# Discrete-time Markov chains

## Stochastic algorithms

Francesco Corona

Department of Computer Science
Federal University of Ceará, Fortaleza





# Important matrices

## Discrete-time Markov chains



## Important matrices

We let $f_{ij}$ be the probability of ever reaching state $j$ from state $i$

- All $f_{ij}$s define the **reachability matrix** $F$

$$f_{jj}^{(n)} = p_{jj}^{(n)} - \sum_{l=1}^{n-1} f_{jj}^{(l)} p_{jj}^{(n-l)}, \quad (\text{for } n \geq 1)$$

$$f_{ij}^{(n)} = p_{ij}^{(n)} - \sum_{l=1}^{n-1} f_{ij}^{(l)} p_{jj}^{(n-l)}, \quad (\text{for } n \geq 1)$$

We use $F$ to compute the probability of visiting recurrent states

- Given that the process starts in a transient state



## Important matrices (cont.)

Let $r_{ij}$ be the expected number of times that state $j$ is visited

$\rightsquigarrow$ Given that the chain starts in state $i$

$$r_{ij} = \langle N_j | X_0 = i \rangle = \langle \Big[ \sum_{n=0}^{\infty} I_n | X_0 = i \Big] \rangle = \sum_{n=0}^{\infty} \langle I_n | X_0 = i \rangle$$

$$= \sum_{n=0}^{\infty} \text{Prob}\{X_n = j | X_0 = i\} = \sum_{n=0}^{\infty} p_{ij}^{(n)}$$

$I_n = 1$ if process is in $j$ at time $n$ and zero otherwise

- $\sum_{n=0}^{\infty} I_n$ is the number of times $j$ is occupied

Let $R$ be the matrix whose $(i,j)$-th element is $r_{ij}$

- Matrix $R$ is called the **potential matrix**

$$R = \sum_{n=0}^{\infty} P^n \tag{1}$$

## Important matrices (cont.)

Let the states be arranged so that transient states precede recurrent states

Upper-left corner of $R$ concerns in/out transitions from/to transient states

- This matrix defines the **fundamental matrix**

$$\rightsquigarrow \quad S$$

Element $s_{ij}$ is the expected number of times a chain is in transient state $j$

- Given that it started in transient state $i$





## Important matrices (cont.)

We have defined the reachability, the potential and the fundamental matrix

They can be used to compute important properties of the chain

- Mean and variance of the RV defining the number of visits to specific transient states, starting from potentially different transient states, before getting absorbed into a recurrent state
- The mean and variance of the RV defining the total number of steps the process makes before being absorbed into a recurrent state (mean time to absorption), given an initial transient state





## Important matrices (cont.)

Easier to first calculate matrix $R$ (and $S$), then from it calculate matrix $F$

We first concentrate on the potential and the fundamental matrix

- We compute the mean number of steps until absorption

We then concentrate on the reachability matrix

- We compute probabilities of absorption





# Potential and fundamental matrices

### Important matrices

## Potential and fundamental matrices

### Mean time to absorption

The elements of $R$ can be infinite, zero, and finite positive real numbers

$$\rightsquigarrow \quad R = \sum_{n=0}^{\infty} P^n$$

Consider some state $i$ and the elements of the $i$-th row of $R$

- $\rightsquigarrow$ $r_{ij}$ is the expected number of times state $j$ is visited
- $\rightsquigarrow$ Starting from state $i$

A number of different possibilities occur, depending on state classification

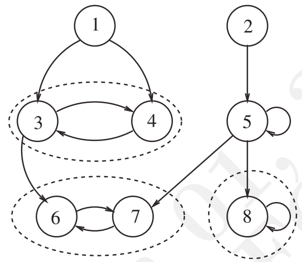
- $\rightsquigarrow$ They depend on both initial $i$ and final $j$ states



## Potential and fundamental matrices (cont.)

### CASE 1: State $i$ is recurrent

Let $i$ be a recurrent state[1]




State 6 and 7 are recurrent

State 8 is recurrent

The Markov process will return to state $i$ an infinite number of times

- $\rightsquigarrow$ Diagonal element in row $i$ and column $i$ is infinite, $\rightsquigarrow$ $r_{ii} = \infty$

The elements in column $j$ for states $j$ that communicate[2] with state $i$

- $\rightsquigarrow$ They must also be infinite, $\rightsquigarrow$ $r_{ij} = \infty$

True for all $j$s in the same closed communicating class, $C(i)$, as $i$

[1] The probability to ever return $f_{ii} = 1$, with mean recurrence time $M_{ii}$.
[2] The probability to ever visit $f_{ij} = 1$, with mean first passage time $M_{ij}$.



## Potential and fundamental matrices (cont.)

It is not possible to go from a recurrent state $i$ to any transient state $j$

$$\rightsquigarrow \quad r_{ij} = 0$$

Nor is possible to go from recurrent state $i$ to any recurrent state[3] $j$

- If $j$ is in a different irreducible class

Thus, all other elements $(i, j)$ of row $i$ of matrix $R$ must be zero

$$\rightsquigarrow \quad r_{ij} = 0$$

Element $r_{ij}$ is the expected number of times that state $j$ is visited

- Conditioned on the fact that the process starts from state $i$
- ($i$ is recurrent)

$$\rightsquigarrow \quad r_{ij} = \begin{cases} \infty, & j \in C(i) \\ 0, & \text{otherwise} \end{cases}$$

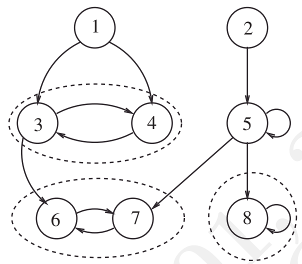
■

[3] The probability to ever visit $f_{ij} < 1$.



## Potential and fundamental matrices (cont.)

### CASE 2: State $i$ is transient and $j$ is recurrent




State 1 and 2 are transient

State 3 and 4 are transient

State 5 is transient

Suppose that transient state[4] $i$ can reach any state of recurrent[5] class $C(i)$

$$\rightsquigarrow \quad r_{ik} = \infty, \quad \text{for all } k \in C(i)$$

After leaving state $i$, the process may enter $C(i)$ and stay in it forever

Suppose transient state $i$ cannot reach any state of recurrent[6] class $C(i)$

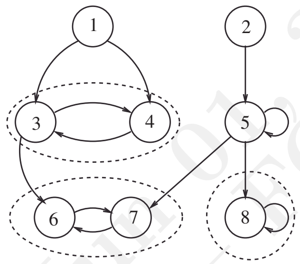
$$\rightsquigarrow \quad r_{ik} = 0, \quad \text{for all } k \in C(j)$$

■

[4] The probability to ever return $f_{ii} < 1$.
[5] The probability to ever visit $f_{ik} = 1$, with mean first passage time $M_{ik}$.
[6] The probability to ever visit $f_{ik} < 1$.

## Potential and fundamental matrices (cont.)

**CASE 3: State $i$ and state $j$ are both transient**




State 1 and 2 are transient

State 3 and 4 are transient

State 5 is transient

We can use the $(i, j)$ elements of the potential matrix $R = \sum_{n=0}^{\infty} P^n$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Potential and fundamental matrices (cont.)

Let states be ordered so that transient states come before recurrent ones

The transition probability matrix

$$\rightsquigarrow \quad P = \begin{pmatrix} T & U \\ 0 & V \end{pmatrix}$$

- $\rightsquigarrow$ Sub-matrix $T$ represents transitions between transient states only
- $\rightsquigarrow$ Sub-matrix $U$ represents transitions from transient to recurrent states
- $\rightsquigarrow$ Sub-matrix $V$ represents transitions between recurrent states only

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Potential and fundamental matrices (cont.)

For some matrix $\overline{U}$, we have

$$\rightsquigarrow \quad P^n = \begin{pmatrix} T^n & \overline{U}(n) \\ 0 & V^n \end{pmatrix}$$

Thus,

$$R = \sum_{n=0}^{\infty} P^n = \begin{pmatrix} \sum_{n=0}^{\infty} T^n & \sum_{n=0}^{\infty} \overline{U}(n) \\ 0 & \sum_{n=0}^{\infty} V^n \end{pmatrix} \equiv \begin{pmatrix} S & \sum_{n=0}^{\infty} \overline{U}(n) \\ 0 & \sum_{n=0}^{\infty} V^n \end{pmatrix}$$

We are interested in the expected number of visits to transient state $j$

- Given the process starts from transient state $i$

These quantities correspond to the element of matrix $S = \sum_{n=0}^{\infty} T^n$

- The only elements of $R$ that are not zero or infinity
- (The fundamental matrix)

■

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Potential and fundamental matrices (cont.)

