Stochastic Network Modeling (SNM)

Llorenç Cerdà-Alabern
Universitat Politècnica de Catalunya
Departament d'Arquitectura de Computadors
llorenc@ac.upc.edu

Parts

Part II

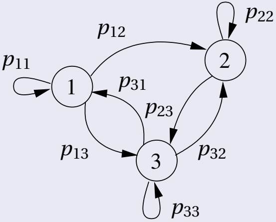
Discrete Time Markov Chains (DTMC)

Outline

Definition of a DTMC

State Transition Diagram

- We are interested in a process that evolve in stages.
- For the model to be tractable, it is convenient to represent the SP by giving all possible states (there may be ∞), and the possible transitions between them:



For the model to be consistent:
$\sum_{\forall j} p_{ij} = 1$

- Mathematically:

$$\boxed{p_{ij} = P\big(X(n) = j \mid X(n-1) = i\big)}$$

Definition of a DTMC

Properties of a DTMC

- The event $X(n) = i$ (at step n the system is in state i) must satisfy (memoryless property):

$$\boxed{\begin{aligned} P(X(n) = j \mid X(n-1) = i, X(n-2) = k, \cdots) = \\ P(X(n) = j \mid X(n-1) = i) \end{aligned}}$$

- If $P\big(X(n) = j \mid X(n-1) = i\big) = P\big(X(1) = j \mid X(0) = i\big)$ for any n we have an homogeneous DTMC. We shall only consider homogeneous DTMC.
- We call one-step transition probabilities to:

$$\boxed{p_{ij} = P\big(X(n) = j \mid X(n-1) = i\big)}$$

- The SP is called a Markov Process (MP) or Markov Chain (MC) depending on the state being continuous or discrete.

Transition Matrix

- Transition probabilities:

$$\boxed{p_{ij} = P\big(X(n) = j \mid X(n-1) = i\big)}$$

- In matrix form:

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

Definition of a DTMC

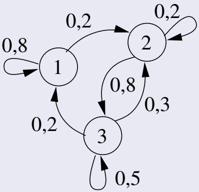
Transition Matrix

- We have
$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}, \text{ where } p_{ij} = P\big(X(n) = j \mid X(n-1) = i\big)$$
- For the model to be consistent, the probability to move from i to any state must be 1. Mathematically:
$$\sum_{\forall j} p_{ij} = \sum_{\forall j} P\big(X(n) = j \mid X(n-1) = i\big) =$$
$$\sum_{\forall j} \frac{P\big(X(n-1) = i \mid X(n) = j\big)\, P\big(X(n) = j\big)}{P(X(n-1) = i)} = \frac{P(X(n-1) = i)}{P(X(n-1) = i)} = \boxed{1}$$
- $\mathbf{P}$ is a stochastic matrix, i.e. a matrix which rows sum 1.

Definition of a DTMC

Example

- Assume a terminal can be in 3 states:
  - State 1: Idle.
  - State 2: Active without sending data.
  - State 3: Active and sending data at a rate v bps.



$$
\begin{array}{cccc}
 & \color{red}{\text{to state}} & & \\
 & \begin{array}{ccc} 1 & 2 & 3 \end{array} & & \\
\mathbf{P} = & \begin{bmatrix} 0{,}8 & 0{,}2 & 0 \\ 0 & 0{,}2 & 0{,}8 \\ 0{,}2 & 0{,}3 & 0{,}5 \end{bmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{array}{c} \color{red}{\text{from}} \\ \color{red}{\text{state}} \end{array}
\end{array}
$$

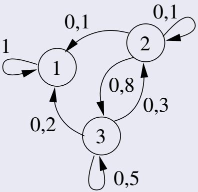
- The average transmission rate (throughput), v_a, is:

$$v_a = P(\text{the terminal is in state 3}) \times v$$

Definition of a DTMC

Absorbing Chains

- It is possible to have chains with absorbing states.
- A state i is absorbing if $p_{ii} = 1$.
- Example: State 1 is absorbing.



$$
\begin{array}{cccccc}
 & \multicolumn{3}{c}{\text{to state}} & & \\
 & 1 & 2 & 3 & & \\
\mathbf{P} = & \multicolumn{3}{c}{\begin{bmatrix} 1 & 0 & 0 \\ 0{,}1 & 0{,}1 & 0{,}8 \\ 0{,}2 & 0{,}3 & 0{,}5 \end{bmatrix}} & \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{array}{c} \text{from} \\ \text{state} \end{array}
\end{array}
$$

Definition of a DTMC

n-step transition probabilities

- Transition probabilities: $p_{ij} = P\big(X(n) = j \mid X(n-1) = i\big)$
- In matrix form:

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

- We define the n-step transition probabilities:

$$p_{ij}(n) = P\big(X(n) = j \mid X(0) = i\big)$$

$$\boxed{\mathbf{P}(n) = \begin{bmatrix} p_{11}(n) & p_{12}(n) & \cdots \\ p_{21}(n) & p_{22}(n) & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}}$$

- $\mathbf{P}$ and $\mathbf{P}(n)$ are stochastic matrices: Their rows sum 1.

Definition of a DTMC

State Probabilities

- Define the probability of being in state i at step n:

$$\pi_i(n) = P(X(n) = i)$$

- In vector form (row vector)

$$\boldsymbol{\pi}(n) = (\pi_1(n), \pi_2(n), \cdots) = (P(X(n) = 1), P(X(n) = 2), \cdots).$$

- Thus, the vector $\boldsymbol{\pi}(n)$ is the distribution of the random variable $X(n)$, and it is called the state probability at step n.

