State the methods of sovling linear recurrence relations in the language of linear algebra

Hanbin Liu 11912410

Consider an arbitrary linear homogeneous relation of degree k with constant coefficients:(The value of $a_1, a_2, ..., a_k$ are given)

$$a_n = \sum_{i=1}^{k} c_i a_{n-i}$$

We define a colume vector of length k: $\eta_n = (a_n, a_{n-1}, ..., a_{n-k+1})^T$. In particular, $\eta_k = (a_k, a_{k-1}, ..., a_1)^T$ is known since a_i are given.($i = 1, 2, ..., k$)

Then, the linear homogeneous relation can be written as matrix multiplication: $\eta_n = A\eta_{n-1}$

i.e.

$$\begin{pmatrix} a_n \\ a_{n-1} \\ ... \\ a_{n-k+1} \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & ... & c_k \\ 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \\ ... \\ a_{n-k} \end{pmatrix} \tag{1}$$

where $A =$

$$\begin{pmatrix} c_1 & c_2 & ... & c_k \\ 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & 1 & 0 \end{pmatrix}$$

This is called Companion matrix or Forbenius matrix. One of its properties is that Companion matrix has the same characterisic polynomial and minimal polynomial.(The minimal polynomial is a monic annihilating polynomial of the matrix whose degree is the lowest.) Now we prove this conclusion. Note that the characteristic polynomial of A is $det(\lambda I - A) = \varphi(\lambda) = \lambda^k - c_1\lambda^{k-1} -$

$c_2\lambda^{k-2} - ... - c_k$. A also can be written as

$$\begin{pmatrix} 0 & 0 & ... & c_k \\ 1 & 0 & ... & c_{k-1} \\ 0 & 1 & ... & ... \\ 0 & ... & 1 & c_1 \end{pmatrix}$$

Suppose $\{\xi_1, \xi_2, ..., \xi_k\}$ is an orthonormal basis and $(A\xi_1, A\xi_2, ..., A\xi_k) = (\xi_1, \xi_2, ..., \xi_k)A$. Note that $\xi_2 = A\xi_1, \xi_3 = A\xi_2 = A^2\xi_1, ..., \xi_k = A\xi_{k-1} = ... = A^{k-1}\xi_1$ and $A\xi_k = c_k\xi_1 + c_{k-1}\xi_2 + ... + c_1\xi_k = c_k\xi_1 + c_{k-1}A\xi_1 + c_{k-2}A^2\xi_1 + ... + c_1A^{k-1}\xi_1$. *i.e.* $A^k\xi_1 = c_k\xi_1 + c_{k-1}A\xi_1 + c_{k-2}A^2\xi_1 + ... + c_1A^{k-1}\xi_1$. Thus, $\varphi(A)\xi_1 = A^k\xi_1 - c_1A^{k-1}\xi_1 - ... - c_k\xi_1 = A^k\xi_1 - (c_1A^{k-1}\xi_1 + c_2A^{k-2}\xi_1 + ... + c_k\xi_1) = A^k\xi_1 - A^k\xi_1 = 0$. For $i = 2, 3, ..., k$, $\varphi(A)\xi_i = \varphi(A)A^{i-1}\xi_1 = A^{i-1}\varphi(A)\xi_1 = A^{i-1}0 = 0$. Therefore, $\varphi(A) = 0$ since $\varphi(A)\xi_i = 0$ for $i = 1, 2, ..., k$ and $\{\xi_1, \xi_2, ..., \xi_k\}$ is a basis of the vector space. In other words, $\varphi(\lambda)$ is a monic annihilating polynomial of A. Suppose there exists another monic annihilating polynomial whose degree $< k : p(\lambda) = \lambda^m + b_{m-1}\lambda^{m-1} + ... + b_0$, then

$$0 = p(A)\xi_1 = A^m\xi_1 + b_{m-1}A^{m-1}\xi_1 + ... + b_0\xi_1 = \xi_{m+1} + b_{m-1}\xi_m + ... + b_0\xi_1 (m < k) \qquad (2)$$

This contradicts that $\xi_1, \xi_2, ..., \xi_k$ are linearly independent. Hence, $\varphi(\lambda)$ is the minimal polynomial of A.

Lemma

A matrix is diagonalizable if and only if its minimal polynomial has only simple roots.

