

Discrete Time Markov Chains, Limiting Distribution and Classification

Bo Friis Nielsen¹

¹DTU Informatics

02407 Stochastic Processes 3, September 14 2021



Today:

- ▶ Discrete time Markov chains - invariant probability distribution
- ▶ Classification of states
- ▶ Classification of chains

Next week

- ▶ Poisson process

Two weeks from now

- ▶ Birth- and Death Processes



Regular Transition Probability Matrices

$$P = \{P_{ij}\}, \quad 0 \leq i, j \leq N$$

Regular: If $P^k > 0$ for some k

In that case $\lim_{n \rightarrow \infty} P_{ij}^{(n)} = \pi_j$

Theorem 4.1 (Page 168) let P be a regular transition probability matrix on the states $0, 1, \dots, N$. Then the limiting distribution $\pi = (\pi_0, \pi_1, \pi_N)$ is the unique nonnegative solution of the equations

$$\pi_j = \sum_{k=0}^N \pi_k P_{kj}, \quad \pi = \pi P$$

$$\sum_{k=0}^N \pi_k = 1, \quad \pi \mathbf{1} = 1$$



Interpretation of π_j 's

- ▶ Limiting probabilities $\lim_{n \rightarrow \infty} P_{ij}^{(n)} = \pi_j$
- ▶ Long term averages $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n P_{ij}^{(m)} = \pi_j$
- ▶ Stationary distribution $\pi = \pi P$



A Social Mobility Example

		Son's Class		
		Lower	Middle	Upper
Father's Class	Lower	0.40	0.50	0.10
	Middle	0.05	0.70	0.25
	Upper	0.05	0.50	0.45

$$P^8 = \begin{pmatrix} 0.0772 & 0.6250 & 0.2978 \\ 0.0769 & 0.6250 & 0.2981 \\ 0.0769 & 0.6250 & 0.2981 \end{pmatrix}$$

$$\pi_0 = 0.40\pi_0 + 0.05\pi_1 + 0.05\pi_2$$

$$\pi_1 = 0.50\pi_0 + 0.70\pi_1 + 0.50\pi_2$$

$$\pi_2 = 0.10\pi_0 + 0.25\pi_1 + 0.45\pi_2$$

$$1 = \pi_0 + \pi_1 + \pi_2$$



Classification of Markov chain states

- ▶ States which cannot be left, once entered - absorbing states
- ▶ States where the return some time in the future is certain - recurrent or persistent states
 - ▶ The mean time to return can be
 - ▶ finite - positive recurrence/non-null recurrent
 - ▶ infinite - null recurrent
- ▶ States where the return some time in the future is uncertain - transient states
- ▶ States which can only be visited at certain time epochs - periodic states



Classification of States

- ▶ j is **accessible** from i if $P_{ij}^{(n)} > 0$ for some n
- ▶ If j is accessible from i and i is accessible from j we say that the two states **communicate**
- ▶ Communicating states constitute equivalence classes (an equivalence relation)
 - ▶ i communicates with j and j communicates with k then i and k communicates



First passage and first return times

We can formalise the discussion of state classification by use of a certain class of probability distributions - first passage time distributions. Define the first passage probability

$$f_{ij}^{(n)} = \mathbb{P}\{X_1 \neq j, X_2 \neq j, \dots, X_{n-1} \neq j, X_n = j | X_0 = i\}$$

This is the probability of reaching j for the *first* time at time n having started in i .

What is the probability of ever reaching j ?

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)} \leq 1$$

The probabilities $f_{ij}^{(n)}$ constitute a probability distribution. On the contrary we cannot say anything in general on $\sum_{n=1}^{\infty} p_{ij}^{(n)}$ (the n -step transition probabilities)



- ▶ A state is recurrent (persistent) if $f_{ii} = \left(\sum_{n=1}^{\infty} f_{ii}^{(n)} \right) = 1$
 - ▶ A state is positive or non-null recurrent if $E(T_i) < \infty$.
 $E(T_i) = \sum_{n=1}^{\infty} n f_{ii}^{(n)} = \mu_i$
 - ▶ A state is null recurrent if $E(T_i) = \mu_i = \infty$
- ▶ A state is transient if $f_{ii} < 1$.
In this case we define $\mu_i = \infty$ for later convenience.
- ▶ A periodic state has nonzero $p_{ij}(nk)$ for some k .
- ▶ A state is ergodic if it is positive recurrent and aperiodic.



