0.1 Построение Жорданова базиса

Для характеристического многочлена справедливо разложение:

$$\frac{1}{P_n(\lambda)} = \frac{1}{(\lambda-\lambda_1)^{k_1}...(\lambda-\lambda_m)^{k_m}} = \sum_{i=1}^{m}\sum_{l=1}^{k_1}\frac{A_l^i}{(\lambda-\lambda_i)^l},\, A_l^i \in R^m$$

После сложения по внутренней сумме:

$$\frac{1}{P_n(\lambda)} = \frac{1}{(\lambda-\lambda_1)^{k_1}...(\lambda-\lambda_m)^{k_m}} = \frac{f_1(\lambda)}{(\lambda-\lambda_1)^{k_1}} + ... + \frac{f_s(\lambda)}{(\lambda-\lambda_s)^{k_s}} + ... + \frac{f_m(\lambda)}{(\lambda-\lambda_m)^{k_m}}$$

где $f_s(\lambda)-$ многочлен степени не выше k_{s-1}, $s=\overline{1,m}$. Умножим на $P_n(\lambda)$:

$$1 = Q_1(\lambda) + ... + Q_m(\lambda)$$

$$Q_s(\lambda) = f_s(\lambda)\cdot\frac{P_n(\lambda)}{(\lambda-\lambda_s)^{k_s}} = f_s(\lambda)\cdot(\lambda-\lambda_1)^{k_1}\cdot...\cdot(\lambda-\lambda_{s-1})^{k_{s-1}}\cdot(\lambda-\lambda_{s+1})^{k_{s+1}}\cdot...\cdot(\lambda-\lambda_m)^{k_m} \quad (1)$$

Рассмотрим множество квадратных матриц одного порядка. Это множество является ассоциативным кольцом с единицей, поэтому

$$A^n\cdot A^m = A^{n+m} = A^m\cdot A^n;\ A^0 \overset{def}{=} E$$

Определены коммутативное и ассоциативное сложение матриц. Нулевую матрицу примем за ноль. Согласно свойствам умножения матриц на числа:

$$A^k\cdot\alpha = \alpha A^k,\, \alpha A^k + \beta A^k = (\alpha+\beta)A^k$$

Таким образом правила приведения подобных членов аналогично правилу для многочленов.

$$A^k + (-1\cdot A^k) = A^k + (-A^k) = 0$$

В качестве символа x в определении многочлена можно взять квадратную матрицу A и получить множество матричных многочленов $\{P_n(A)\}$

$$P_n(A) = a_0E + a_1A + ... + a_nA^n$$

На множестве $\{P_n(A)\}$ сложение и умножение определяются как обычные матричные действия, поэтому $\{P_n(A)\}$ является кольцом.

1. $P_n(A) + P_m(A) = P_m(A) + P_n(A)$
2. $(P_n(A) + P_m(A)) + P_s(A) = P_n(A) + (P_m(A) + P_s(A))$
3. $P_n(A)\cdot P_m(A) = P_m(A)\cdot P_n(A)$
4. $(P_n(A)\cdot P_m(A))\cdot P_s(A) = P_n(A)\cdot(P_m(A)\cdot P_s(A))$
5. $P_n(A)\cdot(P_m(A) + P_s(A)) = P_n(A)\cdot P_m(A) + P_n(A)\cdot P_s(A)$

За ноль в этом множестве принимается нулевая матрица.

Определение 0.1. *Отображение φ кольца K на кольцо K' называется гомоморфизмом, если $\forall a \in K$, $\forall b \in K$:*

$$\varphi(a+b) = \varphi(a) + \varphi(b);\ \varphi(ab) = \varphi(a)\cdot\varphi(b)$$

В отличие от изоморфизма гомоморфизм не обязательно является взаимно однозначным отображением, т.е. не предполагается, что образы K заполняют все кольцо K', и различным элементам из K соответствуют разные элементы из K'.

В силу определения множеств $\{P_n(A)\}$ и $\{P_n(\lambda)\}$, кольца $\{P_n(A)\}$ и $\{P_n(\lambda)\}$ гомоморфны:

$$\varphi : \varphi(P_n(\lambda)) \longrightarrow P_n(A)$$

Неоднозначность отображения φ возникает в силу того, что существуют такие квадратные матрицы $A \neq 0 : \exists n \in N : A^m = 0 \ \forall m \geq n$.