Consider the $(i, j)$-th element of the fundamental matrix $S$

- $\rightsquigarrow$ It holds the expected number of visits to $j$, from $i$
- $\rightsquigarrow$ (Both $i$ and $j$ are transient)

By definition, we have

$$S = \sum_{n=0}^{\infty} T^n = I + T + T^2 + \cdots$$

Thus, we have

$$\rightsquigarrow \quad S - I = T + T^2 + T^3 + \cdots = TS \tag{2}$$

Hence,

$$\rightsquigarrow \quad S - TS = I$$
$$\rightsquigarrow \quad (I - T)S = I$$

Furthermore, it can be shown that $S$ satisfies $S(I - T) = I$

## Potential and fundamental matrices (cont.)

Consider the original series $ST = TS = T + T^2 + T^3 + \cdots$

- (Without identity matrix)

$T^n$ gives $n$-step transition probabilities from/to transient states

- Suppose that the number of transient states is finite
- Matrices $T^n$ must ultimately tend to zero

The chain cannot escape leaving the set of transient states

- At some point, it will enter a recurrent state



## Potential and fundamental matrices (cont.)

Thus, in the limit $n \to \infty$, we have that $T^n \to 0$

- Thus, series $I + T + T^2 + \cdots$ converges

Consider rewriting the expansion

$$I = (\underbrace{I + T + T^2 + \cdots}_{S})(I - T)$$

We have, the fundamental matrix $S$

$$\rightsquigarrow \quad S = \sum_{k=0}^{\infty} T^k = I + T + T^2 + \cdots = (I - T)^{-1}$$

The non-singularity of $(I - T)$ is required to compute $S$



## Potential and fundamental matrices (cont.)

If the number of states is not finite, we have

$$(I - T)X = I, X \geq 0$$

Matrix $S$ is the minimal non-negative solution to the system

■



## Potential and fundamental matrices (cont.)

**Example**

Consider the transition probability matrix of a discrete-time Markov chain

$$P = \begin{pmatrix} 0.4 & 0.2 & 0.0 & 0.2 & 0.0 & 0 & 0.0 & 0.2 \\ 0.3 & 0.3 & 0.0 & 0.0 & 0.1 & 0 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0.0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

Interest in the potential matrix $R$ and in the fundamental matrix $S$

- State 1, 2 and 3 are transient
- State 4 and 5 define an irreducible set
- State 6 is an absorbing state
- State 7 and 8 define an irreducible set

## Potential and fundamental matrices (cont.)

$$P = \begin{pmatrix} 0.4 & 0.2 & 0.0 & 0.2 & 0.0 & 0 & 0.0 & 0.2 \\ 0.3 & 0.3 & 0.0 & 0.0 & 0.1 & 0 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0.0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

- State 1, 2 and 3 are transient
- State 4 and 5 define an irreducible set
- State 6 is an absorbing state
- State 7 and 8 define an irreducible set

Consider row $4-8$ of matrix $R$

All its elements are zero, except for the elements in diagonal blocks

- Transitions among recurrent states of the same closed class
- They are all set to be equal to infinity



## Potential and fundamental matrices (cont.)

$$P = \begin{pmatrix} 0.4 & 0.2 & 0.0 & 0.2 & 0.0 & 0 & 0.0 & 0.2 \\ 0.3 & 0.3 & 0.0 & 0.0 & 0.1 & 0 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0.0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

- State 1, 2 and 3 are transient
- State 4 and 5 define an irreducible set
- State 6 is an absorbing state
- State 7 and 8 define an irreducible set

Consider row $1-3$ of matrix $R$

All its elements in positions 4, 5, 7 and 8 are equal to infinity

The chain will eventually transition from states $1-3$

- Destination 4 and 5, irreducible subset
- Destination 7 and 8, irreducible subset

The process will remain there



## Potential and fundamental matrices (cont.)

$$P = \begin{pmatrix} 0.4 & 0.2 & 0.0 & 0.2 & 0.0 & 0 & 0.0 & 0.2 \\ 0.3 & 0.3 & 0.0 & 0.0 & 0.1 & 0 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0.0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

- State 1, 2 and 3 are transient
- State 4 and 5 define an irreducible set
- State 6 is an absorbing state
- State 7 and 8 define an irreducible set

Consider all the elements of $R$ that are in column 6

There is no path from any transient state to absorbing

- They must be equal to zero



## Potential and fundamental matrices (cont.)

Consider the fundamental matrix $S = (I - T)^{-1}$

We have,

$$\rightsquigarrow \quad T = \begin{pmatrix} 0.4 & 0.2 & 0 \\ 0.3 & 0.3 & 0 \\ 0 & 0 & 0.1 \end{pmatrix}$$

Moreover,

$$\rightsquigarrow \quad (I - T) = \begin{pmatrix} 0.6 & -0.2 & 0 \\ -0.3 & 0.7 & 0 \\ 0 & 0 & 0.9 \end{pmatrix}$$

Thus,

$$\rightsquigarrow \quad S = (I - T)^{-1} = \begin{pmatrix} 1.94 & 0.56 & 0 \\ 0.83 & 1.67 & 0 \\ 0 & 0 & 1.11 \end{pmatrix}$$

## Potential and fundamental matrices (cont.)

Therefore, we can complete the potential matrix $R$

$$R = \begin{pmatrix} 1.94 & 0.56 & 0 & \infty & \infty & 0 & \infty & \infty \\ 0.83 & 1.67 & 0 & \infty & \infty & 0 & \infty & \infty \\ 0 & 0 & 1.11 & \infty & \infty & 0 & \infty & \infty \\ 0 & 0 & 0 & \infty & \infty & 0 & 0 & 0 \\ 0 & 0 & 0 & \infty & \infty & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \infty & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \infty & \infty \\ 0 & 0 & 0 & 0 & 0 & 0 & \infty & \infty \end{pmatrix}$$

Element $(i, j)$ is the expected number of visits to state $j$

- Given that the process stated from state $i$






## Potential and fundamental matrices (cont.)

Let $N_{ij}$ be the RV defining the total number of visits to $j$ (transient)

- Given that the process stated from state $i$

Let $N$ be the matrix with those elements, we have

$$\langle N_{ij} \rangle = s_{ij}, \quad \rightsquigarrow \quad \langle N \rangle = S$$

Moreover, we have

$$\langle N^2 \rangle = S\big[2\text{diag}(S) - I\big]$$

And,

$$\text{Var}(N) = \langle N^2 \rangle - \langle N \rangle^2 = S\big[2\text{diag}(S) - I\big] - \text{sq}(S)$$

$\text{sq}(S)$ is a matrix whose $(i, j)$-th element is $(s_{ij} \times s_{ij})$






## Potential and fundamental matrices (cont.)

Consider the element $(i, j)$ of the fundamental matrix $S$

$$S = (I - T)^{-1} = \begin{pmatrix} 1.94 & 0.56 & 0.0 \\ 0.83 & 1.67 & 0.0 \\ 0.0 & 0.0 & 1.11 \end{pmatrix}$$

Number of visits to transient state $j$, from transient state $i$

Consider the sum of all the elements in row $i$ of matrix $S$

- Mean number of steps from $i$ before absorption
- The $i$-th element of vector $Se$

The variance of the total time to absorption, from $i$

$$(2S - I)Se - \text{sq}(Se)$$

$\text{Sq}(Se)$ is a vector whose $i$-th element is $(Se)_i$

■






## Potential and fundamental matrices (cont.)

**Theorem**

*Consider a discrete-time chain with a finite number of transient states*

*Let the transition probability matrix $P$*

$$P = \begin{pmatrix} T & U \\ 0 & V \end{pmatrix} \tag{3}$$

*Assume that the Markov chain begin at some transient state $i$*

- ⇝ *The mean number of times the chain visits transient state $j$ is given by the $(i, j)$-th element of matrix $S = (I - T)^{-1}$*
- ⇝ *The variance of the number of times the chain visits transit state $j$ is given by the $(i, j)$-th element of matrix $S\big[2diag(S) - I\big] - sq(S)$*
- ⇝ *The mean time to absorption (average number of steps among transient states before visiting a recurrent state) is given by the $i$-th element of vector $Se = (I - T)^{-1}e$*
- ⇝ *The variance of the time to absorption is given by the $i$-th element of vector $(2S - I)Se - sq(Se)$*

■

## Potential and fundamental matrices (cont.)

Suppose that the Markov chain starts from state $i$ with probability $\alpha_i$

⇝ $\alpha$ is the vector of initial probabilities

The results must be modified accordingly

⇝ $\alpha S$ is a vector whose $j$-th component gives the mean number of visits to state $j$ before absorption, $\alpha S\big[2\text{diag}(S) - I\big] - \text{sq}(\alpha S)$ is the variance

⇝ $\alpha S e$ is a real number that gives the expected number of steps before absorption, $\alpha(2S - I)Se - (\alpha Se)^2$ is the variance



## Potential and fundamental matrices (cont.)

The sub-matrix $T$ (usually) represents the subset of transient states

- Yet, the theorem holds, regardless

We can use this theorem to investigate non-absorbing state $i$

- We can let $T$ be the single element $p_{ii}$
- Say, $i = 1$ ($p_{11}$ strictly smaller than $1$)

As a result, matrix $S$ consists also of a single element $1/(1 - p_{ii})$

- Average number of steps the chain remains in $i = 1$
- (It is a non-absorbing state)



## Potential and fundamental matrices (cont.)

The variance can be computed

$$\frac{1}{(1-p_{11})}\Big(\frac{2}{1-P_{11}} - 1\Big) - \Big(\frac{1}{1-p_{11}}\Big)^2$$
$$= \frac{2}{(1-p_{11})^2} - \frac{1}{(1-p_{11})} - \frac{1}{(1-p_{11})^2}$$
$$= \frac{1}{(1-p+11)^2} - \frac{1}{1-p_{11}} = \frac{p_{11}}{(1-p_{11})^2}$$

We derived these results earlier

⇝ Holding time



## Potential and fundamental matrices (cont.)

**Example**

Consider the transition probability matrix of a discrete-time Markov chain

$$P = \begin{pmatrix} 0.4 & 0.2 & 0.0 & 0.2 & 0.0 & 0 & 0.0 & 0.2 \\ 0.3 & 0.3 & 0.0 & 0.0 & 0.1 & 0 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0.0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