- ▶ We can identify subclasses of states with the same properties
- ▶ All states which can mutually reach each other will be of the same type
- ▶ Once again the formal analysis is a little bit heavy, but try to stick to the fundamentals, definitions (concepts) and results



Properties of sets of intercommunicating states

- ▶ (a) i and j has the same period
- ▶ (b) i is transient if and only if j is transient
- ▶ (c) i is null persistent (null recurrent) if and only if j is null persistent



A set C of states is called

- ▶ (a) **Closed** if $p_{ij} = 0$ for all $i \in C, j \notin C$
- ▶ (b) **Irreducible** if $i \leftrightarrow j$ for all $i, j \in C$.

Theorem

Decomposition Theorem The state space S can be partitioned uniquely as

$$S = T \cup C_1 \cup C_2 \cup \dots$$

where T is the set of transient states, and the C_i are irreducible closed sets of persistent states □

Lemma

If S is finite, then at least one state is persistent (recurrent) and all persistent states are non-null (positive recurrent) □



Theorem 4.3 The basic limit theorem of Markov chains

- (a) Consider a recurrent irreducible aperiodic Markov chain. Let $P_{ij}^{(n)}$ be the probability of entering state i at the n th transition, $n = 1, 2, \dots$, given that $X_0 = i$. By our earlier convention $P_{ij}^{(0)} = 1$. Let $f_{ij}^{(n)}$ be the probability of first returning to state i at the n th transition $n = 1, 2, \dots$, where $f_{ij}^{(0)} = 0$. Then

$$\lim_{n \rightarrow \infty} P_{ij}^{(n)} = \frac{1}{\sum_{n=0}^{\infty} n f_{ij}^{(n)}} = \frac{1}{m_i}$$

- (b) under the same conditions as in (a), $\lim_{n \rightarrow \infty} P_{ji}^{(n)} = \lim_{n \rightarrow \infty} P_{ij}^{(n)}$ for all j .



$$P = \begin{bmatrix} 0.6 & 0.4 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.3 & 0.3 & 0.4 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.3 & 0.3 & 0.4 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.3 & 0.3 & 0.4 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.3 & 0.3 & 0.4 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.3 & 0.3 & 0.4 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.3 & 0.3 & 0.4 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.3 & 0.7 \end{bmatrix}$$

With initial probability distribution $\mathbf{p}^{(0)} = (1, 0, 0, 0, 0, 0, 0, 0)$ or $X_0 = 1$.

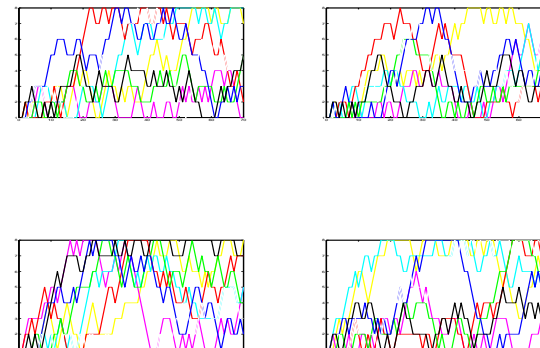


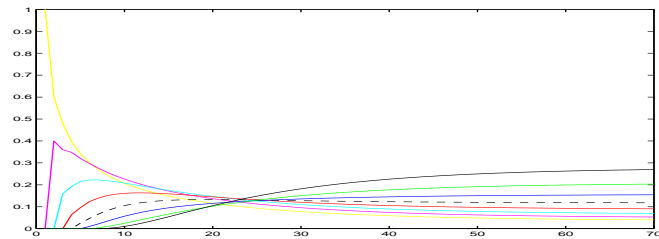
Properties of that chain

- ▶ We have a finite number of states
- ▶ From state 1 we can reach state j with a probability $f_{1j} \geq 0.4^{j-1}, j > 1$.
- ▶ From state j we can reach state 1 with a probability $f_{j1} \geq 0.3^{j-1}, j > 1$.
- ▶ Thus all states communicate and the chain is irreducible. Generally we won't bother with bounds for the f_{ij} 's.
- ▶ Since the chain is finite all states are positive recurrent
- ▶ A look on the behaviour of the chain