Proof

A is diagonalizable $\Leftrightarrow$ its Jordan canonical form is diagonal $\Leftrightarrow$ each Jordan block is of size 1×1 $\Leftrightarrow$ each elementary divisor is of the form $\lambda - \lambda_i$ $\Leftrightarrow$ the minimal polynomial of A has only simple roots.

Therefore, by the property that Companion matrix has the same characteristic polynomial and minimal polynomial and by the lemma, there are two cases. One is that $\varphi(\lambda)$ has only simple roots and then A must be diagonalizable, and the other is that $\varphi(\lambda)$ has multiple roots and then A is not diagonalizable.

Case1: If $\varphi(\lambda)$ has only simple roots, then A is diagonalizable, and thus $A = P^{-1}DP$, where P is an invertible matrix and D is a diagonal matrix. In this way, we can easily calculate A^n. $A^n = (P^{-1}DP)^n = \underbrace{(P^{-1}DP)(P^{-1}DP)...(P^{-1}DP)}_{n} = \underbrace{P^{-1}DPP^{-1}DP...P^{-1}DP}_{n} = P^{-1}\underbrace{DD...D}_{n}P = P^{-1}D^nP$. A is diagonalizable, then $D = diag(\lambda_1, \lambda_2, ..., \lambda_k)$, where λ_i are the eigenvalues of A. By (1), we have that

$$
\begin{aligned}
\begin{pmatrix} a_n \\ a_{n-1} \\ ... \\ a_{n-k+1} \end{pmatrix} &= P^{-1} \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ 0 & 0 & ... & 0 \\ ... & ... & 0 & \lambda_k \end{pmatrix} P \begin{pmatrix} a_{n-1} \\ a_{n-2} \\ ... \\ a_{n-k} \end{pmatrix} \\
&= P^{-1} \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ 0 & 0 & ... & 0 \\ ... & ... & 0 & \lambda_k \end{pmatrix}^{n-k} P \begin{pmatrix} a_k \\ a_{k-1} \\ ... \\ a_1 \end{pmatrix} \\
&= P^{-1} \begin{pmatrix} \lambda_1^{n-k} & 0 & ... & 0 \\ 0 & \lambda_2^{n-k} & ... & 0 \\ 0 & 0 & ... & 0 \\ ... & ... & 0 & \lambda_k^{n-k} \end{pmatrix} P \begin{pmatrix} a_k \\ a_{k-1} \\ ... \\ a_1 \end{pmatrix}
\end{aligned} \tag{3}
$$

Therefore, we can obtain a_n by the equation (3) and matrix multiplication.(The essence of matrix multiplication is linear combination) That is

$$
\begin{aligned}
a_n &= \beta_1\lambda_1^{n-k} + \beta_2\lambda_2^{n-k} + ... + \beta_k\lambda_k^{n-k} \\
&= \alpha_1\lambda_1^n + \alpha_2\lambda_2^n + ... + \alpha_k\lambda_k^n
\end{aligned} \tag{4}
$$

where $\alpha_i = \frac{\beta_i}{\lambda_i^k}$ are all constants.$(i = 1, 2, ..., k)$ By the initial conditions,we can find out the value of $\alpha_1, \alpha_2, ..., \alpha_k$, and thus we obtain a formula for a_n.

Case2: If $\varphi(\lambda)$ has multiple roots, then A is not diagonalizable. But we can turn A into a Jordan canonical form. That is, there exists an invertible matrix P $s.t.$

$$
\begin{aligned}
A &= P^{-1}JP \\
&= P^{-1}\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & ... & \\ & & & J_t \end{pmatrix}P
\end{aligned}
\tag{5}
$$

where J_i $(i = 1, 2, ..., t)$ are Jordan blocks and each J_i is of the form:

$$
\begin{pmatrix}
\lambda_i & 1 & 0 & ... & 0 \\
0 & \lambda_i & 1 & ... & 0 \\
... & ... & ... & ... & ... \\
0 & ... & ... & \lambda_i & 1 \\
0 & ... & ... & 0 & \lambda_i
\end{pmatrix}_{m_i \times m_i}
$$