Definition of a DTMC

State Probabilities

- State probability:

$$\boldsymbol{\pi}(n) = (\pi_1(n), \pi_2(n), \cdots) = (P(X(n)=1), P(X(n)=2), \cdots).$$

- Law of total prob. $P(A) = \sum_n P(A \cap B_n) = \sum_n P(A|B_n) P(B_n)$:

$$\pi_i(n) = \sum_k P(X(n-1)=k)\, P\big(X(n)=i \,\big|\, X(n-1)=k\big) = \sum_k \pi_k(n-1)\, p_{ki}$$

$$\pi_i(n) = \sum_k P(X(0)=k)\, P\big(X(n)=i \,\big|\, X(0)=k\big) = \sum_k \pi_k(0)\, p_{ki}(n)$$

- In matrix form:

$$\boxed{\begin{aligned} \boldsymbol{\pi}(n) &= \boldsymbol{\pi}(n-1)\,\mathbf{P} \\ \boldsymbol{\pi}(n) &= \boldsymbol{\pi}(0)\,\mathbf{P}(n) \end{aligned}}$$

where $\boldsymbol{\pi}(0)$ is the initial distribution.

Definition of a DTMC

State Probabilities

$$\boxed{\begin{aligned} \boldsymbol{\pi}(n) &= \boldsymbol{\pi}(n-1)\,\mathbf{P} \\ \boldsymbol{\pi}(n) &= \boldsymbol{\pi}(0)\,\mathbf{P}(n) \end{aligned}}$$

- Iterating

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(n-1)\,\mathbf{P} = \boldsymbol{\pi}(n-2)\,\mathbf{P}\,\mathbf{P} = \boldsymbol{\pi}(n-3)\,\mathbf{P}\,\mathbf{P}\,\mathbf{P} = \cdots = \boldsymbol{\pi}(0)\,\mathbf{P}^n$$

- Thus:

$$\boxed{\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}^n}$$

Definition of a DTMC

Chapman-Kolmogorov Equations

$$\boxed{p_{ij}(n) = \sum_k p_{ik}(r)\, p_{kj}(n-r)}$$

- Proof:

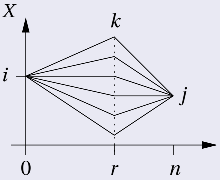
$$\begin{aligned}
p_{ij}(n) &= P\big(X(n)=j \mid X(0)=i\big) = \sum_k P\big(X(n)=j, X(r)=k \mid X(0)=i\big) \\
&= \sum_k \frac{P\big(X(n)=j, X(r)=k, X(0)=i\big)}{P(X(0)=i)} \times \frac{P(X(r)=k, X(0)=i)}{P(X(r)=k, X(0)=i)} \\
&= \sum_k P\big(X(n)=j \mid X(r)=k, X(0)=i\big)\, P\big(X(r)=k \mid X(0)=i\big) \\
&= \sum_k P\big(X(n)=j \mid X(r)=k\big)\, P\big(X(r)=k \mid X(0)=i\big) \\
&= \sum_k p_{ik}(r)\, p_{kj}(n-r)
\end{aligned}$$

$\square$

Chapman-Kolmogorov Equations

$$\boxed{p_{ij}(n) = \sum_k p_{ik}(r)\, p_{kj}(n-r)}$$

- Graphical interpretation:



- In matrix form:

$$\boxed{\mathbf{P}(n) = \mathbf{P}(r)\,\mathbf{P}(n-r)}$$

Definition of a DTMC

Chapman-Kolmogorov Equations

$$\boxed{\mathbf{P}(n) = \mathbf{P}(r)\,\mathbf{P}(n-r)}$$

- Particularly:

$$\mathbf{P}(n) = \mathbf{P}(1)\,\mathbf{P}(n-1) = \mathbf{P}\,\mathbf{P}(n-1) = \mathbf{P}(n-1)\,\mathbf{P}$$

- Iterating:

$$\boxed{\mathbf{P}(n) = \mathbf{P}^n}$$

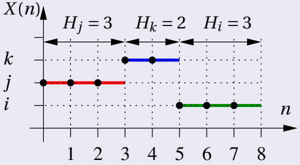
- Thus:

$$\boxed{\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}^n}$$

Definition of a DTMC

Sojourn or Holding Time

- Sojourn or holding time in state k: Is the RV H_k equal to the number of steps that the chain remains in state k before leaving to a different state:



- The Markov property implies:

$$\boxed{H_i(n) = P(H_i = n) = p_{ii}^{n-1}(1 - p_{ii}),\ n \geq 1}$$

- Which is a geometric distribution with mean:

$$\mathrm{E}[H_i] = \sum_{n=1}^{\infty} n P(H_i = n) = \frac{1}{1 - p_{ii}}.$$

Definition of a DTMC

Sojourn or Holding Time

- NOTE: We allow that:

$$p_{ii} = 0 \Rightarrow H_i(n) = I(n=1) = \begin{cases} 1, & n = 1, \\ 0, & \text{otherwise.} \end{cases}, \text{ and}$$

$$p_{ii} = 1 \Rightarrow \mathrm{E}[H_i] = \infty \text{ (absorbing state).}$$

Definition of a DTMC

Theorem

A stochastic process is a DTMC if and only if the sojourn times are geometrically distributed.

Proof.

- We have seen that a DTMC has a sojourn time

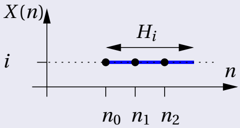
$$\boxed{H_i(n) = P(H_i = n) = p_{ii}^{n-1}(1 - p_{ii}),\ n \geq 1}$$

- Which is geometrically distributed.
- We need to prove that the geometric distribution satisfies the memoryless property (aka Markov property).