The fundamental matrix $S = (I - T)^{-1}$

$$S = (I - T)^{-1} = \begin{pmatrix} 1.94 & 0.56 & 0.0 \\ 0.83 & 1.67 & 0.0 \\ 0.0 & 0.0 & 1.11 \end{pmatrix}$$

Discrete-time Markov chains

UFC/DC SA (CK0191) 2018.1



## Potential and fundamental matrices (cont.)

The mean time to absorption, when starting from state $i = 1, 2, 3$

$$Se = (I - T)^{-1} e = \begin{pmatrix} 1.94 & 0.56 & 0.0 \\ 0.83 & 1.67 & 0.0 \\ 0.0 & 0.0 & 1.11 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2.5 \\ 2.5 \\ 1.11 \end{pmatrix}$$

The variance of the number of times the chain visits $j$ starting from $i$

$$S[2\text{diag}(S) - I] - \text{sq}(S)$$
$$= \begin{pmatrix} 1.94 & 0.56 & 0 \\ 0.83 & 1.67 & 0 \\ 0 & 0 & 1.11 \end{pmatrix} \begin{pmatrix} 2.89 & 0 & 0 \\ 0 & 2.33 & 0 \\ 0 & 0 & 1.22 \end{pmatrix} - \begin{pmatrix} 3.78 & 0.31 & 0 \\ 0.69 & 2.78 & 0 \\ 0 & 0 & 1.23 \end{pmatrix}$$
$$= \begin{pmatrix} 1.83 & 0.98 & 0 \\ 1.71 & 1.11 & 0 \\ 0 & 0 & 1.23 \end{pmatrix}$$

Discrete-time Markov chains

UFC/DC SA (CK0191) 2018.1

Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Potential and fundamental matrices (cont.)

The variance of the total time to absorption from $i$

$$\big[2S - I\big](Se) - \text{sq}(Se)$$
$$= \begin{pmatrix} 2.89 & 1.11 & 0 \\ 1.67 & 2.33 & 0 \\ 0 & 0 & 1.22 \end{pmatrix} \begin{pmatrix} 2.5 \\ 2.5 \\ 1.11 \end{pmatrix} - \begin{pmatrix} 6.25 \\ 6.25 \\ 1.23 \end{pmatrix} = \begin{pmatrix} 3.75 \\ 3.75 \\ 0.23 \end{pmatrix}$$

■

Discrete-time Markov chains

UFC/DC SA (CK0191) 2018.1

Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

# Reachability matrix

Important matrices

Discrete-time Markov chains

UFC/DC SA (CK0191) 2018.1

Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix

We discuss the computation of the elements of the reachability matrix $F$

Consider the $(i, j)$-th element of the reachability matrix $F$

- The probability of ever reaching $j$ from $i$, $f_{ij}$

We shall split elements $f_{ij}$ into categories

- Based on initial and final states

## Reachability matrix (cont.)

**State $i$ and state $j$ are recurrent and belong to the same closed communicating class**

$$\rightsquigarrow \quad f_{ij} = 1$$

**State $i$ and state $j$ are recurrent and belong to the same closed communicating class**

$$\rightsquigarrow \quad f_{ij} = 0$$

**State $i$ is recurrent and state $j$ is transient**

$$\rightsquigarrow \quad f_{ij} = 0$$

**State $i$ and state $j$ are both transient**

$$\rightsquigarrow \quad f_{ij} < 1$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

**State $i$ and state $j$ are both transient** $(f_{ij} < 1)$

**Expected number of hits to transient state $j$ from transient $i$**

$$\rightsquigarrow \quad s_{ij} = \sum_{n=0}^{\infty} p_{ij}^{(n)} = \sum_{n=0}^{\infty} \sum_{l=1}^{n-1} f_{ij}^{(l)} p_{jj}^{(n-l)}$$

For any finite number of states

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)} = \sum_{n=1}^{\infty} \Big[ p_{ij}^{(n)} - \sum_{l=1}^{n-1} f_{ij}^{(n)} p_{jj}^{(n-l)} \Big]$$

Thus,

$$\rightsquigarrow \quad s_{ij} = 1_{[i=j]} + f_{ij} s_{jj}$$

$1_{[i=j]}$ is equal to one if $i = j$, zero otherwise

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

Let $H$ denote the upper left corner of the reachability matrix $F$

- (Transient states to transient states only)

In matrix notation,

$$\rightsquigarrow \quad S = I + H \big[ \text{diag}(S) \big]$$

Alternatively,

$$\rightsquigarrow \quad H = (S - I) \big[ \text{diag}(S) \big]^{-1} \tag{4}$$

The inverse must exist, as $s_{ii} > 0$, for all $i$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

We may write the elements of $S$ in terms of the elements of $H$

$$\rightsquigarrow \quad s_{ii} = \frac{1}{1 - h_{ii}}$$

$$\rightsquigarrow \quad s_{ij} = h_{ij} s_{jj}, \quad \text{(for } i \neq j\text{)}$$

Alternatively,

$$\rightsquigarrow \quad h_{ii} = 1 - \frac{1}{s_{ii}}$$

$$\rightsquigarrow \quad h_{ij} = \frac{s_{ij}}{h_{ij}}, \quad \text{(for } i \neq j\text{)}$$

## Reachability matrix (cont.)

**State $i$ and state $j$ are both transient ($f_{ij} < 1$)**

**Probability of visiting state $j$ a fixed number of times**

- (Assuming a starting state $i$)

Consider visiting $j$ a number $k > 0$ times

We must have,

1. Transitions $i \rightarrow j$ at least once (occurs with probability $h_{ij}$)
2. Returns from $j$ to $j$, $(k-1)$ times (probability $h_{jj}^{k-1}$)
3. No returns $j$ to $j$ again (probability $1 - h_{jj}$)

In matrix terms,

$$\rightsquigarrow \quad H \times \text{diag}(H)^{k-1} \times \big[I - \text{diag}(H)\big]$$

Discrete-time Markov chains

UFC/DC
SA (CK0191)
2018.1

Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

$$H \times \text{diag}(H)^{k-1} \times \big[I - \text{diag}(H)\big]$$

Using $H = (S - I)\big[\text{diag}(S)\big]^{-1}$ and observing that $\text{diag}(H) = I - \big[\text{diag}(S)\big]^{-1}$

$$\rightsquigarrow \quad (S - I)\big[\text{diag}(S)\big]^{-1} \times \big[\text{diag}(S)\big]^{-1} \times (I - \big[\text{diag}(S)\big]^{-1})^{k-1}$$

The probability state $j$ is visited zero times, starting from state $i$

- Zero if $i = j$, and $(1 - h_{ij})$ if $i \neq j$

■

Discrete-time Markov chains

UFC/DC
SA (CK0191)
2018.1

Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

**State $i$ and state $j$ are both transient ($f_{ij} < 1$)**

**Mean number of different transient states before absorption**

- (Into some recurrent class, from some transient state $i$)

It is the sum of probabilities of visiting the different transient states

- The probability of ever hitting state $j$ from $i$ is $h_{ij}$
- The probability of ever visiting $i$ from $i$ is one

The mean number of transient states visited before absorption

$$\rightsquigarrow \quad 1 + \sum_{j \neq i} h_{ij}$$

The sum across row $i$ of $(H - \text{diag}(H) + I)$

Discrete-time Markov chains

UFC/DC
SA (CK0191)
2018.1

Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

In terms of the fundamental matrix $S$

$$\begin{aligned}
\big[H - \text{diag}(H) + I\big]e &= \big[H + (I - \text{diag}(H))\big]e \\
&= ((S - I)\big[\text{diag}(S)\big]^{-1} + (I - (I - \big[\text{diag}(S)\big]^{-1}))e \\
&= ((S - I)\big[\text{diag}(S)\big]^{-1} + \big[\text{diag}(S)\big]^{-1})e \\
&= S\big[\text{diag}(S)\big]^{-1}e
\end{aligned}$$

■

## Reachability matrix (cont.)

**State $i$ is transient and state $j$ is recurrent**

Suppose that some recurrent state $j$ can be reached from transient state $i$

Then, all states in $C(j)$ can be reached from $i$, with same probability

- $C(j)$ is the recurrent class containing state $j$

$$\rightsquigarrow \quad f_{ik} = f_{ij}, \quad \text{for all } k \in C(i)$$

They are called **absorption probabilities**





## Reachability matrix (cont.)

Combine all states in an irreducible recurrent set into an absorbing state

- Compute the probability of entering this state from transient $i$

Assume a state arrangement in normal form

$$P = \begin{pmatrix} T_{11} & T_{12} & T_{13} & \cdots & T_{1N} \\ 0 & R_2 & 0 & \cdots & 0 \\ 0 & 0 & R_3 & \cdots & 0 \\ \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & R_N \end{pmatrix}$$

Replace each block $R_k$ with $1$ (an absorbing state)

Block $T_{1k}$ is replaced by vector $t_k$

- Sum over row $T_{ik}$, $t_k = T_{1k}e$

Block $T_{11}$ is kept unchanged



## Reachability matrix(cont.)

We obtain

$$\rightsquigarrow \quad \overline{P} = \begin{pmatrix} T_{11} & t_2 & t_3 & \cdots & t_N \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