A number of different sample paths X_n 's





$p_j^{(n)}$



For an irreducible aperiodic chain, we have that

$$p_{ij}^{(n)} \rightarrow \frac{1}{\mu_j} \text{ as } n \rightarrow \infty, \text{ for all } i \text{ and } j$$

Three important remarks

- ▶ If the chain is transient or null-persistent (null-recurrent) $p_{ij}^{(n)} \rightarrow 0$
- ▶ If the chain is positive recurrent $p_{ij}^{(n)} \rightarrow \frac{1}{\mu_j}$
- ▶ The limiting probability of $X_n = j$ does not depend on the starting state $X_0 = i$



The stationary distribution

- ▶ A distribution that does not change with n
- ▶ The elements of $\mathbf{p}^{(n)}$ are all constant
- ▶ The implication of this is $\mathbf{p}^{(n)} = \mathbf{p}^{(n-1)}\mathbf{P} = \mathbf{p}^{(n-1)}$ by our assumption of $\mathbf{p}^{(n)}$ being constant
- ▶ Expressed differently $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$



Stationary distribution

Definition

The vector $\boldsymbol{\pi}$ is called a *stationary distribution* of the chain if $\boldsymbol{\pi}$ has entries $(\pi_j : j \in S)$ such that

- ▶ (a) $\pi_j \geq 0$ for all j , and $\sum_j \pi_j = 1$
- ▶ (b) $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$, which is to say that $\pi_j = \sum_i \pi_i p_{ij}$ for all j .

□

VERY IMPORTANT

An irreducible chain has a stationary distribution $\boldsymbol{\pi}$ if and only if all the states are non-null persistent (positive recurrent); in this case, $\boldsymbol{\pi}$ is the unique stationary distribution and is given by $\pi_i = \frac{1}{\mu_i}$ for each $i \in S$, where μ_i is the mean recurrence time of i .



The example chain (random walk with reflecting barriers)

$$P = \begin{bmatrix} 0.6 & 0.4 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.3 & 0.3 & 0.4 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.3 & 0.3 & 0.4 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.3 & 0.3 & 0.4 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.3 & 0.3 & 0.4 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.3 & 0.3 & 0.4 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.3 & 0.3 & 0.4 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.3 & 0.7 \end{bmatrix} \quad \pi = \pi P$$

Elementwise the matrix equation is $\pi_j = \sum_j \pi_j p_{ji}$

$$\begin{aligned} \pi_1 &= \pi_1 \cdot 0.6 + \pi_2 \cdot 0.3 \\ \pi_2 &= \pi_1 \cdot 0.4 + \pi_2 \cdot 0.3 + \pi_3 \cdot 0.3 \\ \pi_3 &= \pi_2 \cdot 0.4 + \pi_3 \cdot 0.3 + \pi_4 \cdot 0.3 \end{aligned}$$



The normalising condition

- ▶ We note that we don't have to use the last equation
- ▶ We need a solution which is a probability distribution

$$\sum_{j=1}^8 \pi_j = 1, \quad \sum_{j=1}^8 \left(\frac{0.4}{0.3}\right)^{j-1} \pi_1 = \pi_1 \sum_{k=0}^7 \left(\frac{0.4}{0.3}\right)^k$$

$$\sum_{i=0}^N a^i = \begin{cases} \frac{1-a^{N+1}}{1-a} & N < \infty, a \neq 1 \\ N+1 & N < \infty, a = 1 \\ \frac{1}{1-a} & N = \infty, |a| < 1 \end{cases}$$

Such that

$$1 = \pi_1 \frac{1 - \left(\frac{0.4}{0.3}\right)^8}{1 - \frac{0.4}{0.3}} \Leftrightarrow \pi_1 = \frac{1 - \frac{0.4}{0.3}}{1 - \left(\frac{0.4}{0.3}\right)^8}$$



$$\begin{aligned} \pi_1 &= \pi_1 \cdot 0.6 + \pi_2 \cdot 0.3 \\ \pi_j &= \pi_{j-1} \cdot 0.4 + \pi_j \cdot 0.3 + \pi_{j+1} \cdot 0.3 \\ \pi_8 &= \pi_7 \cdot 0.4 + \pi_8 \cdot 0.7 \end{aligned}$$