Теорема 0.1 (Гамильтона-Кэли). *Пусть $P_n(\lambda)-$ характерестический многочлен матрицы A, тогда $P_n(A) = 0$.*

В силу построения гомоморфизма между $\{P_n(A)\}$ и $\{P_n(\lambda)\}$ имеет место разложение:

$$P_n(A) = A^n + a_1 \cdot A^{n-1} + ... + a_o \cdot E = (A - \lambda_1 E)^{k_1} \cdot ... \cdot (A - \lambda_m E)^{k_m}$$

где $\lambda_1, ..., \lambda_m-$ корни $P_n(A)$.

Подействуем гомоморфизмом φ на (1) :

$$E = Q_1(A) + ... + Q_m(A)$$
$$Q_s(A) = f_s(A) \cdot (A - \lambda_1 E)^{k_1} \cdot ... \cdot (A - \lambda_{s-1})^{k_{s-1}} \cdot (A - \lambda_{s+1})^{k_{s+1}} \cdot ... \cdot (A - \lambda_m)^{k_m} \qquad (2)$$
$$Q_s(A) - \text{линейные преобразования}$$

Порядок сомножетелей в (2) не важен, т.к. матрицы $(A - \lambda_s E)$ такого вида перестоновочны между собой.

Рассмотрим $Q_i(A)$. Покажем, что $\forall i, \ j = \overline{1{,}m} \longmapsto$

$$Q_i(A) \cdot Q_j(A) = \begin{cases} 0, i \neq j \\ Q_i^2, i = j \end{cases} \quad \text{и } Q_i(A) = Q_i^2(A) \qquad (3)$$

Доказательство. $Q_i(A) \cdot Q_j(A) = f_i(A) \cdot f_j(A) \cdot (A - \lambda_1 E)^{k_1} \cdot ... \cdot (A - \lambda_{i-1} E)^{k_{i-1}} \cdot (A - \lambda_{i-1} E)^{k_{i+1}} \cdot ... \cdot (A - \lambda_m E)^{k_m} \cdot (A - \lambda_1 E)^{k_1} \cdot ... \cdot (A - \lambda_{j-1} E)^{k_{j-1}} \cdot (A - \lambda_{j+1} E)^{k_{j+1}} \cdot ... \cdot (A - \lambda_m E)^{k_m} = M(A) \cdot P_n(A) =$ (Теорема Гамильтона-Кэли) $= 0$

В силу (2):

$$\vec{x} = E\vec{x} = Q_1(\vec{x}) + ... + Q_i(\vec{x}) + ... + Q_m(\vec{x})$$
$$\Rightarrow Q_i(\vec{x}) = (Q_i Q_1)(\vec{x}) + ... + (Q_i^2)(\vec{x}) + ... + (Q_i Q_n)(\vec{x}) = Q_i^2(\vec{x})$$

■

Пусть $R_i = Im Q_i(A), \ i = \overline{1{,}m}$ – образ $Q_i(A)$. Из (3) следует, что R_i- инвариантное подпространство A. Тогда, если $\vec{x} \in R_i \rightarrow \exists \vec{y} \in A, Q_i(\vec{y}) = \vec{x}$, то $A(\vec{x}) = A(Q_i(\vec{y})) = (A \cdot Q_i)(\vec{y}) = (Q_i A)(\vec{y}) = Q_i(A(\vec{y})) \in R_i-$ инвариантное подпространство.

При доказательстве (3) было получено, что:

$$\vec{x} = E\vec{x} = Q_1(\vec{x}) + ... + Q_i(\vec{x}) + ... + Q_m(\vec{x}) = \vec{x_1} + ... + \vec{x_i} + ... + \vec{x_m} \qquad (4)$$

где $x_i = Q_i(\vec{x}) \in R_i, \ i = \overline{1{,}m}$.

(3) означает, что R^n является суммой подпространств R_i. Покажем, что такое разложение единственно:

Доказательство. Предположим, что хотя бы для одного $k = \overline{1{,}m} \ \exists \vec{y_k} = Q_k(z_k) \neq \vec{x_k}$: $\vec{x} = \sum\limits_{k=1}^{m} Q_k(\vec{z_k}) = \vec{y_1} + ... + \vec{y_i} + ... + \vec{y_m}$. Тогда $Q_i(\vec{x}) = \vec{x_i} = Q_i\left(\sum\limits_{k=1}^{m} Q_k(\vec{z_k})\right) \overset{Th \text{ Г.К.}}{=} Q_i^2(\vec{z_i}) = Q_i(\vec{z_i}) = \vec{y_i} \Rightarrow \vec{x_i} = \vec{y_i}$ ■

Т.к. единственное разложение эквивалентно тому, что сумма подпространств прямая, то:

$$\vec{R^n} = R_1 \oplus R_2 \oplus ... \oplus R_m$$

Тогда A в таком базисе будет иметь вид:

$$\begin{Vmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_m \end{Vmatrix} \quad (5)$$

Подпространства R_i называются корневыми подпространствами $\vec{R^n}$.