□

Definition of a DTMC

The geometric distribution satisfies the Markov property (1)

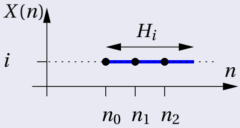


Proof

- Markov property:
  $P\big(X(n_2) = i \mid X(n_1) = i, X(n_0) = i\big) = P\big(X(n_2) = i \mid X(n_1) = i\big)$
- Thus, the Markov property in terms of the sojourn time can be written as:

$$P\big(H_i > n_2 - n_0 \mid H_i > n_1 - n_0\big) = P(H_i > n_2 - n_1)$$

The geometric distribution satisfies the Markov property (2)



$$P\big(H_i > n_2 - n_0 \;\big|\; H_i > n_1 - n_0\big) = P(H_i > n_2 - n_1)$$

- Since

$$P(H_i > k) = 1 - P(H_i \le k) = 1 - \sum_{n=1}^{k} p^{n-1}(1-p) = 1-(1-p)\frac{1-p^k}{1-p} = p^k$$

- We have:

$$P\big(H_i > n_2 - n_0 \;\big|\; H_i > n_1 - n_0\big) = \frac{P(H_i > n_2 - n_0, H_i > n_1 - n_0)}{P(H_i > n_1 - n_0)} =$$

$$\frac{P(H_i > n_2 - n_0)}{P(H_i > n_1 - n_0)} = \frac{p^{n_2 - n_0}}{p^{n_1 - n_0}} = p^{n_2 - n_1} = P(H_i > n_2 - n_1) \qquad \square$$

Part II

Discrete Time Markov Chains (DTMC)

Outline

Transient Solution

Transient Solution

- If we are interested in the transient evolution we shall study $\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}^n$.
- If we can diagonalize **P**, we can obtain the transient evolution in close form.
- **P** can be diagonalized if **P** can be decomposed as:

$$\mathbf{P} = \mathbf{L}^{-1}\,\Lambda\,\mathbf{L}$$

where **L** is some invertible matrix and Λ is the diagonal matrix

$$\Lambda = \mathrm{diag}(\lambda_1, \cdots \lambda_N) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix}$$

with λ_l, $l = 1, \cdots N$ the eigenvalues of **P**.

Eigenvalues

- The eigenvalues λ_l of a matrix $\mathbf{A}$ are scalars that satisfy: $\boldsymbol{l}\mathbf{A} = \lambda_l \boldsymbol{l}$ (or $\mathbf{A}\boldsymbol{r} = \lambda_l \boldsymbol{r}$) for some row vectors $\boldsymbol{l}$ (column vectors $\boldsymbol{r}$), referred to as *left* and *right* eigenvectors, respectively.

$$\boldsymbol{l}\mathbf{A} = \lambda_l \boldsymbol{l} \Rightarrow \boldsymbol{l}(\mathbf{A} - \mathbf{I}\lambda_l) = 0 \Rightarrow \det(\lambda_l \mathbf{I} - \mathbf{A}) = 0$$

$$\mathbf{A}\boldsymbol{r} = \lambda_l \boldsymbol{r} \Rightarrow (\mathbf{A} - \mathbf{I}\lambda_l)\boldsymbol{r} = 0 \Rightarrow \det(\lambda_l \mathbf{I} - \mathbf{A}) = 0$$

- Thus, λ_l solve the characteristic polynomial $\det(\lambda \mathbf{I} - \mathbf{A}) = 0$.
- Note that, in general, *left* and *right* eigenvectors are different, but eigenvalues are the same (they solve the same characteristic polynomial).
- A matrix can be diagonalized if all eigenvalues are single (multiplicity = 1). If a matrix cannot be diagonalized it is called defective.

Transient Solution

Determinants

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}\,a_{22} - a_{12}\,a_{21}$$

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{array}{l} +a_{11}\,a_{22}\,a_{33} + a_{12}\,a_{23}\,a_{31} + a_{21}\,a_{32}\,a_{13} \\ -a_{31}\,a_{22}\,a_{13} - a_{12}\,a_{21}\,a_{33} - a_{23}\,a_{32}\,a_{11} \end{array}$$

- Cofactor Formula: expanding along a row i:

  $$\det \mathbf{A} = \sum_{j=1}^{N} a_{ij}\,(-1)^{i+j} \det M_{ij},$$

  where the minor matrices M_{ij} are obtained removing the row i and column j from $\mathbf{A}$. $(-1)^{i+j} \det M_{ij}$ is called the cofactor of a_{ij}.

Transient Solution

Properties of the determinants

$$\det \mathbf{A} = \prod \text{eigenvalues of } \mathbf{A}$$

$$\text{trace } \mathbf{A} = \sum \text{eigenvalues of } \mathbf{A}$$

where trace $\mathbf{A} = \sum$ elements of the diagonal of $\mathbf{A}$.

Transient Solution

Transient Solution

- Assume a finite DTMC with N states. Then $\mathbf{P} = \mathbf{P}^{N\times N}$.
- Assume that $\mathbf{P}$ can be diagonalized: $\mathbf{P} = \mathbf{L}^{-1}\,\Lambda\,\mathbf{L}$, where Λ is the diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \cdots \lambda_N)$, with λ_l, $l = 1, \cdots N$ the eigenvalues of $\mathbf{P}$.
- Since $\Lambda^n = \mathrm{diag}(\lambda_1^n, \cdots \lambda_N^n)$, we have that

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}^n = \boldsymbol{\pi}(0)\,(\mathbf{L}^{-1}\,\Lambda^n\,\mathbf{L}) = \boldsymbol{\pi}(0)\,(\mathbf{L}^{-1}\,\mathrm{diag}(\lambda_1^n, \cdots \lambda_N^n)\,\mathbf{L})$$