Or

$$\begin{aligned} \pi_2 &= \frac{1 - 0.6}{0.3} \pi_1 \\ \pi_{j+1} &= \frac{1}{0.3} ((1 - 0.3)\pi_j - 0.4\pi_{j-1}) \end{aligned}$$

Can be solved recursively to find:

$$\pi_j = \left(\frac{0.4}{0.3}\right)^{j-1} \pi_1$$



The solution of $\pi = \pi P$

- ▶ More or less straightforward, but one problem
- ▶ if \mathbf{x} is a solution such that $\mathbf{x} = \mathbf{x}P$ then obviously $(k\mathbf{x}) = (k\mathbf{x})P$ is also a solution.
- ▶ Recall the definition of eigenvalues/eigen vectors
- ▶ If $A\mathbf{y} = \lambda\mathbf{y}$ we say that λ is an eigenvalue of A with an associated eigenvector \mathbf{y} . Here \mathbf{y} is a right eigenvector, there is also a left eigenvector



The solution of $\pi = \pi P$ continued

- ▶ The vector π is a left eigenvector of P .
- ▶ The main theorem says that there is a unique eigenvector associated with the eigenvalue 1 of P
- ▶ In practice this means that we can only solve but a normalising condition
- ▶ But we have the normalising condition by $\sum_j \pi_j = 1$ this can be expressed as $\pi \mathbf{1} = 1$. Where

$$\mathbf{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$



Roles of the solution to $\pi = \pi P$

For an irreducible Markov chain, (the condition we need to verify)

- ▶ The stationary solution. If $\mathbf{p}^{(0)} = \pi$ then $\mathbf{p}^{(n)} = \pi$ for all n .
- ▶ The limiting distribution, i.e. $\mathbf{p}^{(n)} \rightarrow \pi$ for $n \rightarrow \infty$ (the Markov chain has to be aperiodic too). Also $\rho_{ij}^{(n)} \rightarrow \pi_j$.
- ▶ The mean recurrence time for state i is $\mu_i = \frac{1}{\pi_i}$.
- ▶ The mean number of visits in state j between two successive visits to state i is $\frac{\pi_j}{\pi_i}$.
- ▶ The long run average probability of finding the Markov chain in state i is π_i . $\pi_i = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \rho_i^{(k)}$ also true for periodic chains.



Example (null-recurrent) chain

$$P = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & \dots \\ 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

For $p_j > 0$ the chain is obviously irreducible. The main theorem tells us that we can investigate directly for $\pi = \pi P$.

$$\pi_1 = \pi_1 p_1 + \pi_2 \quad \pi_2 = \pi_1 p_2 + \pi_3 \quad \pi_j = \pi_1 p_j + \pi_{j+1}$$

$$\pi_1 = \pi_1 p_1 + \pi_2 \quad \pi_2 = \pi_1 p_2 + \pi_3 \quad \pi_j = \pi_1 p_j + \pi_{j+1}$$

we get

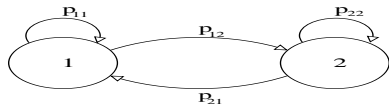
$$\pi_2 = (1 - p_1)\pi_1 \quad \pi_3 = (1 - p_1 - p_2)\pi_1 \quad \pi_j = (1 - p_1 \cdots - p_{j-1})\pi_1$$

$$\pi_j = (1 - p_1 \cdots - p_{j-1})\pi_1 \Leftrightarrow \pi_j = \pi_1 \left(1 - \sum_{i=1}^{j-1} p_i \right) \Leftrightarrow \pi_j = \pi_1 \sum_{i=j}^{\infty} p_i$$

Normalisation

$$\sum_{j=1}^{\infty} \pi_j = 1 \quad \sum_{j=1}^{\infty} \pi_1 \sum_{i=j}^{\infty} p_i = \pi_1 \sum_{i=1}^{\infty} \sum_{j=1}^i p_i = \pi_1 \sum_{i=1}^{\infty} i p_i$$



