

## Companion Matrices

---

Dan Kalman  
American University  
Washington, D.C. 20016  
kalman@american.edu

[www.dankalman.net](http://www.dankalman.net)

## Overview

---

- Definition and Elementary Properties
- Connection with Difference Equations
- Diagonalization & Vandermonde Matrices
- Generalized Fibonacci Numbers
- Jordan Form & Generalized Vandermonde
- Monorootic Case & Pascal Triangle Matrices
- Sums of Integer Powers

These topics  
will be covered  
in a separate  
slideshow file.

## What is a Companion Matrix?

---

- Let  $p$  be monic degree  $n$  polynomial.
- $C_p$  is an  $n \times n$  matrix
- $p$  is the characteristic polynomial of  $C_p$
- $p(t) = \det(tI - C_p)$
- $p(\lambda) = 0$  iff  $\lambda$  an eigenvalue of  $C_p$
- $C_p$  has a simple block structure
- For clarity, look at  $4 \times 4$  case. Things generalize in the obvious way.

---

$$p(t) = t^4 + at^3 + bt^2 + ct + d$$

$$C_p = \left[ \begin{array}{cccc} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -d & -c & -b & -a \end{array} \right]$$

Partitioned form:

$$\left[ \begin{array}{c|c} 0 & I \\ \hline & -\mathbf{c}^T \end{array} \right]$$

where  $\mathbf{c}^T = [d \ c \ b \ a]$

## Difference Equations

---

- Fibonacci Numbers  $0, 1, 1, 2, 3, 5, 8, \dots$
- Each term equals sum of two preceding terms
- Notation:  $F_k$  for  $k$ th term
- $F_{k+2} = F_k + F_{k+1}$
- Related example  $a_0, a_1, a_2, a_3 = 0, 0, 0, 1$ ;  $a_{k+4} = -a_k - a_{k+1} + 2a_{k+2} + a_{k+3}$
- General Case:  $a_{k+n}$  is a linear combination of preceding  $n$  terms, with constant coefficients
- $n$ th order, constant coefficient, linear, homogeneous, difference equation – but we just say *difference equation*

## Window Vectors

---

- View scalar sequence  $\{a_k\}$  as an infinite column vector  $\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \end{bmatrix}$
- Make a vector sequence by sliding an  $n$  term *window* down the column
- $\mathbf{a}_k = \begin{bmatrix} a_k \\ a_{k+1} \\ \vdots \\ a_{k+n-1} \end{bmatrix}$

4 term windows

$$\begin{array}{c} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ \vdots \end{array} \mathbf{a}_0$$

4 term windows

$$\begin{array}{c} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ \vdots \end{array} \mathbf{a}_1$$

4 term windows

$$\begin{array}{c} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ \vdots \end{array} \mathbf{a}_2$$

4 term windows

$$\begin{array}{c} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ \vdots \end{array} \mathbf{a}_7$$

## Companion Matrix $\times$ Window Vector

---

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} a_k \\ a_{k+1} \\ a_{k+2} \\ a_{k+3} \end{bmatrix} = \begin{bmatrix} a_{k+1} \\ a_{k+2} \\ a_{k+3} \\ -a_k - a_{k+1} + 2a_{k+2} + a_{k+3} \end{bmatrix}$$

## Vector Difference Equations

---

- Fact TFAE:

1.  $a_{k+4} = -a_k - a_{k+1} + 2a_{k+2} + a_{k+3}$  for  $k \geq 0$

2.  $\mathbf{a}_{k+1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -1 & 2 & 1 \end{bmatrix} \mathbf{a}_k$  for  $k \geq 0$

3.  $\mathbf{a}_k = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -1 & 2 & 1 \end{bmatrix}^k \mathbf{a}_0$  for  $k \geq 0$

- Observe companion matrix for  $p(t) = t^4 - t^3 - 2t^2 + t + 1$
- Generalizes to arbitrary difference equations

- Any scalar difference equation can be converted to a vector difference equation  $\mathbf{a}_{k+1} = C\mathbf{a}_k$  with  $C$  a companion matrix
- The *solution* of the vector difference equation is  $\mathbf{a}_k = C^k \mathbf{a}_0$
- That means we can find the scalar solution

$$\mathbf{a}_k = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \mathbf{a}_k = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_k \\ a_{k+1} \\ a_{k+2} \\ \vdots \\ a_{k+n-1} \end{bmatrix}$$