Transient Solution

Transient Solution

- But $\mathbf{L}^{-1}\,\mathrm{diag}(\lambda_1^n,\cdots\lambda_N^n)\,\mathbf{L}$ are linear combinations of $\lambda_1^n,\cdots\lambda_N^n$. Thus, the probability of being in state i is given by:

$$\boxed{\pi_i(n) = (\boldsymbol{\pi}(n))_i = \sum_{l=1}^{N} a_i^{(l)}\,\lambda_l^n}$$

where the unknown coefficients $a_i^{(l)}$ can be obtained solving the system of equations:

$$\sum_{l=1}^{N} a_i^{(l)}\,\lambda_l^n = (\boldsymbol{\pi}(n))_i = (\boldsymbol{\pi}(0)\,\mathbf{P}^n)_i,\ n = 0,\cdots N-1$$

Transient Solution

Example

- Assume a DTMC with

$$\mathbf{P} = \begin{bmatrix} 4/5 & 1/5 \\ 2/5 & 3/5 \end{bmatrix}$$

- We want the probability of being in state 2 in n steps starting from state 1: $\pi_2(n)$ with $\boldsymbol{\pi}(0) = \begin{bmatrix} 1 & 0 \end{bmatrix}$.

Transient Solution

Solution

- It can be easily found that the eigenvalues of $\mathbf{P}$ are $\lambda_1 = 1$ and $\lambda_2 = 2/5$.

$$\pi_2(n) = \lambda_1^n a + b\lambda_2^n = a + b(2/5)^n$$

- Imposing the boundary conditions $\pi_i(n) = (\boldsymbol{\pi}(0)\,\mathbf{P}^n)_i$:

$$\pi_2(0) = a + b = (\begin{bmatrix}1 & 0\end{bmatrix}\mathbf{P}^0)_2 = (\mathbf{P}^0)_{12} = 0$$

$$\pi_2(1) = a + b(2/5) = (\begin{bmatrix}1 & 0\end{bmatrix}\mathbf{P}^1)_2 = (\mathbf{P})_{12} = 1/5$$

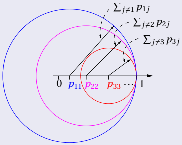
we have that $a = 1/3$, $b = -1/3$, thus:

$$\boxed{\begin{aligned}&\pi_2(n) = 1/3 - 1/3\,(2/5)^n, \quad n \ge 0\\ &\pi_1(n) = 1 - \pi_2(n) = 2/3 + 1/3\,(2/5)^n, \quad n \ge 0\end{aligned}}$$

Transient Solution

Eigenvalues of a Stochastic Matrix

- $\mathbf{P}$ has an eigenvalue equal to 1 ($\mathbf{P}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$, for $\lambda = 1$).
  Proof: $\mathbf{P}\,\mathbf{e} = \mathbf{e}$, where $\mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots \end{bmatrix}^{\mathrm{T}}$ is a column vector of 1 (all rows of $\mathbf{P}$ add to 1). $\square$
- All eigenvalues of $\mathbf{P}$ are $|\lambda_l| \le 1$.
  Proof: Using Gerschgorin's theorem *The eigenvalues of a matrix $\mathbf{P}_{n\times n}$ lie within the union of the n circular disks with center p_{ii} and radius $\sum_{j\neq i}|p_{ij}|$* in $\mathbb{C}$. Since $\sum_j p_{ij} = 1$, the property is proved. $\square$

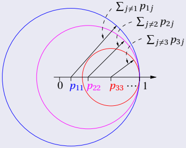


Gershgorin discs of a stochastic matrix

- The eigenvalue $\lambda = 1$ is single if $\mathbf{P}$ is irreducible (Perron-Frobenius theorem). $\mathbf{P}$ is irreducible if all states communicate: for some n, $p_{ij}(n) = (\mathbf{P}^n)_{ij} > 0$, $\forall\, i, j$.

Proof of Gerschgorin's theorem

Gerschgorin's theorem: *The eigenvalues of a matrix $\mathbf{P}_{n\times n}$ lie within the union of the n circular disks with center p_{ii} and radius $\sum_{j\neq i}|p_{ij}|$* in $\mathbb{C}$.



Gershgorin discs of a stochastic matrix

Proof: From $\mathbf{P}\,\boldsymbol{x} = \lambda\,\boldsymbol{x}$ we have

$$\sum_j p_{ij}\,x_j = \lambda\,x_i \quad \forall i \in \{1,\dots,n\}.$$

We choose i such that $|x_i| = \max_j |x_j|$. Thus, $\sum_{j\neq i} p_{ij}\,x_j = \lambda\,x_i - p_{ii}\,x_i$, and

$$|\lambda - p_{ii}| = \left|\sum_{j\neq i} p_{ij}\,\frac{x_j}{x_i}\right| \le \sum_{j\neq i}\left|p_{ij}\,\frac{x_j}{x_i}\right| \le \sum_{j\neq i}|p_{ij}|$$

and the equation $|x-c| \le r$, $x, c \in \mathbb{C}, r \in \mathbb{R}$ is a disk of center c and radius r in $\mathbb{C}$. $\square$

Transient Solution

Chain with a Defective Matrix

- What if **P** cannot be diagonalized? (defective matrix).
- Let λ_l, $l = 1, \cdots L$ be the eigenvalues of $\mathbf{P}^{N\times N}$, each with multiplicity k_l ($k_l \geq 1$, $\sum_l k_l = N$), and a possible eigenvalue $\lambda_1 = 0$ with multiplicity k_1. Then [1]:

$$\boxed{\pi_j(n) = \sum_{m=0}^{k_1-1} a_j^{(1,m)} I(n=m) + \sum_{l=2}^{L} \lambda_l^n \sum_{m=0}^{k_l-1} a_j^{(l,m)} n^m, \quad 1 \leq j \leq N,\ n \geq 0}$$

$I(n=m)$ is the indicator func.: $I(n) = 1$ if $n = m$, $I(n) = 0$ if $n \neq m$.