This process is denoted as **absorbing chain**

■



## Reachability matrix(cont.)

**Example**

Consider the transition matrix $P$ of a discrete-state Markov chain

$$P = \begin{pmatrix} 0.4 & 0.2 & 0.0 & 0.2 & 0.0 & 0 & 0.0 & 0.2 \\ 0.3 & 0.3 & 0.0 & 0.0 & 0.1 & 0 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0.0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

The transition matrix $\overline{P}$ of the absorbing chain

$$P = \begin{pmatrix} 0.4 & 0.2 & 0 & 0.2 & 0 & 0.2 \\ 0.3 & 0.3 & 0 & 0.1 & 0 & 0.3 \\ 0 & 0 & 0.1 & 0.4 & 0 & 0.5 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

■

## Reachability matrix (cont.)

$$\overline{P} = \begin{pmatrix} T_{11} & t_2 & t_3 & \cdots & t_N \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

We simplify the notation and re-write the absorbing chain

$$\rightsquigarrow \quad \overline{P} = \begin{pmatrix} T & B \\ 0 & I \end{pmatrix}$$

The higher-order powers of $\overline{P}$

$$\rightsquigarrow \quad \overline{P}^n = \begin{pmatrix} T^n & (I + T + \cdots + T^{n-1})B \\ 0 & I \end{pmatrix} = \begin{pmatrix} T^n & B_n \\ 0 & I \end{pmatrix}$$

We can compute the absorption probabilities



## Reachability matrix (cont.)

$$\overline{P}^n = \begin{pmatrix} T^n & (I + T + \cdots + T^{n-1})B \\ 0 & I \end{pmatrix} = \begin{pmatrix} T^n & B_n \\ 0 & I \end{pmatrix}$$

The probability of entering the $j$-th irreducible recurrent class $n$ steps after starting from transient state $i$ is given by element $(i, j)$ of matrix $B_n$

- $\rightsquigarrow$ States of the first recurrent class, transitions associated to block $R_2$
- $\rightsquigarrow$ States of the second recurrent class, transitions associated to block $R_3$
- $\rightsquigarrow$ $\cdots$

The probability of ever being absorbed from state $i$ into the $j$-th recurrent

- The $(i, j)$-th element of matrix $A = \lim_{n\to\infty} B_n$



## Reachability matrix (cont.)

Matrix $A = \lim_{n\to\infty} B_n$ is called **absorption probability matrix**

Suppose that the number of transient states is finite

We can obtain $A$ from the fundamental matrix $S$

$$A = \lim_{n\to\infty} B_n = \lim_{n\to\infty} (I + T + \cdots + T^{n-1})B$$
$$= \Big(\sum_{k=0}^{\infty} T^k\Big) B = (I - T)^{-1} B = SB$$

The $(i, j)$-th element, probability of ever reaching $j$-th recurrent class

- Starting from transient state $i$

For every state $k$ in this class, we have that $f_{ik} = a_{ij}$

- We get them all from the inner product $SB$



## Reachability matrix (cont.)

Consider the following absorption probability matrix $A$

$$\overline{A} = \begin{pmatrix} 0 & A \\ 0 & I \end{pmatrix}$$

Then,

$$\rightsquigarrow \quad \overline{P}\,\overline{A} = \begin{pmatrix} T & B \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & A \\ 0 & I \end{pmatrix} = \begin{pmatrix} 0 & A \\ 0 & I \end{pmatrix} = \overline{A}$$

Since,

$$TA + B = TSB + B = (S - I)B + B = SB = A$$

We used $A = SB$ and $TS = (S - I)$

## Reachability matrix (cont.)

Let a column of matrix $\overline{A}$ to be the vector $\overline{\alpha}$

We have,

$$\overline{P}\,\overline{\alpha} = \overline{\alpha}$$

$\overline{\alpha}$ is the right eigenvector associated to a unit eigenvalue of $\overline{P}$



Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

**Example**

Consider the transition matrix $P$ of a discrete-state Markov chain

$$P = \begin{pmatrix} 0.4 & 0.2 & 0.0 & 0.2 & 0.0 & 0 & 0.0 & 0.2 \\ 0.3 & 0.3 & 0.0 & 0.0 & 0.1 & 0 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0.0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

We are interested in the first-return matrix $F$

We can directly determine the elements of rows 4 to 8

- One inside diagonal blocks
- Zero elsewhere

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

We can compute the first $3 \times 3$ block by using $H = (S - I)\big[\text{diag}(S)\big]^{-1}$

$$H = F_{3\times 3} = (S - I)\big[\text{diag}(S)\big]^{-1}$$
$$= \begin{pmatrix} 0.9444 & 0.5556 & 0.0 \\ 0.8333 & 0.6667 & 0.0 \\ 0.0 & 0.0 & 0.11111 \end{pmatrix} \begin{pmatrix} 1.9444 & 0.0 & 0.0 \\ 0.0 & 1.6667 & 0.0 \\ 0.0 & 0.0 & 1.1111 \end{pmatrix}^{-1}$$
$$= \begin{pmatrix} 0.4857 & 0.3333 & 0.0 \\ 0.4286 & 0.4000 & 0.0 \\ 0.0 & 0.0 & 0.1 \end{pmatrix}$$

We still need to find the absorption probabilities

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

We have,

$$\rightsquigarrow \quad A = \lim_{n\to\infty} B_n = (I - T)^{-1} B = SB$$
$$= \begin{pmatrix} 0.6 & -0.2 & 0.0 \\ -0.3 & 0.7 & 0.0 \\ 0.0 & 0.0 & 0.9 \end{pmatrix}^{-1} \begin{pmatrix} 0.2 & 0.0 & 0.2 \\ 0.1 & 0.0 & 0.3 \\ 0.4 & 0.0 & 0.5 \end{pmatrix}$$
$$= \begin{pmatrix} 0.4444 & 0.0 & 0.5556 \\ 0.3333 & 0.0 & 0.6667 \\ 0.4444 & 0.0 & 0.5556 \end{pmatrix}$$

Element $(i, j)$ is the probability of being absorbed into the $j$-th class

- Starting from transient state $i$

(The sum of absorption probabilities is 1, for each state)

## Reachability matrix (cont.)

The complete matrix $F$

$$F = \begin{pmatrix} 0.4857 & 0.3333 & 0 & 0.4444 & 0.4444 & 0 & 0.5556 & 0.5556 \\ 0.4286 & 0.4 & 0 & 0.3333 & 0.3333 & 0 & 0.6667 & 0.6667 \\ 0 & 0 & 0.1 & 0.4444 & 0.4444 & 0 & 0.5556 & 0.5556 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

The results refer to Markov chains with both transient and recurrent states

Transition from transient states to one or more close communicating classes

- (Transition probabilities)

We can use sich results for chains without transient states

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

Consider a finite irreducible Markov chain

We wish to compute the probability of reaching state $i$ before state $j$

- Starting from any other state $k$ ($i \neq j \neq k$)

This can be achieved by splitting states $i$ and $j$ from the other states

- We construct the fundamental matrix of the remainder
- (All states become transient states)

We can then compute the probability that starting in a non-absorbing state $k$, the chain visits an absorbing state $i$ before it visits an absorbing state $j$

The set of all states other than $i$ and $j$ forms an open set

- States $i$ and $j$ can be turned into absorbing states

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Reachability matrix (cont.)

Assume there is a total of $n$ states with some specific ordering

1. Set of $(n-2)$ states not including $i$ or $j$ but including $k$
2. State $i$
3. State $j$

$\leadsto$ The probability that $i$ is visited before $j$
The elements of the vector as product of matrix $S$ and a vector $v_{n-1}$ whose components are the first $(n-2)$ elements of column $(n-1)$ of matrix $P$

- The elements that set the conditional probabilities of entering state $i$
- Given that the Markov chain is in state $k$, $k = 1, 2, \dots, (n-2)$

$\leadsto$ The probability that $j$ is visited before $i$
The elements of the vector as product of matrix $S$ and a vector $v_n$ whose components are the first $(n-2)$ elements of the last column of matrix $P$

## Reachability matrix (cont.)

### Example

Consider the transition probability matrix of a six-state Markov chain

$$P = \begin{pmatrix} 0.25 & 0.25 & 0 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0.50 & 0.5 \end{pmatrix}$$

We have,

$$S = \begin{pmatrix} 0.75 & -0.25 & 0 & 0 \\ 0.50 & 1 & -0.5 & 0 \\ 0 & -0.50 & 1 & -0.5 \\ 0 & 0 & -0.5 & 1 \end{pmatrix}^{-1} = 2/9 \begin{pmatrix} 8 & 3 & 2 & 1 \\ 6 & 9 & 6 & 3 \\ 4 & 6 & 10 & 5 \\ 2 & 3 & 5 & 7 \end{pmatrix}$$



## Reachability matrix (cont.)

We have

$$v_5 = (0.5, 0.0, 0.0, 0.0)^T$$
$$v_6 = (0.0, 0.0, 0.0, 0.5)^T$$

From this we obtain,

$$Sv_5 = (0.8889, 0.6667, 0.4444, 0.2222)^T$$
$$Sv_6 = (0.1111, 0.3333, 0.5556, 0.7778)^T$$

