Diagonalization

Now suppose in $\mu_A(\theta) = (\theta-\lambda_1)^{p_1}\ldots(\theta-\lambda_k)^{p_k}$ we have $p_j = 1$ for $1 \le j \le k$ *(which means then there are only linear factors in the minimal polynomial)* then there are exactly k distinct eigenvalues, with n_j linearly independent eigenvectors for the eigenvalue λ_j, $n_1+\ldots+n_k = n$, dim $\ker(A-\lambda_j Id) = n_j$; these $n_j's$ being exactly the ones obtained in $x_A(\theta) = (\theta-\lambda_1)^{n_1}\ldots(\theta-\lambda_k)^{n_k}$, 1 and $X = \oplus_{j=1}^{r}\ker(A-\lambda_j Id)$.

Since there are $n_1+\ldots+n_k = n$ linearly independent eigen-vectors, let us arrange them in an ordered list $\{x_1,\ldots,x_n\}$ obtaining $Ax_i = \lambda_i x_i$, $1 \le i \le n$.

Then these n vectors being linearly independent form a basis x and forming the matrix

$P = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}_{n\times n}$ *($\because$ each x_i is a column vector)*

we can write this as

$$aP = PD \ (D = \mathtt{diag}(\lambda_1,\ldots,\lambda_n))$$

where a is the matrix for A to start with *(in some basis b)*. But P is invertible *(Since its columns are linearly independent)* and thus $a = PDP^{-1}$ i.e. $D = P^{-1}aP$.

This means that by a change of basis, the matrix a has been represented in some new basis as a diagonal matrix; put differently, there is some basis with respect to which we can have a diagonal matrix representing A, this diagonal matrix has the eigen values λ_i on its columns, each λ_i repeated as many times as it occurs in $x_A(\theta)A$. One says that A is diagonalizable, the matrix P being known as the modal matrix and D being known as the spectral matrix for the operator A or the matrix a.

Examples

Example 1. Take some polynomial $a(\theta) = \theta^n + a_{n-1}\theta^{n-1}+\ldots+a_1\theta+a_0$ of degree n*(if we do have some $a_n \ne 1$, just divide by a_n which is possible since a_n must be non-zero with deg $(a(\theta)) = n$).*

The matrix

$$\begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{n-1} \end{bmatrix}$$

is called the companion matrix of $a(\theta)$. Then

$$\begin{aligned} det(\lambda I - C_a) &= \begin{vmatrix} \lambda & -1 & 0 & \ldots & 0 \\ 0 & \lambda & -1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -1 \\ a_0 & a_1 & a_2 & \ldots & \lambda + a_{n-1} \end{vmatrix} \\ &= \begin{vmatrix} 0 & -1 & 0 & \ldots & 0 \\ 0 & \lambda & -1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -1 \\ a(\lambda) & a_1 & a_2 & \ldots & \lambda + a_{n-1} \end{vmatrix} \\ & \quad \textit{(Adding } \lambda \textit{ times the 2nd column,} \lambda^2 \textit{ times the 3rd column,} \\ & \quad \lambda^{n-1} \textit{ times the last column to the first column)} \\ &= (-1)^{n+1} a(\lambda)(-1)^{n-1} = a(\lambda) (\textit{ expanded in terms of the first column)} \end{aligned}$$

So that $x(\theta) = a(\theta)$ *(By problem 6 of Module 3)*. We also have $\mu(\theta) = a(\theta)$ *($\because$ the minor of a_0 in $(\lambda I - C_a)$ is $(-1)^{n-1}$ and thus d_{n-1}, the gcd of all minors of order $(n-1)$ in $(\lambda I - C_a)$ is 1.)*

If λ_i is an eigenvalue, we have $x(\lambda_i) = 0$ hence $(\lambda_i)^n = -[a_{n-1}(\lambda_i)^{n-1} + \ldots + a_1\lambda_i + a_0]$ and thus, for $x_i = \begin{bmatrix} 1 \\ \lambda_i \\ \vdots \\ (\lambda_i)^{n-1} \end{bmatrix}$ we find $C_a x_i = \lambda_i x_i == \begin{bmatrix} \lambda_i \\ \lambda_i^2 \\ \vdots \\ (\lambda_i)^{n-1} \end{bmatrix}$.

Verifying that x_i is an eigenvector or corresponding to λ_i. We then form the matrix $V = [x_1 \ldots x_n]_{n \times n}$ so that $C_a V = V \text{ diag}(\lambda_1, \ldots, \lambda_n)$.

Now

$$V = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-2} & \dots & \lambda_n^{n-1} \end{bmatrix}$$

which is **Vandermonde matrix** with determinant $\prod_{1\le i<j\le n}(\lambda_j - \lambda_i)$. If all the λ_i are distinct, this determinant is not zero, the modal matrix V is invertible, and we have $V^{-1}C_aV = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. So that $C = C_a$ is diagonalizable.

Example 2. For a complex vector equipped with an inner product space, we ask additional question: given $X \xrightarrow{A} X$, dim $X = n < \infty$, do we have an orthonormal set of n eigenvectors of A?. If so, we call it a normal matrix. Then if $\{e_1, \dots, e_n\}$ happens to be the orthonormal set with n eigenvectors, $\{e_1, \dots, e_n\}$ is a basis and thus a normal matrix is certainly diagonalizable. Further, if $U = \begin{bmatrix} e_1 & \dots & e_n \end{bmatrix}$ (the $n \times n$ matrix with e_i as columns), we can write

$$a = UDU^{-1}$$