Models matricials: cadenes de Markov

Problemes per a la ciència de dades

PID_00262436

Francesc Pozo Montero
Jordi Ripoll Missé

Universitat Oberta de Catalunya

Francesc Pozo Montero

Llicenciat en Matemàtiques per la Universitat de Barcelona (2000) i doctor en Matemàtica Aplicada per la Universitat Politècnica de Catalunya (2005). Ha estat professor associat a la Universitat Autònoma de Barcelona i professor associat, col·laborador i actualment professor agregat a la Universitat Politècnica de Catalunya. A més, és cofundador del Grup d'Innovació Matemàtica E-learning (GIMEL), responsable de diversos projectes d'innovació docent i autor de diverses publicacions. Com a membre del grup de recerca consolidat CoDAlab, centra la recerca en la teoria de control i les aplicacions en enginyeria mecànica i civil, com també en l'ús de la ciència de dades per al monitoratge de la integritat estructural i per al monitoratge de la condició, sobretot en turbines eòliques.

Jordi Ripoll Missé

Llicenciat en Matemàtiques i doctor en Ciències Matemàtiques per la Universitat de Barcelona (2005). Professor col·laborador de la Universitat Oberta de Catalunya des del 2011 i professor del Departament d'Informàtica, Matemàtica Aplicada i Estadística de la Universitat de Girona (UdG) des del 1996, on actualment és professor agregat i desenvolupa tasques de recerca en l'àmbit de la biologia matemàtica (models amb equacions en derivades parcials i dinàmica evolutiva). També ha estat professor i tutor de la UNED en dues etapes, primer al centre associat de Terrassa i actualment al de Girona. Ha participat en nombrosos projectes d'innovació docent, especialment pel que fa a l'aprenentatge de les matemàtiques en línia.

L'encàrrec i la creació d'aquest recurs d'aprenentatge UOC han estat coordinats per la professora: Cristina Cano Bastidas (2019)

Primera edició: febrer 2019

Av. Tibidabo, 39-43, 08035 Barcelona
Disseny: Manel Andreu
Realització editorial: Oberta UOC Publishing, SL

Índex

1. Problemes 5

2. Solucions dels problemes 9

3. Exercicis d'autoavaluació 16

4. Solucions dels exercicis d'autoavaluació 19

1. Problemes

Models matricials: cadenes de Markov

Temes: *Introducció als models matricials en temps discret. Concepte de cadenes de Markov en temps discret. Diagrama d'estats i probabilitats de transició. Evolució en el temps d'una cadena de Markov. Matrius positives i valor propi dominant. Distribucions d'estat estacionàries. Aplicacions.*

Una cadena de Markov és un tipus de procés estocàstic en què el sistema no té memòria, és a dir, l'evolució futura només depèn del present i no de l'estat del sistema en el passat. Concretament, en les cadenes de Markov la distribució de les probabilitats condicionades dels estats futurs (probabilitats de transició) depenen de l'estat present del procés i no de la seqüència d'esdeveniments que han succeït en el passat.

En aquest mòdul només tractarem de les anomenades *cadenes de Markov homogènies*, en què les probabilitats de transició són independents del temps (el que es coneix en altres contextos com a *sistemes dinàmics autònoms*). Si treballem a temps discret $t \geq 0$, tindrem un model matricial lineal, en què la matriu P del sistema es pot descriure per columnes o per files. Habitualment es construeix la matriu (positiva) P mitjançant un diagrama d'estats (*nodes*) i de les probabilitats de transició entre estats (*fletxes entre nodes*).

L'evolució en el temps de la cadena de Markov s'obté per iteració de la fórmula $\vec{x}_{t+1} = P\,\vec{x}_t$, $t \geq 0$, és a dir, multiplicant la matriu del sistema pel vector (columna) dels estats:

$$\begin{pmatrix} x_{t+1}^{(1)} \\ \vdots \\ x_{t+1}^{(n)} \end{pmatrix} = \begin{pmatrix} p_{11} & \dots & p_{1n} \\ \vdots & & \vdots \\ p_{n1} & \dots & p_{nn} \end{pmatrix} \begin{pmatrix} x_t^{(1)} \\ \vdots \\ x_t^{(n)} \end{pmatrix}, \quad t = 0,1,2,\dots$$

En certes hipòtesis sobre la matriu P, la distribució límit del sistema coincideix amb l'estat estacionari que és donat pel vector propi de valor propi 1 de la matriu P. Però aquest vector propi cal que estigui normalitzat, de manera que la suma de les seves components sigui igual a 1: $\vec{x} = P\,\vec{x}$ amb $x^{(1)} + \cdots + x^{(n)} = 1$. Si el vector propi no està normalitzat, només cal que el dividim per la suma de les seves components: $\frac{1}{v_1 + \ldots + v_n}(v_1, \ldots, v_n)$.