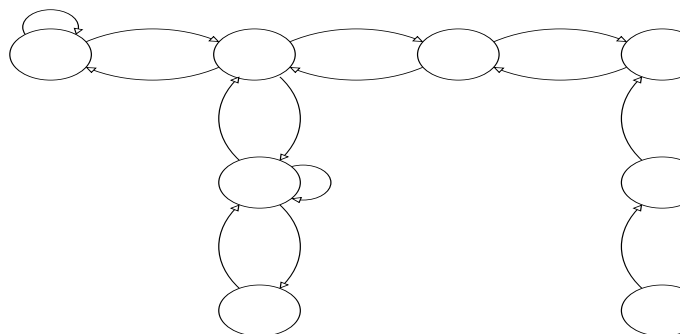


$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

- ▶ Solve sequence of linear equations instead of the whole system
- ▶ Local balance in probability flow as opposed to global balance
- ▶ Nice theoretical construction



Local balance equations



$$\pi_i = \sum_j \pi_j p_{ji} \quad \pi_i \cdot 1 = \sum_j \pi_j p_{ji} \quad \pi_i \sum_j p_{ij} = \sum_j \pi_j p_{ji}$$

$$\sum_j \pi_i p_{ij} = \sum_j \pi_j p_{ji}$$

Term for term we get

$$\pi_i p_{ij} = \pi_j p_{ji}$$

If they are fulfilled for each i and j , the global balance equations can be obtained by summation.



Why reversible?

$$\begin{aligned}\mathbb{P}\{X_{n-1} = i \cap X_n = j\} &= \mathbb{P}\{X_{n-1} = i\} \mathbb{P}\{X_n = j | X_{n-1} = i\} \\ &= \mathbb{P}\{X_{n-1} = i\} p_{ij}\end{aligned}$$

and for a stationary chain

$$\pi_i p_{ij}$$

For a reversible chain (local balance) this is $\pi_i p_{ij} = \pi_j p_{ji} = \mathbb{P}\{X_{n-1} = j\} \mathbb{P}\{X_n = i | X_{n-1} = j\} = \mathbb{P}\{X_{n-1} = j \cap X_n = i\}$ the reversed sequence.



Exercise 10 (16/12/91 ex.1)

In connection with an examination of the reliability of some software intended for use in control of modern ferries one is interested in examining a stochastic model of the use of a control program.

The control program works as "state machine" i.e. it can be in a number of different levels, 4 are considered here. The levels depend on the physical state of the ferry. With every shift of time unit while the program is run, the program will change from level j to level k with probability p_{jk} .



Another look at a similar question

$$\begin{aligned}\mathbb{P}\{X_{n-1} = j | X_n = i\} &= \frac{\mathbb{P}\{X_{n-1} = j \cap X_n = i\}}{\mathbb{P}\{X_n = i\}} \\ &= \frac{\mathbb{P}\{X_{n-1} = j\} \mathbb{P}\{X_n = i | X_{n-1} = j\}}{\mathbb{P}\{X_n = i\}} = \frac{\mathbb{P}\{X_{n-1} = j\} p_{ji}}{\mathbb{P}\{X_n = i\}}\end{aligned}$$

For a stationary chain we get

$$\frac{\pi_j p_{ji}}{\pi_i}$$

The chain is reversible if $\mathbb{P}\{X_{n-1} = j | X_n = i\} = p_{ji}$ leading to the local balance equations

$$p_{ij} = \frac{\pi_j p_{ji}}{\pi_i}$$



Two possibilities are considered:

The program has no errors and will run continuously shifting between the four levels.

The program has a critical error. In this case it is possible that the error is found, this happens with probability q_i , $i = 1, 2, 3, 4$ depending on the level.

The error will be corrected immediately and the program will from then on be without faults.

Alternatively the program can stop with a critical error (the ferry will continue to sail, but without control). This happens with probability r_i , $i = 1, 2, 3, 4$.

In general $q_i + r_i < 1$, a program with errors can thus work and the error is not necessarily discovered. It is assumed that detection of an error, as well as the appearance of a fault happens coincidentally with shift between levels.

The program starts running in level 1, and it is known that the program contains one critical error.



Solution: Question 1

Formulate a stochastic process (Markov chain) in discrete time describing this system.

The model is a discrete time Markov chain. A possible definition of states could be

- 0: The programme has stopped.
- 1-4: The programme is operating safely in level i .
- 5-8: The programme is operating in level $i-4$, the critical error is not detected.