[1] Llorenç Cerdà-Alabern. *Transient Solution of Markov Chains Using the Uniformized Vandermonde Method.* Tech. rep. UPC-DAC-RR-XCSD-2010-2. Universitat Politècnica de Catalunya, Dec. 2010. URL: https://www.ac.upc.edu/app/research-reports/html/research_center_index-XCSD-2010,en.html.

Transient Solution

Example

- Assume a DTMC with

$$\mathbf{P} = \begin{bmatrix} 3/4 & 1/4 & 0 \\ 0 & 3/4 & 1/4 \\ 1 & 0 & 0 \end{bmatrix}$$

- We want the probability of being in state 1 in n steps starting from state 1: $\pi_1(n)$ with $\pi_1(0) = 1$.
- It can be easily found that the eigenvalues of $\mathbf{P}$ are $\lambda_1 = 1$ and $\lambda_2 = 1/4$ with multiplicity 2. We guess:

$$\pi_1(n) = a + 1/4^n(b + c\,n)$$

- Imposing $\pi_1(0) = 1$, $\pi_1(1) = 3/4$, $\pi_1(2) = (3/4)^2$, we have:

$$\boxed{\pi_1(n) = \frac{4}{9} + \frac{1}{4^n}\left(\frac{5}{9} + \frac{2}{3}n\right)}$$

Part II

Discrete Time Markov Chains (DTMC)

Outline

Classification of States

Objective

- Identify the different types of behavior that the chain can have.
- Introduce the concepts of first passage probability and mean recurrence time.

Classification of States

Irreducibility

- A state j is said to communicate with i, $i \leftrightarrow j$, if $p_{ij}(m_1) > 0$, $p_{ji}(m_2) > 0$ for some $m_1, m_2 \geq 0$.
- We define an irreducible closed set, ICS C_k as a set where all states communicate with each other, and have no transitions to other states out of the set:
  $i \leftrightarrow j, \forall i, j \in C_k$ and $p_{ij} = 0, \forall i \in C_k, j \notin C_k$
  (note that for $i \in C_k, j \notin C_k$ we have: $p_{ij}(2) = \sum_k p_{ik} p_{kj} = 0$, since $p_{ik} = 0$ if $k \notin C_k$, and $p_{kj} = 0$ if $k \in C_k$. Thus, $p_{ij}(n) = 0, \forall n$.)
- An absorbing state form an ICS of only one element. This state, i, must have $p_{ii} = 1$, $p_{ij} = 0\ \forall j \neq i$.
- Transient states do not belong to any ICS.
- A MC is irreducible if all the states form a unique ICS.

Classification of States

Irreducibility

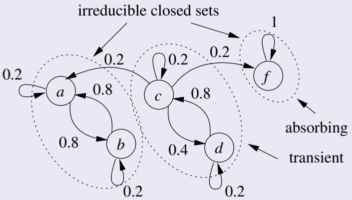
- Assume a MC has M ICSs: By properly numbering the states, we can write $\mathbf{P}$ as an M block diagonal matrix with the probabilities of the transient states in the last rows.
- Example, if $M = 3$:

$$\mathbf{P} = \left[\begin{array}{ccc|c} \mathbf{P}_1 & \mathbf{0} & & \\ & \mathbf{P}_2 & & \\ \mathbf{0} & & \mathbf{P}_3 & \\ \hline \multicolumn{3}{c|}{\text{at least one} > 0} & \mathbf{T} \end{array}\right] \Rightarrow \boldsymbol{\pi}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}^n = \boldsymbol{\pi}(0) \left[\begin{array}{ccc|c} \mathbf{P}_1^n & \mathbf{0} & & \\ & \mathbf{P}_2^n & & \\ \mathbf{0} & & \mathbf{P}_3^n & \\ \hline \multicolumn{3}{c|}{\text{at least one} > 0} & \mathbf{T}^n \end{array}\right]$$

- Note that the M sub-matrices are stochastic (their rows sum 1).

Classification of States

Example



$$\mathbf{P} = \begin{array}{c} \\ a \\ b \\ f \\ c \\ d \end{array} \begin{array}{c} \begin{array}{ccccc} a & b & f & c & d \end{array} \\ \left[\begin{array}{cc|c|cc} 0{,}2 & 0{,}8 & 0 & 0 & 0 \\ 0{,}8 & 0{,}2 & 0 & 0 & 0 \\ \hline 0 & 0 & 1{,}0 & 0 & 0 \\ \hline 0{,}2 & 0 & 0{,}2 & 0{,}2 & 0{,}4 \\ 0 & 0 & 0 & 0{,}8 & 0{,}2 \end{array}\right] \end{array}$$

$$\mathbf{P}^{\infty} = \begin{array}{c} \\ a \\ b \\ f \\ c \\ d \end{array} \begin{array}{c} \begin{array}{ccccc} a & b & f & c & d \end{array} \\ \left[\begin{array}{cc|c|cc} 0{,}5 & 0{,}5 & 0 & 0 & 0 \\ 0{,}5 & 0{,}5 & 0 & 0 & 0 \\ \hline 0 & 0 & 1{,}0 & 0 & 0 \\ \hline 0{,}25 & 0{,}25 & 0{,}5 & 0 & 0 \\ 0{,}25 & 0{,}25 & 0{,}5 & 0 & 0 \end{array}\right] \end{array}$$

- What is the meaning of the probabilities in $\mathbf{P}^{\infty}$? (recall that $(\mathbf{P}^n)_{ij} = p_{ij}(n) = P\big(X(n) = j \mid X(0) = i\big)$).