A partir d'aquesta teoria podem resoldre els següents problemes i exercicis d'autoavaluació:

1. S'ha fet un estudi sobre el moviment dels treballadors d'una empresa multinacional, que té la seu principal a Londres (LON) i tres filials repartides a les ciutats de París (PAR), Milà (MIL) i Barcelona (BCN). S'ha estimat que, cada any, el 70% dels treballadors localitzats en cada una de les ciutats no canvien de lloc de treball d'un any al següent. Segons les dades recopilades, el 30% restant dels treballadors es distribueixen entre cada una de les localitzacions de la manera següent:

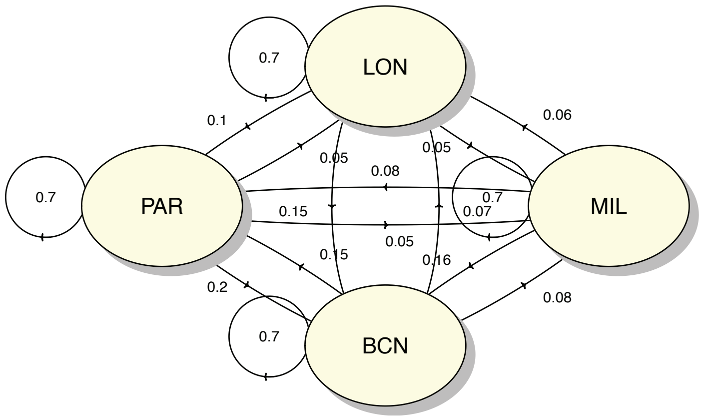
- Anualment, el 10% dels treballadors de LON se'n van a PAR; el 5%, a MIL, i la resta, a BCN.

- Cada any, el 20% dels treballadors de PAR se'n van a BCN i la resta es divideixen a parts iguals entre LON i MIL.

- Cada any, el 6% dels treballadors de MIL se'n van a LON; el 8%, a PAR, i la resta, a BCN.

- Finalment, cada any, el 7% dels treballadors de BCN se'n van a LON; el 8%, a MIL, i la resta, a PAR.

Diagrama d'estats i probabilitats de transició: moviment d'empleats



a) Escriviu la matriu d'aquesta cadena de Markov a partir del diagrama d'estats (4 nodes: LON, PAR, MIL i BCN) i de les probabilitats de transició. Recordeu que la suma de les columnes d'aquesta matriu ha de ser igual a 1.

b) Disposem de les dades per a l'any 2019. La plantilla total de l'empresa és de N = 236 treballadors, repartits així: 46 a LON, 58 a PAR, 70 a MIL i 62 a BCN. Calculeu quants treballadors hi haurà l'any següent a cada ciutat.

c) Considerant la distribució geogràfica de treballadors de l'apartat anterior per a l'any 2019, quants treballadors hi havia el 2018 a Barcelona?

d) Segons aquest model, quants treballadors hi haurà l'any 2025 treballant a cada ciutat?

e) Quin percentatge de treballadors s'espera que hi hagi en cadascuna de les seus a llarg termini?

2. Les enquestes de satisfacció dels estudiants d'una universitat sobre la docència rebuda classifiquen els estudiants en tres categories:

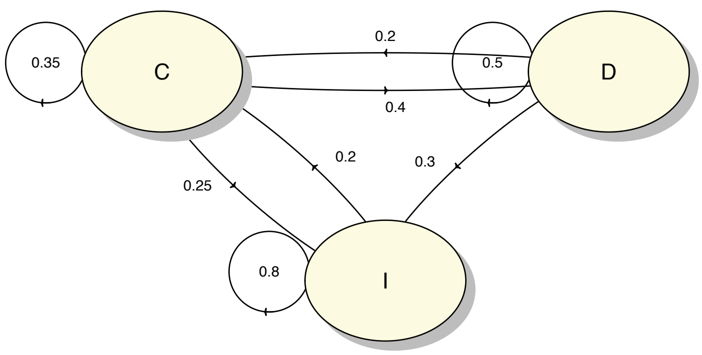
$$\text{contents } \overset{\circ\circ}{\smile}, \text{ descontents } \overset{\circ\circ}{\frown} \text{ i indiferents } \overset{\circ\circ}{-}.$$

A partir de les dades recopilades durant els darrers anys, s'han pogut estimar les probabilitats de canviar d'opinió que tenen els estudiants sobre la docència que han rebut. Les dades, en forma de matriu, són les següents:

$$\begin{array}{cc} & \begin{array}{ccc} C & D & I \end{array} \\ P = & \begin{pmatrix} 35\% & 20\% & 20\% \\ 40\% & 50\% & 0 \\ 25\% & 30\% & 80\% \end{pmatrix} \begin{array}{c} C \\ D \\ I \end{array} \end{array}$$

on cada element de la matriu, p_{ij}, $1 \leq i,j \leq 3$, és la probabilitat de transició de l'estat j a l'estat i en una unitat de temps. El diagrama dels tres estats amb les probabilitats de transició corresponents és:

Diagrama d'estats i probabilitats de transició: satisfacció estudiants



a) Comproveu que la matriu transposada P^T té el vector propi $\vec{v} = (1,1,1)^T$ de valor propi $\lambda_1 = 1$. Recordeu que els valors propis d'una matriu i els de la seva transposada són els mateixos, però els vectors propis poden ser diferents.

b) Calculeu els altres dos valors propis de P sabent que el determinant i la traça de la matriu compleixen que $\det(P) = 1 \cdot \lambda_2 \cdot \lambda_3$ i $\mathrm{tr}(P) = 1 + \lambda_2 + \lambda_3$.

c) Quina serà la proporció entre *contents : descontents : indiferents* a llarg termini? Expresseu els percentatges (%) de cada categoria.

2. Solucions dels problemes

1. a) La matriu de les probabilitats de canviar de seu, és a dir, dels moviments dels treballadors de l'empresa multinacional és:

$$P = \begin{array}{c} \begin{array}{cccc} LON & PAR & MIL & BCN \end{array} \\ \begin{pmatrix} 0.70 & 0.05 & 0.06 & 0.07 \\ 0.10 & 0.70 & 0.08 & 0.15 \\ 0.05 & 0.05 & 0.70 & 0.08 \\ 0.15 & 0.20 & 0.16 & 0.70 \end{pmatrix} \end{array} \begin{array}{c} LON \\ PAR \\ MIL \\ BCN \end{array}$$

b) Si multipliquem la matriu P per les dades del problema, és a dir, pel vector del nombre de treballadors en cada seu l'any 2019, obtindrem la predicció, segons aquest model, del nombre de treballadors que hi haurà a cada seu l'any 2020:

Nota

El resultat final es pot arrodonir a l'enter més proper. Com que són prediccions mitjanes, s'obtenen decimals.

$$\begin{array}{c} \begin{array}{cccc} LON & PAR & MIL & BCN \end{array} \\ \begin{pmatrix} 0.70 & 0.05 & 0.06 & 0.07 \\ 0.10 & 0.70 & 0.08 & 0.15 \\ 0.05 & 0.05 & 0.70 & 0.08 \\ 0.15 & 0.20 & 0.16 & 0.70 \end{pmatrix} \end{array} \begin{array}{c} 2019 \\ \begin{pmatrix} 46 \\ 58 \\ 70 \\ 62 \end{pmatrix} \end{array} = \begin{array}{c} 2020 \\ \begin{pmatrix} 43.64 \\ 60.10 \\ 59.16 \\ 73.10 \end{pmatrix} \end{array} \begin{array}{c} LON \\ PAR \\ MIL \\ BCN \end{array}$$

c) Segons el model, per saber el nombre de treballadors que hi havia a cada seu l'any 2018 hem de resoldre el sistema lineal següent:

$$\begin{array}{c} \begin{array}{cccc} LON & PAR & MIL & BCN \end{array} \\ \begin{pmatrix} 0.70 & 0.05 & 0.06 & 0.07 \\ 0.10 & 0.70 & 0.08 & 0.15 \\ 0.05 & 0.05 & 0.70 & 0.08 \\ 0.15 & 0.20 & 0.16 & 0.70 \end{pmatrix} \end{array} \begin{array}{c} 2018 \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \end{array} = \begin{array}{c} 2019 \\ \begin{pmatrix} 46 \\ 58 \\ 70 \\ 62 \end{pmatrix} \end{array} \begin{array}{c} LON \\ PAR \\ MIL \\ BCN \end{array}$$

i obtenim $x_1 = 49.98716$, $x_2 = 56.78659$, $x_3 = 87.61808$ i $x_4 = 41.60816$. Per tant, a Barcelona hi havia $41.60816 \simeq 42$ treballadors de l'empresa.

d) Per saber els treballadors que hi haurà l'any 2025 a partir de les dades del 2019, hem d'iterar la cadena de Markov sis vegades:

$$P \cdot x_{2019} = x_{2020}\ , \quad P \cdot x_{2020} = x_{2021}\ , \quad P \cdot x_{2021} = x_{2022}$$

$$P \cdot x_{2022} = x_{2023}\ , \quad P \cdot x_{2023} = x_{2024}\ , \quad P \cdot x_{2024} = x_{2025}\ ,$$