The probability that state $i = 5$ is reached before state $j = 6$ is 0.8889

- Given initial state $k = 1$

The probability that state $j$ is reached before state $i$ is 0.1111



## Reachability matrix (cont.)

Consider combining $v_5$ and $v_6$ into a single $4 \times 2$

⇝ We obtain what we called matrix $B$

We are computing matrix $SB = A$ of absorption probabilities



# Distributions

## Discrete-time Markov chains

## Distributions

We study the probability distributions defined on the states of a chain

- $\rightsquigarrow$ We consider the homogeneous discrete-time Markov chains

Focus on the probability that the chain is in some state at some step





## Distributions (cont.)

Let $\pi_i(n)$ be the probability that the chain is in state $i$ at time $n$

$$\rightsquigarrow \quad \pi_i(n) = \text{Prob}\{X_n = i\}$$

- $\rightsquigarrow$ Vector $\pi$ is a row-vector of probabilities
- $\rightsquigarrow$ $\pi(n) = [\pi_1(n), \pi_2(n), \dots, \pi_i(n), \dots]$

State probabilities at time $n$ are derived from transition matrices

- Given the initial probability distribution at time $0$, $\pi(0)$

Using the laws of probability, we get

$$\pi_i(n) = \sum_{\text{all } k} \text{Prob}\{X_n = i | X_0 = k\}\pi_k(0) \tag{5a}$$

$$= \sum_{\text{all } k} p_{ki}^{(n)} \pi_k(0) \tag{5b}$$





## Distributions (cont.)

In matrix notation, we have

$$\rightsquigarrow \quad \pi(n) = \pi(0)P^{(n)} = \pi(0)P^n$$

- $\rightsquigarrow$ $\pi(0)$ denotes the initial state distribution
- $\rightsquigarrow$ $P^{(n)} = P^n$, as the chain is homogeneous

The distribution $\pi(n)$ is referred to as **transient distribution** of the chain

- It gives the probability of the process states at a particular time $n$
- It is transient, as it is dropped as the chain goes to step $n + 1$





## Distributions (cont.)

We now consider three state distributions

- $\rightsquigarrow$ **Limiting distributions**
- $\rightsquigarrow$ **Stationary distributions**
- $\rightsquigarrow$ **Steady-state distributions**

## Distributions (cont.)

**Definition**

***Stationary distribution***

*Let $P$ be the transition probability matrix of a discrete-time Markov chain*

*Let vector $z \in \mathcal{R}$ of elements $z_j$ indicate the probability of states $j$*

$$0 \leq z_j \leq 1$$
$$\sum_{\text{all } j} z_j = 1$$

*$z$ is said to be a **stationary distribution** if and only if $zP = z$*

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

$$\rightsquigarrow \quad z = zP = zP^2 = \cdots = zP^n = \cdots$$

Suppose that $z$ is chosen to be the initial probability distribution

$$\rightsquigarrow \quad \pi_j(0) = z_j \ \text{(for all } j\text{)}$$

Then, we have,

$$\rightsquigarrow \quad \pi_j(n) = z_j \ \text{(for all } n\text{)}$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

This distribution resembles vector $z = zP$ (introduced for irreducible chains)

- ⇝ The components of that vector $z$ must be strictly positive
- ⇝ Its existence implies that the states are positive-recurrent

Any non-zero vector that satisfies $z = Pz$ is called an **invariant vector**

- Its elements are not necessarily probabilities are

It can be any left-hand eigenvector associated to a unit eigenvalue of $P$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

**Example**

Let $P$ be the transition matrix of a discrete-time Markov chain

$$P = \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

Consider the following distributions

$$z = (1/2, 1/2, 0, 0)$$
$$z = (0, 0, 1/2, 1/2)$$
$$z = (\alpha/2, \alpha/2, (1-\alpha)/2, (1-\alpha)/2) \ [\text{for } 0 \geq \alpha \geq 1]$$

Are these distributions stationary distributions?

## Distributions (cont.)

Each of these distributions satisfy $z = Pz$ and are stationary distributions

$$z = zP$$

$$(1/2, 1/2, 0, 0) = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix}^T \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

$$(0, 0, 1/2, 1/2) = \begin{pmatrix} 0 \\ 0 \\ 1/2 \\ 1/2 \end{pmatrix}^T \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

$$(\alpha/2, \alpha/2, (1-\alpha)/2, (1-\alpha)/2) = \begin{pmatrix} \alpha/2 \\ \alpha/2 \\ (1-\alpha)/2 \\ (1-\alpha)/2 \end{pmatrix}^T \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

[for $0 \geq \alpha \geq 1$]





## Distributions (cont.)

Also vector $(1, 1, -1, 1)$ is an invariant vector

$$z = zP$$

$$(1, 1, -1, -1) = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}^T \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

But, it is not a stationary distribution





## Distributions (cont.)

Discrete-time Markov chain often have a unique stationary distribution





## Distributions (cont.)

**Example**

Let $P$ be the transition matrix of a discrete-time Markov chain

$$P = \begin{pmatrix} 0.45 & 0.50 & 0.05 \\ 0.15 & 0.65 & 0.20 \\ 0.00 & 0.50 & 0.50 \end{pmatrix}$$

We are interested in the stationary distribution of this process

$z = Pz$ defines the homogeneous systems of equations $z(P - I) = 0$

- ⇝ The system must have a singular coefficient matrix $(P - I)$
- ⇝ Otherwise, $z = 0$

## Distributions (cont.)

$$0 = z(P - I) = (z_1, z_2, z_3)\left[\begin{pmatrix} 0.45 & 0.50 & 0.05 \\ 0.15 & 0.65 & 0.20 \\ 0.00 & 0.50 & 0.50 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]$$

The singular coefficient matrix gives two linearly independent equations

$$\begin{aligned} z_1 &= 0.45 z_1 + 0.15 z_2 \\ z_2 &= 0.50 z_1 + 0.65 z_2 + 0.50 z_3 \end{aligned}$$

(They are the first two equations out of three possible ones)



Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

$$\begin{aligned} z_1 &= 0.45 z_1 + 0.15 z_2 \\ z_2 &= 0.50 z_1 + 0.65 z_2 + 0.50 z_3 \end{aligned}$$

Let (momentarily) $z_1 = 1$, we have

$$z_2 = 0.55/0.15 = 3.\overline{6}$$

Substitute $z_1$ and $z_2$ into the second equation

We get,

$$z_3 = 2\big[-0.5 + 0.35(0.55/0.15)\big] = 1.5\overline{6}$$

Thus, the computed solution is $(1, 3.\overline{6}, 1.5\overline{6})$

- It must be normalised to sum to 1
- We divide it by $z_1 + z_2 + z_3$

$$\rightsquigarrow \quad z \approx (0.160, 0.588, 0.250)$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Normalisation converts the system of equations with singular coefficient matrix into another one with non-singular coefficient matrix and non-zero RHS

$$(z_1, z_2, z_3)\begin{pmatrix} -0.55 & 0.50 & 1 \\ 0.15 & -0.35 & 1 \\ 0.00 & 0,50 & 1 \end{pmatrix} = (0, 0, 1)$$

The third equation has been replaced by the normalisation equation

The resulting system has a unique, non-zero solution

$\rightsquigarrow$ The stationary distribution of the chain

$$\begin{aligned} zP &= (0.160, 0.588, 0.250)\begin{pmatrix} 0.45 & 0.50 & 0.05 \\ 0.15 & 0.65 & 0.20 \\ 0.00 & 0.50 & 0.50 \end{pmatrix} \\ &= (0.160, 0.588, 0.250) \\ &= z \end{aligned}$$

■

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Finite and irreducible Markov chains have a unique stationary distribution

- No proper subset of states that is closed
- That is, no absorbing states

We replaced one of the equations in $z(P - I) = 0$ by the closure

$\rightsquigarrow$ A non-singular coefficient matrix and non-zero RHS

$\rightsquigarrow$ The resulting solution is therefore unique

## Distributions (cont.)



We now study the existence of $\lim_{n\to\infty} \pi(n) = \lim_{n\to\infty} \pi(0)P^{(n)}$

- For some selected initial probability distribution $\pi(0)$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

**Definition**

***Limiting distribution***

*Let $P$ be the transition probability matrix of a discrete-time Markov chain*

*Let $\pi(0)$ be an initial probability distribution*

*Suppose the existence of the limit*

$$\lim_{n\to\infty} P^{(n)} = \lim_{n\to\infty} P^n$$

*Then, it also exists the probability distribution $\pi$*

$$\pi = \lim_{n\to\infty} \pi(n) = \lim_{n\to\infty} \pi(0)P^{(n)} = \pi(0) \lim_{n\to\infty} P^{(n)} = \pi(0) \lim_{n\to\infty} P^n$$

*This distribution is called the* ***limiting distribution*** *of the Markov chain*

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Let the states of the chain be positive recurrent and aperiodic (ergodic)

- $\rightsquigarrow$ In this case, the limiting distribution always exists
- $\rightsquigarrow$ Moreover, it is unique

The result also holds for finite, irreducible and aperiodic chains

- $\rightsquigarrow$ (As the first two properties imply positive recurrence)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

**Example**

Let $P$ be the transition matrix of a discrete-time Markov chain

$$P = \begin{pmatrix} 0.8 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

We have already computed $\lim_{n\to\infty} P^{(n)} = \lim_{n\to\infty} P^n$

$$\lim_{n\to\infty} P^n = \begin{pmatrix} 0.7625 & 0.16875 & 0.06875 \\ 0.7625 & 0.16875 & 0.06875 \\ 0.7625 & 0.16875 & 0.06875 \end{pmatrix}$$