$(\lambda_i \neq \lambda_j$ if $i \neq j)$

We now show that $J_i^n =$

$$
\begin{pmatrix}
\lambda_i^n & C_n^1\lambda_i^{n-1} & C_n^2\lambda_i^{n-2} & ... & C_n^{m_i-1}\lambda_i^{n-(m_i-1)} \\
0 & \lambda_i^n & C_n^1\lambda_i^{n-1} & ... & C_n^{m_i-2}\lambda_i^{n-(m_i-2)} \\
... & ... & ... & ... & ... \\
0 & ... & ... & \lambda_i^n & C_n^1\lambda_i^{n-1} \\
0 & ... & ... & 0 & \lambda_i^n
\end{pmatrix}_{m_i \times m_i}
\tag{6}
$$

This can be done by induction. For $n = 1$, it is obvious. Suppose the equation is true for n-1, it then follows that

$$
\begin{aligned}
J_i^n &= J_i^{n-1} J_i \\
&= \begin{pmatrix} \lambda_i^{n-1} & C_{n-1}^1 \lambda_i^{n-2} & C_{n-1}^2 \lambda_i^{n-3} & ... & C_{n-1}^{m_i-1} \lambda_i^{n-m_i} \\ 0 & \lambda_i^{n-1} & C_{n-1}^1 \lambda_i^{n-2} & ... & C_{n-1}^{m_i-2} \lambda_i^{n-(m_i-1)} \\ ... & ... & ... & ... & ... \\ 0 & ... & ... & \lambda_i^{n-1} & C_{n-1}^1 \lambda_i^{n-2} \\ 0 & ... & ... & 0 & \lambda_i^{n-1} \end{pmatrix} \begin{pmatrix} \lambda_i & 1 & 0 & ... & 0 \\ 0 & \lambda_i & 1 & ... & 0 \\ ... & ... & ... & ... & ... \\ 0 & ... & ... & \lambda_i & 1 \\ 0 & ... & ... & 0 & \lambda_i \end{pmatrix} \\
&= \begin{pmatrix} \lambda_i^n & C_n^1 \lambda_i^{n-1} & C_n^2 \lambda_i^{n-2} & ... & C_n^{m_i-1} \lambda_i^{n-(m_i-1)} \\ 0 & \lambda_i^n & C_n^1 \lambda_i^{n-1} & ... & C_n^{m_i-2} \lambda_i^{n-(m_i-2)} \\ ... & ... & ... & ... & ... \\ 0 & ... & ... & \lambda_i^n & C_n^1 \lambda_i^{n-1} \\ 0 & ... & ... & 0 & \lambda_i^n \end{pmatrix}
\end{aligned} \tag{7}
$$

Note that we used $C_{n-1}^{m-1} + C_{n-1}^m = C_n^m$ in the matrix multiplication. Hence, the equation is true for n.

By mathematical induction, (6) is true for $n \in \mathbb{Z}^+$.

Since $C_n^m = \frac{n(n-1)(n-2)...(n-m+1)}{m!}$, we can rewrite it as $C_n^m = d_0 n^m + d_1 n^{m-1} + ... + d_{m-1} n$, where d_i are constants($i = 0, 1, ..., m-1$). Now, let's power both sides of the equation (5), we have that $A^n = (P^{-1} J P)^n = P^{-1} J^n P$. And $J^n =$

$$
\begin{pmatrix} J_1^n & & & \\ & J_2^n & & \\ & & ... & \\ & & & J_t^n \end{pmatrix}
$$

By (1), we have that

$$
\begin{aligned}
\begin{pmatrix} a_n \\ a_{n-1} \\ \dots \\ a_{n-k+1} \end{pmatrix} &= P^{-1} \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \dots & \\ & & & J_t \end{pmatrix} P \begin{pmatrix} a_{n-1} \\ a_{n-2} \\ \dots \\ a_{n-k} \end{pmatrix} \\
&= P^{-1} \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \dots & \\ & & & J_t \end{pmatrix}^{n-k} P \begin{pmatrix} a_k \\ a_{k-1} \\ \dots \\ a_1 \end{pmatrix} \\
&= P^{-1} \begin{pmatrix} J_1^{n-k} & & & \\ & J_2^{n-k} & & \\ & & \dots & \\ & & & J_t^{n-k} \end{pmatrix} P \begin{pmatrix} a_k \\ a_{k-1} \\ \dots \\ a_1 \end{pmatrix}
\end{aligned} \tag{8}
$$