és a dir, per passar de l'any 2019 a l'any 2025 hem de multiplicar per la potència sisena P^6 de la matriu:

$$\begin{array}{c} \begin{array}{cccc} LON & PAR & MIL & BCN \end{array} \\ \begin{pmatrix} 0.70 & 0.05 & 0.06 & 0.07 \\ 0.10 & 0.70 & 0.08 & 0.15 \\ 0.05 & 0.05 & 0.70 & 0.08 \\ 0.15 & 0.20 & 0.16 & 0.70 \end{pmatrix}^6 \end{array} \begin{array}{c} 2019 \\ \begin{pmatrix} 46 \\ 58 \\ 70 \\ 62 \end{pmatrix} \end{array} = \begin{array}{c} 2025 \\ \begin{pmatrix} 40.26034 \\ 66.55989 \\ 42.97274 \\ 86.20703 \end{pmatrix} \end{array} \begin{array}{c} \\ LON \\ PAR \\ MIL \\ BCN \end{array}$$

e) Per calcular la distribució límit de treballadors a cada seu, hem de calcular el vector propi $\vec{x}$ de valor propi $\lambda = 1$ de la matriu P:

$$\begin{array}{c} \begin{array}{cccc} LON & PAR & MIL & BCN \end{array} \\ \begin{pmatrix} 0.70 & 0.05 & 0.06 & 0.07 \\ 0.10 & 0.70 & 0.08 & 0.15 \\ 0.05 & 0.05 & 0.70 & 0.08 \\ 0.15 & 0.20 & 0.16 & 0.70 \end{pmatrix} \end{array} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \begin{array}{c} LON \\ PAR \\ MIL \\ BCN \end{array}$$

Aquest sistema té infinites solucions. Si n'escollim una que no sigui nul·la i n'expressem les components en percentatges, obtindrem:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 16.9\% \\ 28.7\% \\ 17.5\% \\ 36.9\% \end{pmatrix} \begin{array}{c} LON \\ PAR \\ MIL \\ BCN \end{array}$$

Recordeu que qualsevol múltiple no nul d'un vector propi també és vector propi i, per aquest motiu, podem normalitzar-lo (dividir el vector per la suma de les seves components) i expressar-lo en percentatges.
El codi R necessari per resoldre tot l'exercici és el següent:

```
> # 1. Moviment dels treballadors d'una empresa multinacional
> Etiquetes <- c(" LON", "PAR", "MIL", "BCN")
> MatriuP <- matrix(c(0.7,0.05,0.06,0.07,0.1,0.7,0.08,0.15,
+                      0.05,0.05,0.7,0.08,0.15,0.2,0.16,0.7), 4,4,
+                      byrow=T, dimnames = list(Etiquetes, Etiquetes))
> # install.packages("markovchain")
> library(markovchain)
> # Creem un nou objecte que és una cadena de Markov (per columnes)
> CM_treballadors <- new("markovchain", states = Etiquetes, byrow = F,
+                     transitionMatrix = MatriuP,
+                     name = "Moviment dels treballadors entre les seus de l'empresa")
> CM_treballadors

Moviment dels treballadors entre les seus de l'empresa
 A  4 - dimensional discrete Markov Chain defined by the following states:
 LON, PAR, MIL, BCN
 The transition matrix  (by cols)  is defined as follows:
     LON  PAR  MIL  BCN
LON 0.70 0.05 0.06 0.07
PAR 0.10 0.70 0.08 0.15
MIL 0.05 0.05 0.70 0.08
BCN 0.15 0.20 0.16 0.70

> # install.packages("diagram")
> library(diagram)
> # (a) Dibuix del diagrama d'estats i probabilitats de transició
> plotmat(MatriuP, pos = c(1,2,1), lwd = 1, box.lwd = 1,cex.txt = 0.7,
+   box.size = 0.09, box.type = "circle", box.prop = 0.75,
+   box.col = " light yellow", arr.length=.05, arr.width=.1,
+   self.cex = .5, self.shifty = .01, self.shiftx = -.13,
+   main = "Diagrama d'estats i probabilitats de transició: moviment de treballadors")
> # (b) Dades del 2019:
> library(expm);
> x2019 <- matrix(c(46,58,70,62), 4, 1, dimnames = list(Etiquetes))
> x2019

    [,1]
LON   46
PAR   58
MIL   70
BCN   62

> # Distribució de treballadors l'any 2019
> x2020 <- MatriuP %*% x2019 # multipliquem per la matriu per anar un any endavant
> x2020

     [,1]
LON 43.64
PAR 60.10
```

```
MIL 59.16
BCN 73.10

> # Distribució de treballadors l'any 2020
> # (c) Dades del 2018
> x2018 <- solve(MatriuP,x2019)  # solucionem el sistema per anar un any enrere
> x2018

        [,1]
LON 49.98716
PAR 56.78659
MIL 87.61808
BCN 41.60816

> # Distribució de treballadors l'any 2018
> # (d) Dades del 2025
> CM_treballadors^6  # cadena de Markov al cap de sis anys (2025-2019= 6)

Moviment dels treballadors entre les seus de l'empresa^6
 A  4 - dimensional discrete Markov Chain defined by the following states:
 LON, PAR, MIL, BCN
 The transition matrix  (by cols)  is defined as follows:
          LON       PAR       MIL       BCN
LON 0.2258980 0.1525205 0.1578021 0.1609144
PAR 0.2678989 0.3188126 0.2571385 0.2862213
MIL 0.1549842 0.1579423 0.2309254 0.1696457
BCN 0.3512189 0.3707246 0.3541340 0.3832185

> x2025 <- (MatriuP %^% 6) %*% x2019
> x2025

        [,1]
LON 40.26034
PAR 66.55989
MIL 42.97274
BCN 86.20703

> # (e) Estat estacionari = distribució límit d'estats
> V <- eigen(MatriuP)
> V # valors i vectors propis de la matriu del sistema

eigen() decomposition
$values
[1] 1.0000000+0.0000000i 0.6422344+0.0071586i 0.6422344-0.0071586i
[4] 0.5155311+0.0000000i

$vectors
             [,1]                   [,2]                   [,3]          [,4]
[1,] -0.3204833+0i  0.0198523+0.4321114i  0.0198523-0.4321114i  0.1085121+0i
[2,] -0.5452498+0i  0.4791925-0.2714381i  0.4791925+0.2714381i  0.5294608+0i
```

```
[3,] -0.3310336+0i -0.6733400+0.0000000i -0.6733400+0.0000000i  0.1832046+0i
[4,] -0.7002927+0i  0.1742952-0.1606734i  0.1742952+0.1606734i -0.8211775+0i

> vaps <- V$values
> abs(vaps) # valors propis en valor absolut

[1] 1.0000000 0.6422743 0.6422743 0.5155311

> veps <- V$vectors
> steadystate <- abs(veps[,1])
> (steadystate <- 100*steadystate/sum(steadystate))

[1] 16.89369 28.74184 17.44983 36.91464
```