Classification of States

Example

$$\mathbf{P} = \left[\begin{array}{c|c|c|c} \mathbf{P}_1 & & & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{P}_2 & & \\ & & \mathbf{P}_3 & \\ \hline \multicolumn{3}{c|}{\text{at least one} > 0} & \mathbf{T} \end{array}\right]$$

Theorem *The multiplicity of the eigenvalue $\lambda = 1$ is equal to the number of irreducible closed sets.*

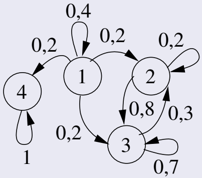
Proof The characteristic polynomial of $\mathbf{P}$ is equal to the product of the characteristic polynomials of the sub-matrices $\mathbf{P}_i$ and $\mathbf{T}$. Since $\mathbf{P}_i$ are irreducible stochastic, each will have a single eigenvalue equal to 1. For the transitorial states it must be $\lim_{n\to\infty} \mathbf{T}^n = \mathbf{0}$. Thus, all the eigenvalues of $\mathbf{T}$ must be $|\lambda| < 1$. $\square$

NOTE: in the closed form solution there is only one unknown associated with $\lambda = 1$, otherwise $\sum_{m=0}^{k_l-1} a_j^{(l,m)} n^m$ will diverge as $n \to \infty$ (i.e. $a_j^{(l,m)} = 0, m > 0$), and $a_j^{(l,0)} = \lim_{n\to\infty} \pi_j(n)$.

Classification of States

Transient and Recurrent

- Recurrent: States that, being visited, they are visited again with probability 1. They are visited an infinite number of times when $n \to \infty$.
- Transient: States that, being visited, have a probability > 0 of never being visited again. They are visited a finite number of times when $n \to \infty$.
- Absorbing: A single (recurrent) state where the chain remains with probability $= 1$.



State 1 is transient

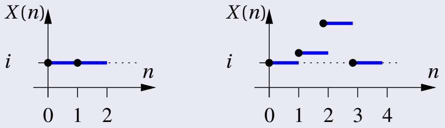
States 2 and 3 are recurrent

State 4 is absorbing

First Passage (Transition) Probabilities

- To derive a classification criteria, we shall study the distribution of the number of steps to go for the first time from a state i another state j. Definition:

$$\boxed{f_{ii}(n) = P\begin{pmatrix}\text{first transition into state } i \\ \text{in } n \text{ steps starting from } i\end{pmatrix}}$$



first transition in 1 step

first transition in 3 steps

- Do not confuse with the n-step transition probability $p_{ii}(n)$, where the state i can be visited in the intermediate states.

Relation between $f_{ii}(n)$ and $p_{ii}(n)$

- $f_{ii}(n)$ and $p_{ii}(n)$ satisfy:

$$\boxed{\begin{aligned} f_{ii}(1) &= p_{ii}(1) \\ p_{ii}(n) &= \sum_{l=1}^{n} f_{ii}(l)\, p_{ii}(n-l),\ n >= 1 \end{aligned}}$$

- The probability that the MC eventually enters state i starting from i is given by:

$$\boxed{f_{ii} = \sum_{n=1}^{\infty} f_{ii}(n)}$$

- If $f_{ii} = 1$ we say i is a recurrent state.
- If $f_{ii} < 1$ we say i is a transient state.

Generalization to Any State Pair

- Analogously to $f_{ii}(n)$, we define the probability of the first passage to state j starting from any state i in n steps: $f_{ij}(n)$.
- $f_{ij}(n)$ and $p_{ij}(n)$ satisfy:

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

Recursive Equation for the First Passage Probabilities

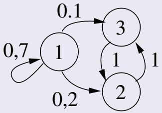
- Recall that the The probability that the MC eventually enters state j starting from i is given by: $f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n)$
- f_{ij} can be computed as follows: Assume we are in i. With probability p_{ij} we will go to j in one step. Otherwise, we will go to k, $k \neq j$, and then we will reach j with probability f_{kj}. Thus:

$$\boxed{f_{ij} = p_{ij} + \sum_{k \neq j} p_{ik} f_{kj}}$$

- If there are more than 1 absorbing states, we can compute the probability to reach them using this method (if there is only 1, say j, then $f_{ij} = 1$, $\forall i$).

Classification of States

Example: Recurrence Times Using the Definition



$$\mathbf{P} = \begin{bmatrix} 0,7 & 0,2 & 0,1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$f_{21}(n) = f_{31}(n) = 0$$
$$f_{11}(n) = 0,7\, I(n=1)$$
$$f_{22}(n) = f_{33}(n) = I(n=2)$$
$$f_{23}(n) = f_{32}(n) = I(n=1)$$

$$f_{12}(n) = \begin{cases} 0,2, & n=1 \\ 0,7^{n-1}\,0,2 + 0,7^{n-2}\,0,1, & n>1 \end{cases}$$

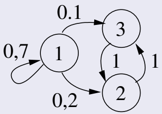
$$f_{13}(n) = \begin{cases} 0,1, & n=1 \\ 0,7^{n-1}\,0,1 + 0,7^{n-2}\,0,2, & n>1 \end{cases}$$

$$f_{11} = 0,7$$
$$f_{12} = f_{13} = 1 \qquad f_{22} = f_{23} = 1$$
$$f_{32} = f_{33} = 1 \qquad f_{21} = f_{31} = 0$$

- State 1 is transient. States 2 and 3 are recurrent.

Classification of States

Example: First Passage Probability Using Recursion



$$\mathbf{P} = \begin{bmatrix} 0{,}7 & 0{,}2 & 0{,}1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

- We have:

$$f_{12} = p_{11} f_{12} + p_{12} + p_{13} f_{32}$$

- Clearly $f_{32} = 1$, thus:

$$f_{12} = 0{,}7 f_{12} + 0{,}2 + 0{,}1 \times 1 \Rightarrow f_{12} = 1$$

as before.