Note that each J_i^{n-k} is of the form

$$
\begin{pmatrix}
\lambda_i^{n-k} & C_{n-k}^1 \lambda_i^{n-k-1} & C_{n-k}^2 \lambda_i^{n-k-2} & \dots & C_{n-k}^{m_i-1} \lambda_i^{n-k-(m_i-1)} \\
0 & \lambda_i^{n-k} & C_{n-k}^1 \lambda_i^{n-k-1} & \dots & C_{n-k}^{m_i-2} \lambda_i^{n-k-(m_i-2)} \\
\dots & \dots & \dots & \dots & \dots \\
0 & \dots & \dots & \lambda_i^{n-k} & C_{n-k}^1 \lambda_i^{n-k-1} \\
0 & \dots & \dots & 0 & \lambda_i^{n-k}
\end{pmatrix}_{m_i \times m_i}
$$

Further, we can write J_i^{n-k} as

$$
\begin{pmatrix}
\lambda_i^{n-k} & e_1(n\lambda_i^{n-k}) + e_0\lambda_i^{n-k} & f_2(n^2\lambda_i^{n-k}) + f_1(n\lambda_i^{n-k}) + f_0\lambda_i^{n-k} & \dots & s \\
0 & \lambda_i^{n-k} & e_1(n\lambda_i^{n-k}) + e_0\lambda_i^{n-k} & \dots & \dots \\
\dots & \dots & \dots & \dots & \dots \\
0 & \dots & \dots & \lambda_i^{n-k} & e_1(n\lambda_i^{n-k}) + e_0\lambda_i^{n-k} \\
0 & \dots & \dots & 0 & \lambda_i^{n-k}
\end{pmatrix}_{m_i \times m_i}
$$

where $s = h_{m_i-1}(n^{m_i-1}\lambda_i^{n-k}) + h_{m_i-2}(n^{m_i-2}\lambda_i^{n-k}) + \dots + h_0\lambda_i^{n-k}$ and $e_1, e_0, f_2, f_1, f_0, h_{m_i-1}, \dots, h_0$ are constants. For instance, $e_1 = \frac{1}{\lambda_i}, e_0 = -\frac{k}{\lambda_i}$, similarly we can obtain other coefficients but there is no need to calculate them, just writing the matrix as this form is OK.

Therefore, we can obtain a_n by the equation (8) and matrix multiplication.(The essence of matrix multiplication is linear combination) That is

$$
\begin{aligned}
a_n &= \beta_{1,0} \cdot \lambda_1^{n-k} + \beta_{1,1} \cdot n\lambda_1^{n-k} + ... + \beta_{1,m_1-1} \cdot n^{m_1-1}\lambda_1^{n-k} \\
&+ \beta_{2,0} \cdot \lambda_2^{n-k} + \beta_{2,1} \cdot n\lambda_2^{n-k} + ... + \beta_{2,m_2-1} \cdot n^{m_2-1}\lambda_2^{n-k} \\
&... \\
&+ \beta_{t,0} \cdot \lambda_t^{n-k} + \beta_{t,1} \cdot n\lambda_t^{n-k} + ... + \beta_{t,m_t-1} \cdot n^{m_t-1}\lambda_t^{n-k} \\
&= \alpha_{1,0} \cdot \lambda_1^{n} + \alpha_{1,1} \cdot n\lambda_1^{n} + ... + \alpha_{1,m_1-1} \cdot n^{m_1-1}\lambda_1^{n} \\
&+ \alpha_{2,0} \cdot \lambda_2^{n} + \alpha_{2,1} \cdot n\lambda_2^{n} + ... + \alpha_{2,m_2-1} \cdot n^{m_2-1}\lambda_2^{n} \\
&... \\
&+ \alpha_{t,0} \cdot \lambda_t^{n} + \alpha_{t,1} \cdot n\lambda_t^{n} + ... + \alpha_{t,m_t-1} \cdot n^{m_t-1}\lambda_t^{n} \\
&= \sum_{i=1}^{t}(\sum_{j=0}^{m_i-1} \alpha_{i,j}n^j)\lambda_i^n \qquad \text{This is what we learned in lecture notes !}
\end{aligned} \tag{9}
$$

where $\alpha_{i,j} = \frac{\beta_{i,j}}{\lambda_i^k}$ are all constants.$(i = 1, 2, ..., t; j = 0, 1, ..., m_i - 1)$ By the initial conditions,we can find out the value of $\alpha_{i,j}$, and thus we obtain a formula for a_n.