2. a) Comprovem que $P^T \cdot \vec{v} = \vec{v}$ amb $\vec{v} = (1,1,1)^T$:

$$\begin{array}{c} \begin{array}{ccc} C & D & I \end{array} \\ \begin{pmatrix} 0.35 & 0.40 & 0.25 \\ 0.20 & 0.50 & 0.30 \\ 0.20 & 0 & 0.80 \end{pmatrix} \end{array} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \begin{array}{c} C \\ D \\ I \end{array}$$

Per tant, la matriu transposada P^T té el valor propi $\lambda_1 = 1$ i, en conseqüència, la matriu del sistema P també té el valor propi $\lambda_1 = 1$.

b) Calculem primer la traça i el determinant de la matriu: $\mathrm{tr}(P) = 1.65$ i $\det(P) = 0.075$. Per trobar els altres dos valors propis λ_2, λ_3 podem resoldre les equacions de la següent manera. De la segona equació tenim que $\lambda_2 = 0.65 - \lambda_3$ i llavors la primera equació esdevé $0.075 = (0.65 - \lambda_3)\lambda_3$, que és una equació de segon grau amb dues solucions diferents: $\lambda = 0.5$ o $\lambda = 0.15$.

c) Per determinar la distribució límit d'estudiants en cada categoria, hem de calcular el vector propi de valor propi $\lambda = 1$, $P \cdot \vec{x} = \vec{x}$:

$$\begin{array}{c} \begin{array}{ccc} C & D & I \end{array} \\ \begin{pmatrix} 0.35 & 0.20 & 0.20 \\ 0.40 & 0.50 & 0 \\ 0.25 & 0.30 & 0.80 \end{pmatrix} \end{array} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \begin{array}{c} C \\ D \\ I \end{array}$$

De les infinites solucions d'aquest sistema, en triem una que no sigui el vector zero i l'expressem en percentatges:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 23.5\% \\ 18.8\% \\ 57.7\% \end{pmatrix} \begin{matrix} C \\ D \\ I \end{matrix}$$

