to this example. (Always, with repeated or complex eigenvalues.

- Lemma. Ta=ca -> f(T)a = f(c)a, f in F[x].
- Proof: $f=a_n x^n + a_{n-1}x^{n-1}+\ldots+a_1 x+a_0$
  - $f(T)=a_n T^n + a_{n-1}T^{n-1}+\ldots+a_1 T+a_0 I$
  - $f(T)(a)= a_n T^n(a)+ a_{n-1}T^{n-1}(a)+\ldots+a_1 T(a)+a_0(a)$
  - $=a_n c^n a+ a_{n-1}c^{n-1}a+\ldots+a_1 ca+a_0 a$
  - $= (a_n c^n+ a_{n-1}c^{n-1}+\ldots+a_1 c+a_0)a$
  - =f(c)a.
- Lemma. T linear operator on the f.d. space V. $c_1,\ldots c_k$ distinct characteristic values of T. $W_1,\ldots,W_k$ respective characteristic spaces

  If $W= W_1+\ldots+W_k$, then $\dim W = \dim W_1+\ldots+\dim W_k$ (i.e., independent).

  If B_i basis, then $\{B_1,\ldots,B_k\}$ is a basis of W.

- Proof: $W = W_1 + \ldots + W_k$
  - dim $W \leq \dim W_1 + \ldots + \dim W_k$ in general
  - We prove independence first:
    - Suppose $b_1 + \ldots + b_k = 0$, b_i in W_i. $Tb_i = c_i b_i$.
    - $0 = f(T)(0) = f(T) b_1 + \ldots + f(T) b_k = f(c_1) b_1 + \ldots + f(c_k) b_k$
    - Choose f_i in F[x] so that $f_i(c_j) = 1$ $(i=j)$ 0 $(i \neq j)$ (from Lagrange)
    - $0 = f_i(T)(0) = f_i(T) b_1 + \ldots + f_i(T) b_k = f_i(c_1) b_1 + \ldots + f_i(c_k) b_k = b_i$.
  - B_i basis. Let $B = \{B_1, \ldots, B_k\}$
    - B spans W.
    - B is linearly independent:
    - If $$\sum c_1^i B_1^i + \sum c_2^i B_2^i + \cdots + \sum c_k^i B_k^i = 0$$

- If not all $c^j_i=0$, then we have $b_1+\ldots+b_k=0$ for some b_is. However, $b_i=0$ as above. This is a contradiction. Thus all $c^j_i=0$.

- Theorem 2. $T:V^n \to V^n$. $c_1,\ldots c_k$ distinct characteristic values of T. $W_i=\text{null}(T-c_i I)$.

TFAE.

1. T is diagonalizable.
2. $f_T = (x-c_1)^{d1}\ldots. (x-c_k)^{dk}$. $\dim W_i=d_i$.
3. $\dim V = \dim W_1+\ldots+\dim W_k$

- Proof: (i)->(ii) done already
  - (ii)->(iii). $d_1+\ldots.+d_k = \deg f_T = n$.
  - (iii)->(i) $W= W_1+\ldots+W_k$. W is a subspace of V.
  - $\dim V = \dim W \to V = W$.
  - $V= W_1+\ldots+W_k$. V is spanned by characteristic vectors and hence T is diagonalizable.

- Lesson here: Algorithm for diagonalizability:
  - Method 1: Determine W_i-> dim W_i -> sum d_is -> equal to dim V -> yes: diagonalizable. no: nondiagonalizable.
  - Method 2: Find characteristic polynomial of f.
    - Completely factored?: -> no: not diagonalizable.
    - -> yes:d_i factor degree-> compute W_i. -> d_i=dim W_i? -> no: not diagonalizable. yes: check for all i.
- Usually, a small perturbations makes nondiagonalizable matrix into diagonalizable matrix if F=C.