The elements in a probability vector are in the interval $[0, 1]$ and sum to 1

Multiply $\lim_{n\to\infty} P^n$ by any probability vector to get

$$\rightsquigarrow \quad \pi = (0.7625, 0.16875, 0.06875)$$

This is the limiting distribution of the chain

## Distributions (cont.)

Limiting distribution may exist when the transition matrix is reducible

- All the states need be positive recurrent



Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

**Example**

Let $P$ be the transition matrix of a discrete-time Markov chain

$$P = \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

We can compute $\lim_{n\to\infty} P^{(n)} = \lim_{n\to\infty} P^n$

$$\lim_{n\to\infty} P^{(n)} = \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$



Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Consider the following (initial) probability vectors

$$\begin{aligned} &\rightsquigarrow (1, 0, 0, 0) \\ &\rightsquigarrow (0, 0, 0.5, 0.5) \\ &\rightsquigarrow (\alpha, 1-\alpha, 0, 0), \ \text{(for } 0 \leq \alpha \leq 1) \\ &\rightsquigarrow (0.375, 0.375, 0.125, 0.125) \end{aligned}$$

All satisfy the necessary conditions for a limiting probability distribution

$$\begin{aligned} (1, 0, 0, 0) \lim_{n\to\infty} P^{(n)} &= (0.5, 0.5, 0, 0) \\ (0, 0, 0.5, 0.5) \lim_{n\to\infty} P^{(n)} &= (0, 0, 0.5, 0.5) \\ (\alpha, 1-\alpha, 0, 0) \lim_{n\to\infty} P^{(n)} &= (0.5, 0.5, 0, 0), \ \text{(for } 0 \leq \alpha \leq 1) \\ (0.375, 0.375, 0.125, 0.125) \lim_{n\to\infty} P^{(n)} &= (0.375, 0.375, 0.125, 0.125) \end{aligned}$$

■



Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Aperiodicity is a necessary property

## Distributions (cont.)

**Example**

Let $P$ be the transition matrix of a discrete-time Markov chain

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

It is easy to observe that $\lim_{n\to\infty} P^n$ does not exist

- $\rightsquigarrow$ The chain has no limiting distribution
- $\rightsquigarrow$ Successive powers of $P$ oscillate

The successive powers of $P$

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}}_{P} \to \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{P^2} \to \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{P^3} \to \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}}_{P^4 = P} \to \cdots$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

**Definition**

***Steady-state distribution***

*Let $P$ be the transition probability matrix of a discrete-time Markov chain*

*A limiting distribution $\pi$ is a steady-state distribution if it converges to a vector whose components are strictly positive and sum to one (probability)*

- $\rightsquigarrow$ $\pi_i > 0$ *for all states $i$, and* $\sum \pi_i = 1$

*This must be true independently of the initial distribution $\pi(0)$*

*If a steady-state distribution exists, it is unique*

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

**Example**

Let $P$ be the transition matrix of a discrete-time Markov chain

$$P = \begin{pmatrix} 0.4 & 0.6 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

Different initial distributions gave different limiting distributions

$$(1,0,0,0) \lim_{n\to\infty} P^{(n)} = (0.5, 0.5, 0, 0)$$
$$(0,0,0.5,0.5) \lim_{n\to\infty} P^{(n)} = (0, 0, 0.5, 0.5)$$
$$(\alpha, 1-\alpha, 0, 0) \lim_{n\to\infty} P^{(n)} = (0.5, 0.5, 0, 0), \text{ (for } 0 \leq \alpha \leq 1)$$
$$(0.375, 0.375, 0.125, 0.125) \lim_{n\to\infty} P^{(n)} = (0.365, 0.375, 0.125, 0.125)$$

This Markov chain has no limiting distribution

- Thus, it has no steady-state distribution

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Component $i$ in the steady-state distribution of a Markov chain is understood as **long-term proportion** of time the process spends in state $i$

This is equivalent to the probability that a random observer sees the Markov chain in state $i$, after the process has evolved over a long period of time

Existence of the steady-state distribution implies two convergences

- Vector $\pi(n)$
- Matrix $P^{(n)}$

They must converge independently of the initial distribution $\pi(0)$

## Distributions (cont.)

Steady state distributions are often given different names

- ⇝ **Equilibrium distributions**
- ⇝ **Long-run distributions**



Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Consider a Markov chain with a steady-state distribution $\pi$

- This is also the unique **stationary distribution**

We have,

$$\rightsquigarrow \quad \pi = \pi(0) \lim_{n \to \infty} P^{(n)} = \pi(0) \lim_{n \to \infty} P^{(n+1)}$$
$$= [\pi(0) \lim_{n \to \infty} P^{(n)}]P = \pi P$$

This implies that $\pi = \pi P$, thus $\pi$ is the (unique) stationary vector

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Equation $\pi = P\pi$ is called the **global balance equation** of the process

- ⇝ It is understood as if it equates the states in- and out-flow
- ⇝ (From queueing literature)

Consider the $i$-th equation

$$\pi_i = \sum_{\text{all } j} \pi_j p_{ji}$$

It can be re-written

$$\rightsquigarrow \quad \pi_i(1 - p_{ii}) = \sum_{j, j \neq i} \pi_j p_{ji}$$

Or,

$$\rightsquigarrow \quad \pi_i \sum_{j, j \neq i} p_{ij} = \sum_{j, j \neq i} \pi_j p_{ji}$$

- ⇝ The LHS, total flow from state $i$ into states $j$ other than $i$
- ⇝ The RHS, total flow from all states $j \neq i$ into state $i$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

**Example**

Let $P$ be the transition matrix of a discrete-time Markov chain

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \text{ (for } 0 < p < 1 \text{ and } 0 < q < 1)$$

We are interested in the steady-state distribution of the chain

- ⇝ The process $\{X_n, n \geq 0\}$ has two states (say, 0 and 1)

Let $\pi_0(0) = \text{Prob}\{X_0 = 0\}$ be the probability that the chain begins in 0

- We know that $\text{Prob}\{X_0 = 1\} = 1 - \pi_0(0) = \pi_1(0)$

We are interested in the distribution at time step $n$

## Distributions (cont.)

We have,

$$\rightsquigarrow \quad \text{Prob}\{X_{n+1}=0\} = \text{Prob}\{X_{n+1}=0|X_n=0\}\text{Prob}\{X_n=0\} + \text{Prob}\{X_{n+1}=0|X_n=1\}\text{Prob}\{X_n=1\}$$

Let $n=0$

$$\rightsquigarrow \quad \text{Prob}\{X_1=0\}$$
$$= \text{Prob}\{X_1=0|X_0=0\}\text{Prob}\{X_0=0\}+\text{Prob}\{X_1=0|X_0=1\}\text{Prob}\{X_0=1\}$$
$$= (1-p)\pi_0(0) + q[1-\pi_0(0)]$$
$$= (1-p-q)\pi_0(0) + q$$

Let $n=1$

$$\rightsquigarrow \quad \text{Prob}\{X_2=0\}$$
$$= \text{Prob}\{X_2=0|X_0=0\}\text{Prob}\{X_1=0\}+\text{Prob}\{X_2=0|X_1=1\}\text{Prob}\{X_1=1\}$$
$$= (1-p)[(1-p-q)\pi_0(0)+q] + q[1-(1-p-q)\pi_0(0)-q]$$
$$= (1-p-q)[(1-p-q)\pi_0(0)+q]+q$$
$$= (1-p-q)^2\pi_0(0) + (1-p-q)q + q$$



Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

We may suspect that for $n=n$, we could get

$$\text{Prob}\{X_n=0\} = (1-p-q)^n\pi_0(0) + q\sum_{j=0}^{n-1}(1-p-q)^j$$
$$\text{Prob}\{X_n=1\} = -\text{Prob}X_n = 0$$

$\sum_{j=0}^{n-1}(1-p-q)^j = \dfrac{1-(1-p-q)^n}{p+q}$ is the sum of a finite geometric series

Thus,

$$\rightsquigarrow \quad \text{Prob}\{X_n=0\} = (1-p-q)^n\pi_0(0) + \frac{q}{p+q} - q\frac{(1-p-q)^n}{p+q}$$
$$= \frac{q}{p+q} + (1-p-q)^n\Big[\pi_0(0) - \frac{q}{p+q}\Big]$$

And,

$$\rightsquigarrow \quad \text{Prob}\{X_n=1\} = \frac{p}{p+q} + (1-p-q)^n\Big[\pi_1(0) - \frac{p}{p+q}\Big]$$

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

We can now determine the limiting distribution, as $n\to\infty$

$$\lim_{n\to\infty}\text{Prob}\{X_n=0\} = \frac{q}{p+q}$$
$$\lim_{n\to\infty}\text{Prob}\{X_n=1\} = \frac{p}{p+q}$$

Moreover,

$$\lim_{n\to\infty}P^n = \lim_{n\to\infty}\begin{pmatrix}1-p & p\\ q & 1-q\end{pmatrix}^n = \frac{1}{p+q}\begin{pmatrix}q & p\\ q & p\end{pmatrix}$$

We used

$$\frac{1}{p+q}\begin{pmatrix}q & p\\ q & p\end{pmatrix}\begin{pmatrix}1-p & p\\ q & 1-q\end{pmatrix} = \frac{1}{p+q}\begin{pmatrix}q & p\\ q & p\end{pmatrix}$$