El codi R necessari per resoldre tot l'exercici és el següent:

```
> # 2. Enquestes de satisfacció dels estudiants
> Etiquetes <- c(" C", "D", " I")
> MatriuP <- matrix(c(0.35,0.20,0.20,0.40,0.50,0,0.25,0.30,0.80),
+                     3,3, byrow=T, dimnames = list(Etiquetes, Etiquetes))
> # Creem un nou objecte que és una cadena de Markov (per columnes)
> CM_estudiants <- new("markovchain", states = Etiquetes, byrow = F,
+                        transitionMatrix = MatriuP,
+                        name = "Enquestes de satisfacció dels estudiants")
> CM_estudiants

Enquestes de satisfacció dels estudiants
 A  3 - dimensional discrete Markov Chain defined by the following states:
 C, D, I
 The transition matrix  (by cols)  is defined as follows:
     C   D   I
C 0.35 0.2 0.2
D 0.40 0.5 0.0
I 0.25 0.3 0.8

> # Dibuix del diagrama d'estats i probabilitats de transició
> plotmat(MatriuP, pos = c(2,1), lwd = 1, box.lwd = 1,cex.txt = 0.7,
+         box.size = 0.09, box.type = "circle", box.prop = 0.75,
+         box.col = " light yellow", arr.length=.05, arr.width=.1,
+         self.cex = .5, self.shifty = .01, self.shiftx = -.13,
+ main = "Diagrama d'estats i probabilitats de transició: satisfacció estudiants")
> # (a) La matriu P i la seva transposada tenen valor propi = 1
> vep <- matrix(c(1,1,1),3,1)
> t(MatriuP) %*% vep # la matriu transposada té vector propi (1,1,1)

  [,1]
C    1
D    1
I    1

> # (b) La resta de valors propis:
> V <- eigen(MatriuP)
> V # valors i vectors propis de la matriu del sistema

eigen() decomposition
$values
```

```
[1] 1.00 0.50 0.15

$vectors
          [,1]          [,2]        [,3]
[1,] 0.3617281 -2.386516e-16  0.65561007
[2,] 0.2893825 -7.071068e-01 -0.74926865
[3,] 0.8862339  7.071068e-01  0.09365858

> VT <- eigen(t(MatriuP))
> VT # valors i vectors propis de la matriu transposada

eigen() decomposition
$values
[1] 1.00 0.50 0.15

$vectors
          [,1]       [,2]       [,3]
[1,] 0.5773503 -0.6948499 -0.9169493
[2,] 0.5773503 -0.5500895  0.2821382
[3,] 0.5773503  0.4632333  0.2821382

> # (c) Proporció límit en percentatges
> vaps <- V$values
> abs(vaps) # valors propis en valor absolut

[1] 1.00 0.50 0.15

> veps <- V$vectors
> steadystate <- abs(veps[,1])
> (steadystate <- 100*steadystate/sum(steadystate))

[1] 23.52941 18.82353 57.64706
```

3. Exercicis d'autoavaluació

1. La dinàmica d'una cadena de Markov per a una població dividida en dos grups (individus sans i individus malalts) és donada per la matriu

$$P = \begin{pmatrix} 0.25 & r \\ 0.75 & 1-r \end{pmatrix},$$

on r és la probabilitat de canviar del segon grup al primer en cada període de temps. Per quin valor de r la proporció entre els individus sans i els individus malalts serà, a la llarga, de 2 : 4?

◯ $r = 0.625$

◯ $r = 0.375$

◯ $r = 0.125$

◯ Cap resposta anterior és correcta.

2. Una població, on els individus estan classificats en contents ☺ o enfadats ☹, evoluciona segons una cadena de Markov. De mitjana, cada any el 30% dels contents s'indignen i, en canvi, el 20% dels enfadats recuperen el bon humor. A llarg termini, quina d'aquestes afirmacions és certa?

◯ La proporció de contents i enfadats serà de 30% : 70%.

◯ Hi haurà el mateix nombre de contents que d'enfadats.

◯ La proporció de contents i enfadats serà de 40% : 60%.

◯ La proporció entre contents i enfadats dependrà de les dades inicials del problema.

3. Mensualment es recullen dades sobre l'opinió que tenen els habitants d'un país sobre dues candidates a la presidència del govern: la senyora Marianna i la senyora Petra. Considerem que alguns són partidaris de Marianna i d'altres de Petra i que el percentatge dels que continuen creient en la seva candidata d'un mes al següent és del 90% i del 75%, respectivament. A la llarga, quin serà el percentatge d'individus favorables a cada candidata?

◯ Marianna : Petra = 88.2% : 11.8%

◯ Marianna : Petra = 71.4% : 28.6%

◯ Marianna : Petra = 45.4% : 54.6%

◯ Marianna : Petra = 21.7% : 78.3%

4. Els habitants d'una regió metropolitana es classifiquen en dos grups: residents a la ciutat (centre) i residents als suburbis (perifèria). Cada any, el 5% de la població de la ciutat se'n va cap als suburbis i el 3% de la població dels suburbis es muden a la ciutat. Escriviu la matriu de Markov del model i calculeu la distribució de població que hi haurà a la llarga.

◯ Ciutat : Suburbi = 62.5% : 37.5%

◯ Ciutat : Suburbi = 37.5% : 62.5%

◯ Ciutat : Suburbi = 5% : 95%

◯ Ciutat : Suburbi = 97% : 3%

5. Considereu una cadena de Markov per a la distribució geogràfica d'una espècie d'aus migratòries distribuïda en tres hàbitats diferents. S'han recopilat dades sobre les probabilitats (mitjanes) de transició: el 23% de les aus del primer hàbitat s'hi queden i el 55% emigren al segon; el 29% de les aus del segon hàbitat s'hi queden i el 65% emigren al primer. Finalment, el 48% i el 50% de les aus del tercer hàbitat emigren cap al primer i el segon, respectivament. Segons aquest model, quin dels tres hàbitats serà el més poblat a llarg termini?