- As we have seen,  $C^k$  can be expressed simply by diagonalizing  $C$ .
- So let us analyze the eigenvalues and eigenvectors of companion matrices.
- First step: verify the characteristic polynomial formula

## 1,1 Minor

---

$$p(t) = d + ct + bt^2 + at^3 + t^4$$

$$C_p = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -d & -c & -b & -a \end{bmatrix}$$

Lower right submatrix is companion matrix for

$$t^3 + at^2 + bt + c = \frac{p(t) - d}{t}$$

## Characteristic Polynomial of $C_p$

$$p(t) = d + ct + bt^2 + at^3 + t^4$$

$$\det(tI - C_p) = \begin{vmatrix} t & -1 & 0 & 0 \\ 0 & t & -1 & 0 \\ 0 & 0 & t & -1 \\ d & c & b & t+a \end{vmatrix}$$

- Expand by minors in first column
- Cofactor of  $d$  is 1
- Cofactor of  $t$  is characteristic polynomial for the companion matrix of  $c + bt + at^2 + t^3 = (p(t) - d)/t$ .
- So if we know the theorem holds for  $3 \times 3$  matrices, that shows that the cofactor of  $t$  is  $(p(t) - d)/t$ .

Combined with the other cofactor, we obtain a final form for the determinant:  $t(p(t) - d)/t + d = p(t)$ .

This can be formalized as an induction proof for any degree  $n$ .

## Finding Eigenvectors

- What is the companion matrix for  $p(t) = t^n - 1$ ?
- Answer: the circulant matrix  $W$
- Recall: if  $\lambda$  is an eigenvalue of  $W$ , a nonzero eigenvector is given by  $[1 \ \lambda \ \lambda^2 \ \lambda^3 \ \dots \ \lambda^{n-1}]^T$
- This is a property that holds for all companion matrices
- We'll see that next.

## Key Identity

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -c_0 & -c_1 & -c_2 & \cdots & -c_{n-2} & -c_{n-1} \end{bmatrix} \begin{bmatrix} 1 \\ t \\ t^2 \\ \vdots \\ t^{n-2} \\ t^{n-1} \end{bmatrix} = \begin{bmatrix} t \\ t^2 \\ t^3 \\ \vdots \\ t^{n-1} \\ t^n - p(t) \end{bmatrix}$$

- Matrix is  $C_p$
- $p(t) = c_0 + c_1 t + \cdots + c_{n-1} t^{n-1} + t^n$  is the characteristic polynomial
- Define  $\mathbf{v}(t) = [1 \ t \ t^2 \ \cdots \ t^{n-1}]^T$  and  $\mathbf{e}_n = [0 \ 0 \ \cdots \ 0 \ 1]^T$
- Key identity: 
$$C_p \mathbf{v}(t) = t \mathbf{v}(t) - p(t) \mathbf{e}_n$$
- The value  $t = \lambda$  is an eigenvalue iff  $p(\lambda) = 0$ .
- In that case, the vector  $\mathbf{v}(\lambda) = [1 \ \lambda \ \lambda^2 \ \lambda^3 \ \cdots \ \lambda^{n-1}]^T$  is an eigenvector.

### One Dimensional Eigenspaces

- Suppose  $\lambda$  is an eigenvalue of  $C_p$  (ie, a root of  $p$ ).
- We know one eigenvector is given by multiples of  $[1 \ \lambda \ \lambda^2 \ \lambda^3 \ \cdots \ \lambda^{n-1}]^T$
- Are there others?
- To answer this, we can find the rref of  $\lambda I - C_p$ : the number of nonpivot columns equals the number of free variables in the solution of the homogeneous equation, and that is also the number of linearly independent solutions.
- Equivalently, the number of zero rows in the rref is equal to the number of linearly independent solutions. (Number of pivot cols = number of pivots = number of nonzero rows. So number of nonpivot columns = number of 0 rows.)
- Exercise: For any scalar  $\lambda$  (not just an eigenvalue), the first  $n-1$  rows of  $\lambda I - C_p$  are linearly independent.
- Exercise: The first  $n-1$  rows of the rref of  $\lambda I - C_p$  are nonzero
- Conclusion: If  $\lambda$  is an eigenvalue of  $C_p$  the corresponding eigenspace is one dimensional.