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Let $\alpha\in[0,1]$

We have that $(\alpha, 1-\alpha)$ times $\lim_{n\to\infty}P^n$ equals $[q/(p+q), p(p+q)]$

That is,

$$(\alpha, 1-\alpha)\begin{pmatrix}q/(p+q) & p/(p+q)\\ q/(p+q) & p/(p+q)\end{pmatrix} = [q/(p+q), p/(p+q)]$$

Hence, $[q/(p+q), p/(p+q)]$ is the unique steady-state distribution

## Distributions (cont.)

This fact can also be verified differently

The following must hold

$$[q/(p+q), p/(p+q)] \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} = [q/(p+q), p(p+q)]$$

■



## Distributions (cont.)

The unique stationary distribution and limiting distribution may coincide

- This is happens in some cases

In some other cases, the chain may possess only a stationary distribution

- No limiting distribution

We explore various classes of Markov chains



## Distributions (cont.)

**Irreducible Markov chains**: **Null recurrent or transient**

For such Markov chains there is no stationary probability vector

System $z = zP$ has only the trivial solution

- $z$, all components are zero

If a limiting distribution exists, its components are zeros

■



## Distributions (cont.)

**Irreducible Markov chains**: **Positive recurrent**

The system of equations $z = zP$ has a unique and strictly positive solution

The solution is the stationary probability distribution $\pi$

$$\rightsquigarrow \quad \pi_j = 1/M_{jj} \tag{6}$$

$M_{jj}$ is the mean recurrence time of $j$ (finite, for positive-recurrent states)

Multiply both sides of $M = E + P[M - \text{diag}(M)]$ by $\pi$

$$\begin{aligned} \rightsquigarrow \quad \pi M &= \pi E + \pi P[M - \text{diag}(M)] \\ &= e^T + \pi[M - \text{diag}(M)] \\ &= e^T + \pi M - \pi\text{diag}(M) \end{aligned}$$

Thus, $\pi\text{diag}(M) = e^T$

## Distributions (cont.)

Consider an irreducible chain with unique stationary probability vector

- The states are positive-recurrent

Consider an irreducible and positive-recurrent Markov chain

- It does not necessarily have a limiting distribution

■

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

**Example**

Consider an irreducible positive-recurrent discrete-time Markov chain

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

We are interested in the stationary and the limiting distributions

Vector $(1/4, 1/4, 1/4, 1/4)$ is the unique stationary distribution

The chain does not have a limiting distribution

- Whatever the initial distribution

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

This Markov chain is also periodic, with period $p = 4$

- If in state 1 at time $n$ it will transition to state 2 at time $n+1$, to state 3 at time $n+2$, to state 4 to state $n+3$ and back to state 1 at time $n+4$
- This process never stabilises to a limiting distribution, it will alternate
- $\lim_{n\to\infty} P^{(n)}$ does not exist

A unique distribution does not imply the existence of a limiting distribution

■

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

**Irreducible and aperiodic Markov chains**

We have two possible situations

States of irreducible and aperiodic discrete-time chains are all null-recurrent

$$\rightsquigarrow \quad \pi_j = 0 \text{ (for all } j)$$

The limiting distribution always exists but no stationary distribution

- Moreover, it is independent of the initial distribution

The state space must be infinite

## Distributions (cont.)

States of irreducible and aperiodic chains are all positive-recurrent (ergodic)

$$\rightsquigarrow \quad \pi_j > 0 \text{ (for all } j)$$

Probabilities $\pi_j$ define a stationary distribution

- They are uniquely determined

$$\rightsquigarrow \quad \pi_j = \sum_{\text{all } j} \pi_j p_{ij}, \quad \sum_j \pi_j = 1$$

In matrix notation,

$$\rightsquigarrow \quad \pi = \pi P, \quad \pi e = 1$$



Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

### Irreducible and ergodic Markov chains

States of ergodic discrete-time chains are all positive-recurrent and ergodic

The distribution $\pi(n)$ converges to a limiting distribution $\pi$

- Because of irreducibility

The limiting (steady-state) distribution is the unique stationary distribution

Thus, from the probability that the process is in state $i$ at step $n$

$$\pi_n(n) = \sum_{\text{all } k} p_{ki}^{(n)} \pi_k(0)$$

$$\rightsquigarrow \quad \pi_j(n+1) = \sum_{\text{all } i} p_{ij} \pi_i(n)$$

It follows that for $n \to \infty$ on both sides

$$\rightsquigarrow \quad \pi_j = \sum_{\text{all } j} p_{ij} \pi_j$$



Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

In matrix notation, we have

$$\pi = \pi P, \text{ (with } \pi > 0 \text{ and } ||\pi||_1 = 1) \tag{7}$$

It can be shown that, for $n \to \infty$, the rows of the $n$-step transition matrix $P^{(n)} = P^n$ all converge to the exact same vector of stationary probabilities

By letting $p_{ij}^{(n)}$ be the $(i,j)$-th element of $P^{(n)}$, we get

$$\pi_j = \lim_{n \to \infty} p_{ij}^{(n)} \text{ (for all } i \text{ and } j)$$

(We saw that with the simplified weather model)



Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Some performance quantities from steady-state probability vectors

- (Irreducible and ergodic Markov processes, only)

$\rightsquigarrow$ Mean time spent in state $j$ in a fixed interval $\tau$ at steady-state

$$\rightsquigarrow \quad v_j(\tau) = \pi_j \tau$$

The steady state probability $\pi_i$ is the portion of time in state $j$, averaged over the long-run

$\rightsquigarrow$ Mean number of steps between successive hits to state $j$, $1/\pi_j$

$\rightsquigarrow$ Mean time spent in state $i$ at steady-state and between two successive hits to state $j$

$$\rightsquigarrow \quad v_{ij} = \pi_i / \pi_j$$

## Distributions (cont.)

**Irreducible and periodic Markov chains**

We investigate the effect of periodicity in seeking limiting distributions

- (And high order powers of single-step transition matrix)

Consider a irreducible discrete-time Markov chain

Let the chain have period/index $p$

The number of single-step transitions needed to return, by any path, on some state after leaving it is a multiple of some integer $p > 1$



## Distributions (cont.)

For such chains we can permute rows and columns of the transition matrix

⇝ The **normal form**

This corresponds to partitioning and ordering of the states in $p$ subsets

- The **cyclic classes** of the chain

$$P = \begin{pmatrix} 0 & P_{12} & 0 & \cdots & 0 \\ 0 & 0 & P_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & P_{p-1,p} \\ P_{p1} & 0 & 0 & \cdots & 0 \end{pmatrix} \qquad (8)$$

The only non-zero sub-matrices are $P_{12}, P_{23}, \dots, P_{p1}$

- Diagonal sub-matrices $P_{ii}$ are all square

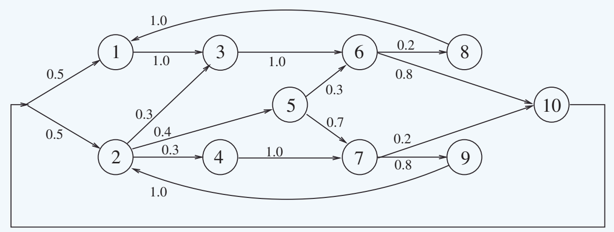
According to this ordering, a chain in a state from subset $i$ exits this subset in the next time step and then it enters a state from subset $(i \bmod p) + 1$



## Distributions (cont.)

**Example**

Consider the transition diagram of a discrete-time Markov chain



The states have been ordered according to their respective classes

The process has period $p = 4$ and four cyclic classes

⇝ $C_1 = \{1, 2\}$

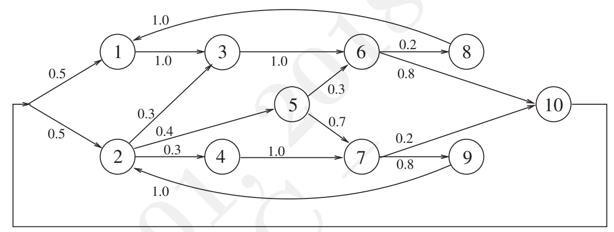
⇝ $C_2 = \{3, 4, 5\}$

⇝ $C_3 = \{6, 7\}$

⇝ $C_4 = \{8, 9, 10\}$



## Distributions (cont.)



From states in class $C_i$ the chain can only go to states in class $C_{(i \bmod p)+1}$

⇝ $C_1 = \{1, 2\}$

⇝ $C_2 = \{3, 4, 5\}$

⇝ $C_3 = \{6, 7\}$

⇝ $C_4 = \{8, 9, 10\}$

The process will return to initial state only after $4 \cdot n$ steps $(n = 1, 2, \cdots)$

## Distributions (cont.)

Consider the transition matrix $P$ of the process

$$P = \begin{pmatrix} 0 & 0 & 1.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.3 & 0.4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.3 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 0.8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.8 & 0.2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Suppose that our interest is in the behaviour of the chain at time $n \to \infty$

We wish to investigate the behaviour of $P^n$, as $n \to \infty$, the existence of the limiting distribution and the existence of a stationary distribution

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Consider the transition matrix $P$, with $p = 4$ cycles

$$P = \begin{pmatrix} 0 & 0 & 1.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.3 & 0.4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.3 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 0.8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.8 & 0.2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & P_{12} & 0 & 0 \\ 0 & 0 & P_{23} & 0 \\ 0 & 0 & 0 & P_{34} \\ P_{41} & 0 & 0 & 0 \end{pmatrix}$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