◯ El primer.

◯ El segon.

◯ El tercer.

◯ No es pot saber quin hàbitat serà el més poblat.

6. Una cadena de Markov per als consumidors d'un producte comercialitzat per dues empreses (Nestlé i Unilever) és donada per aquesta matriu: $P = \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}$, on $0 \leq p,q \leq 1$ són les probabilitats de continuar comprant setmanalment el mateix producte de cada una de les empreses, respectivament. A partir de la fórmula per les potències de la matriu $P^t = V \cdot D^t \cdot V^{-1}$, $t \geq 0$, en què la matriu V són els vectors propis per columnes i la matriu diagonal D són els valors propis de P, calculeu l'expressió de la matriu de transició al cap de 52 setmanes:

$$P^{52} = \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}^{52}$$

en funció de les probabilitats de transició p i q tals que $p+q \neq 2$.

○ $P^{52} = \frac{1}{2-p-q}\begin{pmatrix} 1-q & 1 \\ 1-p & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (p+q-1)^{52} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1-p & q-1 \end{pmatrix}$

○ $P^{52} = \begin{pmatrix} p^{52} & 0 \\ 0 & q^{52} \end{pmatrix}$

○ $P^{52} = \frac{1}{p+q-2}\begin{pmatrix} 1-p & 1 \\ 1-q & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (p+q-1)^{52} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1-p & 1-q \end{pmatrix}$

○ $P^{52} = \begin{pmatrix} (1-p)^{52} & 0 \\ 0 & (1-q)^{52} \end{pmatrix}$

7. Considereu la matriu P de l'exercici anterior per a la cadena de Markov dels consumidors d'un producte comercialitzat per dues empreses (Nestlé i Unilever). Trobeu per a quin cas es compleix el límit de la cadena de Markov:

$$\lim_{t\to\infty}\begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}^t = \frac{1}{2-(p+q)}\begin{pmatrix} 1-q & 1-q \\ 1-p & 1-p \end{pmatrix}$$

en funció de les probabilitats de transició p i q.

○ Si $p+q \neq 2$ i $p+q \neq 1$.

○ Si $p+q \neq 1$ i $p+q \neq 0$.

○ Si $p+q \neq 2$ i $p+q \neq 0$.

○ Cap resposta anterior és correcta.

Nota de l'exercici 7

Fixeu-vos que la matriu límit té les dues columnes iguals i que els elements de cada columna sumen 1.

4. Solucions dels exercicis d'autoavaluació

1. La probabilitat r de recuperar-se de la malaltia es calcula resolent la següent equació:

$$\begin{array}{c} \textbf{sans} \\ \textbf{malalts} \end{array} \begin{pmatrix} 0.25 & r \\ 0.75 & 1-r \end{pmatrix} \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

i el valor de r és:

○ $r = 0.625$

⊗ $r = 0.375$

○ $r = 0.125$

○ Cap resposta anterior és correcta.

2. Per trobar el percentatge de contents i enfadats que hi haurà segons el model, hem de resoldre aquest sistema:

$$\begin{array}{c} \textbf{contents} \\ \textbf{enfadats} \end{array} \begin{pmatrix} 70\% & 20\% \\ 30\% & 80\% \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Una de les solucions del sistema és $(x_1,x_2) = (20,30) \equiv (20,30)/50$ i, per tant, l'afirmació correcta és:

○ La proporció de contents i enfadats serà de 30% : 70%.

○ Hi haurà el mateix nombre de contents que d'enfadats.

⊗ La proporció de contents i enfadats serà de 40% : 60%.

○ La proporció entre contents i enfadats dependrà de les dades inicials del problema.

3. Per saber el percentatge d'individus favorables a cada candidata que hi haurà segons el model, hem de resoldre aquest sistema:

$$\begin{array}{c} \textbf{Marianna} \\ \textbf{Petra} \end{array} \begin{pmatrix} 90\% & 25\% \\ 10\% & 75\% \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Una de les solucions del sistema és $(x_1,x_2) = (25,10) \equiv (25,10)/35$ i, per tant, la proporció serà:

○ Marianna : Petra = 88.2% : 11.8%

⊗ Marianna : Petra = 71.4% : 28.6%

○ Marianna : Petra = 45.4% : 54.6%

○ Marianna : Petra = 21.7% : 78.3%

4. A partir de les dades del model, la distribució geogràfica de població que hi haurà entre el centre de la ciutat i els suburbis s'obté resolent:

$$\begin{array}{c} \textbf{ciutat} \\ \textbf{suburbi} \end{array} \begin{pmatrix} 95\% & 3\% \\ 5\% & 97\% \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Una de les solucions del sistema és $(x_1,x_2) = (3,5) \equiv (3,5)/8$ i, per tant, la proporció serà:

○ Ciutat : Suburbi = 62.5% : 37.5%

⊗ Ciutat : Suburbi = 37.5% : 62.5%

○ Ciutat : Suburbi = 5% : 95%

○ Ciutat : Suburbi = 97% : 3%

5. Amb les dades recopilades podem saber la distribució geogràfica d'aquesta espècie en cada hàbitat. La distribució límit s'obté resolent el sistema:

$$\begin{array}{c} \textbf{1r hàbitat} \\ \textbf{2n hàbitat} \\ \textbf{3r hàbitat} \end{array} \begin{pmatrix} 23\% & 65\% & 48\% \\ 55\% & 29\% & 50\% \\ 22\% & 6\% & 2\% \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

o agrupant les variables, en què obtenim un sistema lineal homogeni:

$$\begin{array}{c} \textbf{1r hàbitat} \\ \textbf{2n hàbitat} \\ \textbf{3r hàbitat} \end{array} \begin{pmatrix} -77\% & 65\% & 48\% \\ 55\% & -71\% & 50\% \\ 22\% & 6\% & -98\% \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Podeu resoldre el sistema lineal pel mètode de Gauss, per exemple. Si normalitzem una de les solucions obtingudes, obtindrem que la proporció en cada hàbitat serà: 44.28% (1r), 43.14% (2n) i 12.58% (3r). La resposta correcta és:

⊗ El primer.

○ El segon.

○ El tercer.

○ No es pot saber quin hàbitat serà el més poblat.

6. Per calcular la potència $t = 52$ (setmanes) d'una cadena de Markov amb dos estats:

$$\begin{array}{r} \textbf{Nestlé} \\ \textbf{Unilever} \end{array} \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}$$

hem de calcular els valors propis i els vectors propis de la matriu. Els valors propis són $\lambda_1 = 1$ i $\lambda_2 = p + q - 1$ i els vectors propis són $\vec{v}_1 = (1-q, 1-p)$ i $\vec{v}_2 = (1, -1)$. Llavors, usant la fórmula per les potències:

$$\begin{array}{r} \textbf{Nestlé} \\ \textbf{Unilever} \end{array} \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}^t = \begin{pmatrix} 1-q & 1 \\ 1-p & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & p+q-1 \end{pmatrix}^t \begin{pmatrix} 1-q & 1 \\ 1-p & -1 \end{pmatrix}^{-1}$$

i calculant la inversa, obtindrem la solució demanada:

⊗ $P^{52} = \frac{1}{2-p-q} \begin{pmatrix} 1-q & 1 \\ 1-p & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (p+q-1)^{52} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1-p & q-1 \end{pmatrix}$

○ $P^{52} = \begin{pmatrix} p^{52} & 0 \\ 0 & q^{52} \end{pmatrix}$

○ $P^{52} = \frac{1}{p+q-2} \begin{pmatrix} 1-p & 1 \\ 1-q & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (p+q-1)^{52} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1-p & 1-q \end{pmatrix}$

○ $P^{52} = \begin{pmatrix} (1-p)^{52} & 0 \\ 0 & (1-q)^{52} \end{pmatrix}$

7. Per obtenir el límit proposat quan $t \to \infty$ de la cadena de Markov amb dos estats:

$$\begin{array}{r} \textbf{Nestlé} \\ \textbf{Unilever} \end{array} \begin{pmatrix} p & 1-q \\ 1-p & q \end{pmatrix}$$

hem d'assegurar que la matriu del sistema té valor propi dominant, és a dir, $|\lambda_1| > |\lambda_2|$. Com que els valors propis són $\lambda_1 = 1$ i $\lambda_2 = p + q - 1$, la condició es compleix per a:

$$|p+q-1| < 1\,, \quad -1 < p+q-1 < 1\,, \quad 0 < p+q < 2$$

Dit d'una altra manera, les dues úniques cadenes de Markov que no compleixen el límit són:

$$\begin{array}{c} \textbf{Nestlé} \\ \textbf{Unilever} \end{array} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{i} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

corresponents a dos casos extrems. En el primer cas ($p = q = 1$), els consumidors sempre compren el producte de la mateixa marca; en el segon cas ($p = q = 0$), cada setmana tots els consumidors canvien de marca del producte que van comprar la setmana passada. Per tant, la resposta correcta és:

○ Si $p + q \neq 2$ i $p + q \neq 1$.

○ Si $p + q \neq 1$ i $p + q \neq 0$.

⊗ Si $p + q \neq 2$ i $p + q \neq 0$.

○ Cap resposta anterior és correcta.