## Diagonalization

---

$$C_p \mathbf{v}(t) = t \mathbf{v}(t) - p(t) \mathbf{e}_n$$

- Roots of  $p$  are eigenvalues
- If  $p(\lambda) = 0$  then  $\mathbf{v}(\lambda)$  is an eigenvector
- Converse holds: eigenspace is one dimensional
- Diagonalization occurs iff  $p$  has  $n$  distinct roots
- $P = [ \mathbf{v}(\lambda_1) \ \mathbf{v}(\lambda_2) \ \cdots \ \mathbf{v}(\lambda_n) ]$   
is Vandermonde matrix  $V(\lambda_1, \lambda_2, \dots, \lambda_n)$
- $\det V(\lambda_1, \lambda_2, \dots, \lambda_n) = \prod_{i < j} (\lambda_j - \lambda_i)$

## Diagonalizable Case

---

- $C_p = PDP^{-1}$  where  $D$  is diagonal
- Possible IFF  $p$  has distinct roots
- $(PDP^{-1})^k = (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) = PD^kP^{-1}$

$$\bullet \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}$$

- Solution to difference equation:

$$\mathbf{a}_k = [1 \ 0 \ 0 \ \cdots \ 0] PD^k P^{-1} \mathbf{a}_0$$

## Generalized Fibonacci Numbers

---

- $n$ th order difference equation
- Initial terms  $a_0 = a_1 = \dots = a_{n-2} = 0$ ;  $a_{n-1} = 1$
- Vector form:  $\mathbf{a}_{k+1} = C_p \mathbf{a}_k$ ;  $\mathbf{a}_0 = \mathbf{e}_n$
- Solution:  $a_k =$

$$\begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} P \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix} P^{-1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

where  $P = V(\lambda_1, \lambda_2, \dots, \lambda_n)$

## Simplification

---

- $\begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} P D^k = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix} D^k = \begin{bmatrix} \lambda_1^k & \lambda_2^k & \dots & \lambda_n^k \end{bmatrix}$
- $P^{-1} \mathbf{e}_n$  is the solution to  $P \mathbf{x} = \mathbf{e}_n$
- Use Cramer's rule - all determinants are Vandermonde
- $x_j = 1/p'(\lambda_j)$
- Solution to difference equation:

$$a_k = \frac{\lambda_1^k}{p'(\lambda_1)} + \frac{\lambda_2^k}{p'(\lambda_2)} + \dots + \frac{\lambda_n^k}{p'(\lambda_n)}$$

## Cramer's Rule

$$x_j = \frac{\det \begin{bmatrix} 1 & \cdots & 0 & \cdots & 1 \\ \lambda_1 & \cdots & 0 & \cdots & \lambda_n \\ \lambda_1^2 & \cdots & 0 & \cdots & \lambda_n^2 \\ \vdots & & \vdots & & \vdots \\ \lambda_1^{n-2} & \cdots & 0 & \cdots & \lambda_n^{n-2} \\ \lambda_1^{n-1} & \cdots & 1 & \cdots & \lambda_n^{n-1} \end{bmatrix}}{\det \begin{bmatrix} 1 & \cdots & 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_j & \cdots & \lambda_n \\ \lambda_1^2 & \cdots & \lambda_j^2 & \cdots & \lambda_n^2 \\ \vdots & & \vdots & & \vdots \\ \lambda_1^{n-2} & \cdots & \lambda_j^{n-2} & \cdots & \lambda_n^{n-2} \\ \lambda_1^{n-1} & \cdots & \lambda_j^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix}} = (-1)^{n+j} \frac{\prod_{\mu < \nu, \mu \neq j, \nu \neq j} (\lambda_\nu - \lambda_\mu)}{\prod_{\mu < \nu} (\lambda_\nu - \lambda_\mu)}$$

- All the factors  $(\lambda_\nu - \lambda_\mu)$  in the numerator cancel with factors in the denominator
- That leaves only factors of the form  $(\lambda_j - \lambda_\mu)$  with  $\mu < j$  or  $(\lambda_\nu - \lambda_j)$  with  $\nu > j$  in the denominator
- The denominator is thus equal to  $(\lambda_j - \lambda_1)(\lambda_j - \lambda_2) \cdots (\lambda_j - \lambda_{j-1})(\lambda_{j+1} - \lambda_j)(\lambda_{j+2} - \lambda_j) \cdots (\lambda_n - \lambda_j)$
- Reversing the last  $n - j$  factors this produces  $(-1)^{n-j} \prod_{\mu \neq j} (\lambda_j - \lambda_\mu)$