Classification of States

Mean Recurrence Time

- If $f_{ii} = 1$ we say i is a recurrent state.
- If $f_{ii} < 1$ we say i is a transient state.
- When $f_{ii} = 1$, we define the mean recurrence time m_{ii} as:

$$\boxed{m_{ii} = \sum_{n=1}^{\infty} n f_{ii}(n)}$$

- m_{ii} is the average number of steps to eventually reach i starting from i. If $f_{ii} < 1$ (transient state) then we define $m_{ii} = \infty$.
- Classification of recurrent states ($f_{ii} = 1$):
  - If $m_{ii} = \infty$ the state is null recurrent: it takes an ∞ time to reach the state after leave it. Can only happen in chains with an infinite number of states.
  - If $m_{ii} < \infty$ the state is positive recurrent: the state is reached in a finite time after leave it.

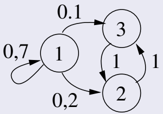
Classification of States

Property of States

In finite MC:

1. States can be only of type positive recurrent or transient.
2. At least one state must be positive recurrent.
3. There are not null recurrent states.

- Example:



- State 1 is transient. States 2 and 3 are positive recurrent.

Generalization to Any State Pair

- When $f_{ij} = 1$, the average number of steps to eventually reach j starting from i, m_{ij} is given by:

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

- If state j can not be reached starting from state i with probability one (if $f_{ij} < 1$), then we define $m_{ij} = \infty$.

Classification of States

Recursive Equation for the Mean Recurrence Time

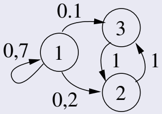
- Recall that the mean recurrence time $m_{ij} = \sum_{n\geq 1} n f_{ij}(n)$ is the average number of steps to eventually reach j starting from i, i.e. it is the mean first passage time from state i to j.
- When $f_{ij} = 1$, m_{ij} can be computed as follows: Assume we are in i. With probability p_{ij} we will go to j in one step. Otherwise, we will go to k, $k \neq j$, and then it will take m_{kj} steps to reach j. Thus:

$$\boxed{m_{ij} = p_{ij} + \sum_{k\neq j} p_{ik}\,(1 + m_{kj}) = 1 + \sum_{k\neq j} p_{ik}\, m_{kj}}$$

since $\sum_j p_{ij} = 1$.

Classification of States

Example: Mean Recurrence Time Using Recursion



$$\mathbf{P} = \begin{bmatrix} 0{,}7 & 0{,}2 & 0{,}1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

- We have:

$$m_{12} = p_{12} + p_{11}\,(1 + m_{12}) + p_{13}\,(1 + m_{32}) = 1 + p_{11}\,m_{12} + p_{13}\,m_{32}$$

- Clearly $m_{32} = 1$, thus:

$$m_{12} = 1 + 0{,}7\,m_{12} + 0{,}1 \times 1 \Rightarrow m_{12} = 11/3.$$

Classification of States

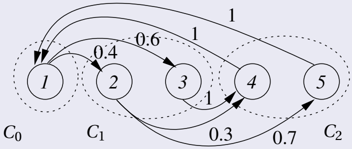
Periodic states

- A recurrent state j is periodic with period $d > 1$ if j can only be reached after leaving it with a multiple of d steps.
- If $d = 1$ the state is aperiodic.
- Any periodic irreducible chain can be partitioned in d cyclic classes $C_0, \cdots C_{d-1}$ such that at each step a transition occur from class C_i to $C_{(i+1) \bmod d}$.
- By properly numerating the states, the transition matrix can be written as (the sub-matrices $\mathbf{A}_i$ may not be square):

$$\mathbf{P} = \begin{array}{c} \\ C_0 \\ C_1 \\ \cdots \\ C_{d-1} \end{array} \begin{array}{c} \begin{array}{ccccc} C_0 & C_1 & C_2 & \cdots & C_{d-1} \end{array} \\ \left[\begin{array}{ccccc} 0 & \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{A}_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{A}_{d-1} & 0 & 0 & \cdots & 0 \end{array} \right] \end{array}$$

Classification of States

Example



$$\mathbf{P} = \begin{bmatrix} 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 0.3 & 0.7 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{P}^2 = \begin{bmatrix} 0 & 0 & 0 & 0.72 & 0.28 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \end{bmatrix}, \mathbf{P}^3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 0.72 & 0.28 \\ 0 & 0 & 0 & 0.72 & 0.28 \end{bmatrix}, \mathbf{P}^4 = \begin{bmatrix} 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 0.72 & 0.28 \\ 0 & 0 & 0 & 0.72 & 0.28 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \cdots$$

- In periodic chains $\mathbf{P}^n$ does not converge.

Part II

Discrete Time Markov Chains (DTMC)

Outline

Steady State

Limiting Distribution

- Probability of being in state i at step n:
$$\pi_i(n) = P(X(n) = i)\,.$$
In vector form (row vector)
$$\boldsymbol{\pi}(n) = (\pi_1(n), \pi_2(n), \cdots).$$
- The evolution of the chain depends on the initial distribution $\boldsymbol{\pi}(0)$.
- If we are interested in the transient evolution we shall study
$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}^n.$$
- If we are interested the steady state we shall be interested in the limiting distribution (if the limit exists):
$$\boldsymbol{\pi}(\infty) = (\pi_1(\infty), \pi_2(\infty), \cdots)$$

Steady State

Limiting Distribution

Assume an irreducible chain with positive recurrent states.

- With infinite steps, we look for a probability converging to a value that depends only on the final state:
$$\pi_j(\infty) = \sum_i \pi_i(0) \lim_{n\to\infty} p_{ij}(n), \forall j \text{ and for any } \boldsymbol{\pi}(0),$$
which implies:
$$\pi_j(\infty) = \lim_{n\to\infty} p_{ij}(n) \sum_i \pi_i(0) = p_{ij}(\infty), \forall j \Rightarrow$$
$$\mathbf{P}(\infty) = \lim_{n\to\infty} \mathbf{P}^n = \begin{bmatrix} \boldsymbol{\pi}(\infty) \\ \cdots \\ \boldsymbol{\pi}(\infty) \end{bmatrix}$$
- If this limit exists, we call $\mathbf{P}(\infty)$ the limiting matrix, and $\boldsymbol{\pi}(\infty)$ the limiting distribution.