The transition matrix \mathbf{A} is

$$\mathbf{A} = \begin{bmatrix} 1 & \vec{0} & \vec{0} \\ \vec{0} & \mathbf{P} & \mathbf{0} \\ \vec{r} & \text{Diag}(\mathbf{q}_i)\mathbf{P} & \text{Diag}(\mathbf{S}_i)\mathbf{P} \end{bmatrix}$$



Question 1 - continued

The model is a discrete time Markov chain. Where $\mathbf{P} = \{p_{ij}\}$

$$\vec{r} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} \quad \text{Diag}(\mathbf{S}_i) = \begin{bmatrix} S_1 & 0 & 0 & 0 \\ 0 & S_2 & 0 & 0 \\ 0 & 0 & S_3 & 0 \\ 0 & 0 & 0 & S_4 \end{bmatrix} \quad S_i = 1 - r_i - q_i$$

$$\text{Diag}(\mathbf{q}_i) = \begin{bmatrix} q_1 & 0 & 0 & 0 \\ 0 & q_2 & 0 & 0 \\ 0 & 0 & q_3 & 0 \\ 0 & 0 & 0 & q_4 \end{bmatrix}$$



Question 1 - continued

Or without matrix notation:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_{11} & p_{12} & p_{13} & p_{14} & 0 & 0 & 0 & 0 \\ 0 & p_{21} & p_{22} & p_{23} & p_{24} & 0 & 0 & 0 & 0 \\ 0 & p_{31} & p_{32} & p_{33} & p_{34} & 0 & 0 & 0 & 0 \\ 0 & p_{41} & p_{42} & p_{43} & p_{44} & 0 & 0 & 0 & 0 \\ r_1 & q_1 p_{11} & q_1 p_{12} & q_1 p_{13} & q_1 p_{14} & S_1 p_{11} & S_1 p_{12} & S_1 p_{13} & S_1 p_{14} \\ r_2 & q_2 p_{21} & q_2 p_{22} & q_2 p_{23} & q_2 p_{24} & S_2 p_{21} & S_2 p_{22} & S_2 p_{23} & S_2 p_{24} \\ r_3 & q_3 p_{31} & q_3 p_{32} & q_3 p_{33} & q_3 p_{34} & S_3 p_{31} & S_3 p_{32} & S_3 p_{33} & S_3 p_{34} \\ r_4 & q_4 p_{41} & q_4 p_{42} & q_4 p_{43} & q_4 p_{44} & S_4 p_{41} & S_4 p_{42} & S_4 p_{43} & S_4 p_{44} \end{bmatrix}$$



Solution question 2

Characterise the states in the Markov chain.

With reasonable assumptions on \mathbf{P} (i.e. irreducible) we get

State	0	1	2	3	4	5	6	7	8
	Absorbing	Positive recurrent	Positive recurrent	Positive recurrent	Positive recurrent	Transient	Transient	Transient	Transient



Solution question 3

We now consider the case where the stability of the system has been assured, i.e. the error has been found and corrected, and the program has been running for long time without errors. The parameters are as follows.

$$P_{i,i+1} = 0.6 \quad i = 1, 2, 3 \quad P_{i,i-1} = 0.2 \quad i = 2, 3, 4$$
$$P_{i,j} = 0 \quad |i - j| > 1 \quad q_i = 10^{-3i} \quad r_i = 10^{-3i-5}$$

Characterise the stochastic process, that describes the stable system.

The system becomes stable by reaching one of the states 1-4. The process is ergodic from then on. The process is a reversible ergodic Markov chain in discrete time.



Solution question 4

For what fraction of time will the system be in level 1. We obtain the following steady state equations

$$\pi_j = 3^{j-1} \pi_1$$

$$\sum_{i=1}^4 3^{i-1} \pi_1 = 1 \Leftrightarrow 40 \pi_1 = 1$$

$$\pi_1 = \frac{1}{40}$$

The sum $\sum_{i=1}^4 3^{i-1}$ can be obtained by using $\sum_{i=1}^4 3^{i-1} = \frac{1-3^4}{1-3} = 40$.

$$\sum_{i=1}^4 3^{i-1} \pi_1 = 1 \Leftrightarrow \frac{1-3^4}{1-3} \pi_1 = 1$$