- Therefore

$$x_j = \frac{(-1)^{n+j}}{(-1)^{n-j} \prod_{\mu \neq j} (\lambda_j - \lambda_\mu)} = \frac{1}{\prod_{\mu \neq j} (\lambda_j - \lambda_\mu)}$$

• OTOH, we know  $p(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n)$

• The product rule shows that

$$p'(t) = (t - \lambda_1)'(t - \lambda_2) \cdots (t - \lambda_n) \\ + (t - \lambda_1)(t - \lambda_2)' \cdots (t - \lambda_n) + \cdots \\ + (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_{n-1})'$$

• Now each differentiated quantity on the right equals 1. Therefore

$$p'(t) = (t - \lambda_2)(t - \lambda_3) \cdots (t - \lambda_n) \\ + (t - \lambda_1)(t - \lambda_3) \cdots (t - \lambda_n) + \cdots \\ + (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_{n-1})$$

• In words this is the sum of all products of  $n - 1$  of the linear factors of  $p$ . So what happens if we replace  $t$  with  $\lambda_j$ ?

• Answer: Any term with the factor  $(t - \lambda_j)$  vanishes, leaving only the term that excludes that factor. Then changing  $t$  to  $\lambda_j$  produces  $\prod_{\mu \neq j} (\lambda_j - \lambda_\mu)$ .

• Conclusion:  $p'(\lambda_j) = \prod_{\mu \neq j} (\lambda_j - \lambda_\mu) = 1/x_j$ .

## Example: Fibonacci Numbers

•  $a_{k+2} = a_k + a_{k+1}$

•  $p(t) = t^2 - t - 1$

• Roots:  $\phi = \frac{1 + \sqrt{5}}{2}$ ;  $\bar{\phi} = \frac{1 - \sqrt{5}}{2}$

•  $p'(t) = 2t - 1$

•  $p'(\phi) = \sqrt{5}$ ;  $p'(\bar{\phi}) = -\sqrt{5}$

•  $a_k = F_k = \frac{1}{\sqrt{5}} (\phi^k - \bar{\phi}^k)$

## Another Example

---

- $a_{k+4} = -a_k - a_{k+1} + 2a_{k+2} + a_{k+3}$

$$= (-1, -1, 2, 1) \cdot (a_k, a_{k+1}, a_{k+2}, a_{k+3})$$

- $(a_0, a_1, a_2, a_3) = (0, 0, 0, 1)$

- First few terms:

$$0, 0, 0, 1, 1, 3, 4, 8, 12, 21, 33, \dots$$

## Solution

---

- $p(t) = t^4 - t^3 - 2t^2 + t + 1 = (t^2 - 1)(t^2 - t - 1)$

- Roots:  $1, -1, \phi, \bar{\phi} = \lambda_i; i = 1, 2, 3, 4$

- $p'(t) = 4t^3 - 3t^2 - 4t + 1 = \sum_{i=1}^4 \prod_{j \neq i} (t - \lambda_j)$

- $p'(1) = -2, p'(-1) = -2$

- $p'(\phi) = (\phi - 1)(\phi + 1)(\phi - \bar{\phi}) = (\phi^2 - 1)\sqrt{5} = \phi\sqrt{5}$

- $p'(\bar{\phi}) = -\bar{\phi}\sqrt{5},$

### Solution continued

---

$$a_k = \frac{1}{-2} + \frac{(-1)^k}{-2} + \frac{\phi^k}{\phi\sqrt{5}} - \frac{\bar{\phi}^k}{\bar{\phi}\sqrt{5}}$$

$$a_k = \frac{(-1)^{k-1} - 1}{2} + \frac{\phi^{k-1} - \bar{\phi}^{k-1}}{\sqrt{5}} = \begin{cases} F_{k-1} - 1 & \text{for even } k \\ F_{k-1} & \text{for odd } k \end{cases}$$

$$\begin{aligned} \{a_k\} &= 0, 0, 0, 1, 1, 3, 4, 8, 12, 21, 33, \dots \\ \{F_{k-1}\} &= 1, 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \dots \end{aligned}$$