By taking successive powers of $P$, we obtain

$$P^2 = \begin{pmatrix} 0 & 0 & P_{12}P_{23} & 0 \\ 0 & 0 & 0 & P_{23}P_{34} \\ P_{34}P_{41} & 0 & 0 & 0 \\ 0 & P_{41}P_{12} & 0 & 0 \end{pmatrix}$$

$$P^3 = \begin{pmatrix} 0 & 0 & 0 & P_{12}P_{23}P_{34} \\ P_{23}P_{34}P_{41} & 0 & 0 & 0 \\ 0 & P_{34}P_{41}P_{12} & 0 & 0 \\ 0 & 0 & P_{41}P_{12}P_{23} & 0 \end{pmatrix}$$

$$P^4 = \begin{pmatrix} P_{12}P_{23}P_{34}P_{41} & 0 & 0 & 0 \\ 0 & P_{23}P_{34}P_{41}P_{12} & 0 & 0 \\ 0 & 0 & P_{34}P_{41}P_{12}P_{23} & 0 \\ 0 & 0 & 0 & P_{41}P_{12}P_{23}P_{34} \end{pmatrix}$$

After four steps the transition matrix has taken a block-diagonal form

- $\rightsquigarrow$ Moreover, each individual block is row-stochastic

Each block represents an irreducible, recurrent and aperiodic chain

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Important matrices
Potential and fundamental matrices
Reachability matrix
Distributions

## Distributions (cont.)

Consider beginning in state $C_1$

- $\rightsquigarrow$ A state $C_2$ may be reached in one step
- $\rightsquigarrow$ A state $C_3$ may be reached in two steps
- $\rightsquigarrow$ A state $C_4$ may be reached in three steps

After four steps, the Markov chain is back to state $C_1$

- $\rightsquigarrow$ At any time $4n$, with $n = 1, 2, \dots$

Each block represents the transition probability matrix

- $\rightsquigarrow$ An irreducible, recurrent and aperiodic chain

(Aperiodic as one single-step in the new chain is four original steps)

## Distributions (cont.)

We can apply the theory of irreducible, recurrent and aperiodic chains

- Each block can be treated individually, as $n \to \infty$

We have that for each block $\lim_{n\to\infty} P^{4n}$ exists

$$P^4 = \begin{pmatrix}
.60 & .40 & & & & & & & & \\
.31 & .69 & & & & & & & & \\
& & .72 & .120 & .16 & & & & & \\
& & .37 & .270 & .36 & & & & & \\
& & .75 & .225 & .30 & & & & & \\
& & & & & .768 & .232 & & & \\
& & & & & .478 & .522 & & & \\
& & & & & & & .200 & .000 & .800 \\
& & & & & & & .080 & .464 & .452 \\
& & & & & & & .142 & .232 & .626
\end{pmatrix}$$

In addition, we can compute $\lim_{n\to\infty} P^{4n}$





## Distributions (cont.)

That is,

$$\lim_{n\to\infty} P^{4n} = \begin{pmatrix}
.437 & .563 & & & & & & & & \\
.437 & .563 & & & & & & & & \\
& & .606 & .17 & .224 & & & & & \\
& & .606 & .17 & .224 & & & & & \\
& & .606 & .17 & .224 & & & & & \\
& & & & & .673 & .327 & & & \\
& & & & & .673 & .327 & & & \\
& & & & & & & .135 & .261 & .604 \\
& & & & & & & .135 & .261 & .604 \\
& & & & & & & .135 & .261 & .604
\end{pmatrix}$$

Each diagonal block can be treated as a transition probability matrix

- A finite, aperiodic and irreducible Markov process

The limiting distribution coincides with the stationary distribution

- It can be computed for each block individually ($n \to \infty$)





## Distributions (cont.)

By concatenating the four stationary distributions yields a vector $z = zP$

We obtain the vector

$$z = (0.4366, 0.5634, 0.6056, 0.1690, 0.2254, 0.6732, 0.3268, 0.1367, 0.2614, 0.6039)$$

The stationary distribution is obtained by normalisation

$$(0.1092, 0.1408, 0.1514, 0.0423, 0.0563, 0.1683, 0.0817, 0.0337, 0.0654, 0.1510)$$

The result assumes that the chain spends equal time in all periodic classes

⇝ This is correct

Irreducible, positive-recurrent chains has a strictly positive distribution

⇝ The distribution is unique (the only possible one)

■





## Distributions (cont.)

### Reducible Markov chains

We consider in some detail chains that are reducible

⇝ Multiple transient and irreducible closed classes

Such chains possess multiple stationary distributions

Also, any linear combination of the stationary distributions is stationary

⇝ This allows to treat each irreducible closed class individually

## Distributions (cont.)

**Example**

Consider the discrete-time Markov chain with transition matrix $P$

$$P = \begin{pmatrix} 0.4 & 0.2 & 0.0 & 0.2 & 0.0 & 0 & 0.0 & 0.2 \\ 0.3 & 0.3 & 0.0 & 0.0 & 0.1 & 0 & 0.2 & 0.1 \\ 0.0 & 0.0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0.0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

These chain has 3 transient states and 3 irreducible classes

- $\rightsquigarrow$ Transient states (1, 2 and 3)
- $\rightsquigarrow$ Irreducible classes ($\{4,5\}$, $\{6\}$ and $\{7,8\}$)



## Distributions (cont.)

The stationary distributions of the three irreducible classes is unique

$$(0.625, 0.375) = (0.625, 0.375) \begin{pmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{pmatrix}$$

$$(1.000) = (1.000)(1)$$

$$(0.500, 0.500) = (0.500, 0.500) \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}$$

These distributions can be filled with zeros to produce a vector length 8

- This vector is a stationary distribution for the process



## Distributions (cont.)

$$(0.625, 0.375) = (0.625, 0.375) \begin{pmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{pmatrix}$$

Consider the case of the first stationary distribution

$$\rightsquigarrow \quad (0, 0, 0, 0.625, 0.375, 0, 0, 0)$$

$$\times \begin{pmatrix} 0.4 & 0.2 & 0 & 0.2 & 0 & 0 & 0 & 0.2 \\ 0.3 & 0.3 & 0 & 0 & 0.1 & 0 & 0.2 & 0.1 \\ 0 & 0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

$$= (0, 0, 0, 0.625, 0.375, 0, 0, 0)$$

The same applies to the other two distributions



## Distributions (cont.)

A linear combination of the three distributions is a stationary distribution

- $\rightsquigarrow$ (Upon normalisation)

Consider adding up all three distributions and normalising

$$\rightsquigarrow \quad \underbrace{(0, 0, 0, 0.6250, 0.3750, 1.0000, 0.5000, 0.5000)}_{\text{unnormalised}}$$

$$\rightsquigarrow \quad \underbrace{(0, 0, 0, 0.2083, 0.1250, 0.3333, 0.1667, 0.1667)}_{\text{normalised}}$$

## Distributions (cont.)

Consider the powers of the transition probability matrix $P$

$$(0, 0, 0, 0.2083, 0.1250, 0.3333, 0.1667, 0.1667)$$
$$\times \begin{pmatrix} 0.4 & 0.2 & 0 & 0.2 & 0 & 0 & 0 & 0.2 \\ 0.3 & 0.3 & 0 & 0 & 0.1 & 0 & 0.2 & 0.1 \\ 0 & 0 & 0.1 & 0.3 & 0.1 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.7 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$
$$= (0, 0, 0, 0.2083, 0.1250, 0.3333, 0.1667, 0.1667)$$



## Distributions (cont.)

In the limit $n \to \infty$, we get

$$\lim_{n\to\infty} P^n = \begin{pmatrix} 0 & 0 & 0 & 0.2778 & 0.1667 & 0 & 0.2778 & 0.2778 \\ 0 & 0 & 0 & 0.2083 & 0.1250 & 0 & 0.3334 & 0.3333 \\ 0 & 0 & 0 & 0.2788 & 0.1667 & 0 & 0.2778 & 0.2778 \\ 0 & 0 & 0 & 0.6250 & 0.3750 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.6250 & 0.3750 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{pmatrix}$$

Not all the rows are equal, yet they all are stationary distributions

- ⇝ They equal some linear combination of the three solutions
- ⇝ (Need re-normalisation due to rounding approximations)



## Distributions (cont.)

Transition probabilities into transient states

$$\rightsquigarrow \quad \lim_{n\to\infty} p_{ij}^{(n)} = 0 \text{ (zero)}$$

Transition probabilities from ergodic states to the same state

$$\rightsquigarrow \quad \lim_{n\to\infty} p_{jj}^{(n)} > 0 \text{ (strictly positive)}$$

Transition probabilities from any state $i$ into an ergodic state $j$

$$\lim_{n\to\infty} p_{ij}^{(n)} = f_{ij} \lim_{n\to\infty} p_{jj}^{(n)}$$

(We must first compute elements $f_{ij}$ of reachability matrix $F$)



## Distributions (cont.)

The elements $f_{ij}$ of the reachability matrix $F$ have been calculated earlier

$$\lim_{n\to\infty} p_{14}^{(n)} = 0.4444 \cdot 0.6250 = 0.2778$$
$$\lim_{n\to\infty} p_{28}^{(n)} = 0.6667 \cdot 0.5000 = 0.3333$$
$$\lim_{n\to\infty} p_{35}^{(n)} = 0.4444 \cdot 0.3750 = 0.1667$$

... and so on

■