Steady State

Example

$$\mathbf{P} = \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.7 & 0.2 & 0.1 \\ 0.5 & 0.3 & 0.2 \end{bmatrix} \qquad \mathbf{P}^2 = \begin{bmatrix} 0.764 & 0.168 & 0.068 \\ 0.760 & 0.170 & 0.070 \\ 0.752 & 0.174 & 0.074 \end{bmatrix}$$

$$\mathbf{P}^4 = \begin{bmatrix} 0{,}7628 & 0{,}1686 & 0{,}0686 \\ 0{,}7620 & 0{,}1690 & 0{,}0690 \\ 0{,}7604 & 0{,}1698 & 0{,}0698 \end{bmatrix} \qquad \mathbf{P}^8 = \begin{bmatrix} 0.762500 & 0.168750 & 0.068750 \\ 0.762499 & 0.168750 & 0.068750 \\ 0.762497 & 0.168752 & 0.068752 \end{bmatrix}$$

$\ldots$

$$\Rightarrow \boldsymbol{\pi}(\infty) = (0.76250, 0.16875, 0.06875)$$

Stationary distribution

- We have:

$$\pi_i(n) = P(X(n)=i) = \sum_k P(X(n-1)=k)\, P\big(X(n)=i \mid X(n-1)=k\big)$$
$$= \sum_k \pi_k(n-1)\, p_{ki}$$

- In matrix form: $\boldsymbol{\pi}(n) = \boldsymbol{\pi}(n-1)\,\mathbf{P}$
- If $\pi_i(n) = \pi_i(n-1) = \pi_i$, we call π_i the stationary probability of state i, and $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots)$, the stationary distribution of the chain.
- In matrix form (Global balance equations):

$$\boxed{\begin{array}{l} \boldsymbol{\pi} = \boldsymbol{\pi}\,\mathbf{P} \\ \boldsymbol{\pi}\,\mathbf{e} = 1, \mathbf{e} = \begin{bmatrix} 1 & 1 & \cdots \end{bmatrix}^{\mathrm{T}} \end{array}}$$

- Thus, the stationary distribution is the left-hand eigenvector corresponding to the unit eigenvalue of $\mathbf{P}$.
- $\boldsymbol{\pi}(n) = \boldsymbol{\pi} \Rightarrow \boldsymbol{\pi}(n+1) = \boldsymbol{\pi}(n)\,\mathbf{P} = \boldsymbol{\pi}\,\mathbf{P} = \boldsymbol{\pi} \Rightarrow \boxed{\boldsymbol{\pi}(k) = \boldsymbol{\pi},\ k \geq n}$

Stationary distribution

- Do not confuse the limiting distribution $\boldsymbol{\pi}(\infty)$ and the stationary distribution $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$.
- $\boldsymbol{\pi}(\infty)$ and $\boldsymbol{\pi}$ may not be the same, e.g. in periodic chains $\boldsymbol{\pi}(\infty)$ does not exists ($\mathbf{P}$ does not converge), but we can compute the stationary distribution.
- Example: the periodic chain

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

has the stationary distribution

$$\boldsymbol{\pi} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Numerical Solution

- Replace one equation method:

$$\boxed{\begin{aligned}&\boldsymbol{\pi} = \boldsymbol{\pi}\,\mathbf{P}\\ &\boldsymbol{\pi}\,\mathbf{e} = 1, \mathbf{e} = \begin{bmatrix}1 & 1 & \cdots\end{bmatrix}^{\mathrm{T}}\end{aligned}}$$

- We solve the equation $\boldsymbol{\pi}(\mathbf{I}-\mathbf{P}) = 0$ replacing the last equation by $\boldsymbol{\pi}\,\mathbf{e} = 1$:

$$\boxed{\boldsymbol{\pi}\begin{bmatrix} p_{11}-1 & p_{12} & \cdots p_{1n-1} & 1\\ p_{21} & p_{22}-1 & \cdots p_{2n-1} & 1\\ \vdots & \vdots & \vdots & \vdots\\ p_{n1} & p_{n2} & \cdots p_{nn-1} & 1\end{bmatrix} = \begin{bmatrix}0 & 0 & \cdots & 0 & 1\end{bmatrix}}$$

Steady State

Numerical Solution

- Replace one equation method: $\mathbf{P} = \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.7 & 0.2 & 0.1 \\ 0.5 & 0.3 & 0.2 \end{bmatrix}$
- With octave (matlab clone):

```
octave:1> P=[0.8,0.15,0.05;0.7,0.2,0.1;0.5,0.3,0.2];
octave:2> s=size(P,1); # number of rows.
octave:3> [zeros(1,s-1),1] / ...
> [eye(s,s-1)-P(1:s,1:s-1), ones(s,1)]
ans =
0.762500   0.168750   0.068750
```

- With R

```
> P <- matrix(nc=3, byr=T, c(0.8,0.15,0.05,0.7,0.2,0.1,0.5,0.3,0.2))
> s <- nrow(P)
> solve(t(cbind(P[,1:(s-1)]-diag(nr=s,nc=s-1), rep(1,s))),
+ c(rep(0,s-1),1))
[1] 0.76250 0.16875 0.06875
```

NOTE: $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P} \Rightarrow \boldsymbol{\pi}^{\mathrm{T}} = \mathbf{P}^{\mathrm{T}}\boldsymbol{\pi}^{\mathrm{T}}$. The transpose operator in R is `t()`.