Теорема 0.2. $\forall s = \overline{1,m} : R_s = ker(A - \lambda_s E)^{k_s}$ $\forall \vec{x} \in R_i \longmapsto (A - \lambda_i E)^{k_i}\vec{x} = 0$

Доказательство. Пусть $\vec{x} \in R_s \Rightarrow \exists \vec{y} \in R_s : \vec{x} = Q_i(\vec{y})$ в силу инвариантности R_s. Тогда $(A - \lambda_s E)^{k_s}\vec{x} = (A - \lambda_s E)^{k_s} \cdot f_s(A) \cdot (A - \lambda_1 E)^{k_1} \cdot ... \cdot (A - \lambda_{s-1}E)^{k_{s-1}} \cdot (A - \lambda_{s+1}E)^{k_{s+1}} \cdot (A - \lambda_m E)^{k_m}\vec{y} = f_s(A) \cdot P_n(A)\vec{y} = 0 \Rightarrow R_s \subseteq ker(A - \lambda_s E)^{k_s}$.

Пусть $\vec{x} \in ker(A - \lambda_s E)^{k_s}$. Тогда $\forall j \neq s : Q_j(\vec{x}) = 0$, поскольку множитель $(A - \lambda_s E)^{k_s}$ как множитель входит в представление Q_j. Поэтому из (4) в этом случае: $\vec{x} = 0 + ... + Q_s(\vec{x}) + ... + 0 \Rightarrow \vec{x} \in R_s \Rightarrow ker(A - \lambda_s E)^{k_s} \subseteq R_s$ ■

Рассмотрим структуру корневого подпространства. Покажем, что

$$dim(R_s = ker(A - \lambda_s E)^{k_s}) = k_s$$

Лемма 0.1. *Пусть B является линейным преобразованием $\vec{R^n}$ и $R = ker(B^l)$, $n < l$. Тогда, если $\exists \vec{x} \in R : B^{l-1}\vec{x} \neq 0$, то $dimR \geq l$.*

Доказательство. Рассмотрим систему векторов $\vec{x}, B\vec{x}, ..., B^{l-1}\vec{x} \in R$. Ни один из векторов этой системы не равен нулю. Покажем, что эта система линейно независима. С этой целью рассмотрим нулевую линейную комбинацию этих векторов.

$$a_0\vec{x} + a_1(B\vec{x}) + ... + a_{n-1}(B^{l-1}\vec{x}) = 0 \quad (6)$$

Подействуем последовательно $l - 1$ раз преобразованием B на (6):

$$\begin{cases} a_0(B\vec{x}) + a_1(B^2\vec{x}) + ... + a_{n-2}(B^{l-1}\vec{x}) = 0 \\ ... \\ a_0(B^{l-2}\vec{x}) + a_1(B^{l-1}\vec{x}) + 0 + ... + 0 = 0 \\ a_0(B^{l-1}\vec{x}) = 0 \end{cases}$$

$(B^{l-1}\vec{x}) \neq 0$ по условию $\Rightarrow a_0 = a_1 = ... = a_{l-1} = 0 \Rightarrow$ Вектора ЛНЗ

Таким образом в R лежит как минимум l ЛНЗ векторов, а значит базис в R не может содержать меньше, чем l векторов $\Rightarrow dimR \geq l$.

Было доказано, что пространства R_i, $i = \overline{1,s}$ образуют прямую сумму, равную $\vec{R^n}$, поэтому размерность $\vec{R^n}$ является суммой размерностей подпространств, которые составляют эту прямую сумму. Т.к. $k_1 + k_2 + ... + k_s = n$, то $\forall i \longmapsto dimR_i = k_i$, поскольку если $\exists j : dimR_j > k_j$, то тогда должно существовать R_i, у которого размерность меньше, чем k_i, что в силу леммы невозможно. ■