Soluciones selectas

Hidden Markov Models for Time Series Zucchini et al.

Jesús Muñoz Cano

1 de Octubre de 2025

Capítulo 1

Problema 1

(a) Let X be a random variable which is distributed as a (δ_1, δ_2)-mixture of two distributions with expectations μ_1, μ_2, and variances σ_1^2, σ_2^2, respectively, where $\delta_1 + \delta_2 = 1$.

i. Show that $\mathrm{Var}(X) = \delta_1\sigma_1^2 + \delta_2\sigma_2^2 + \delta_1\delta_2(\mu_1 - \mu_2)^2$.

ii. Show that a (non-trivial) mixture of two Poisson distributions with distinct means is overdispersed, that is, $\mathrm{Var}(X) > \mathbb{E}(X)$.

(b) Now suppose that X is a mixture of $m \geq 2$ distributions, with means μ_i and variances σ_i^2, for $i = 1, 2, \ldots, m$. The mixing distribution is $\boldsymbol{\delta}$.

i. Show that

$$\mathrm{Var}(X) = \sum_{i=1}^{m} \delta_i\sigma_i^2 + \sum_{i<j} \delta_i\delta_j(\mu_i - \mu_j)^2.$$

Hint: use either $\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ or the conditional variance formula,

$$\mathrm{Var}(X) = \mathbb{E}(\mathrm{Var}(X \mid C)) + \mathrm{Var}(\mathbb{E}(X \mid C)).$$

ii. Describe the circumstances in which $\mathrm{Var}(X)$ equals the linear combination $\sum_{i=1}^{m} \delta_i\sigma_i^2$.

Solución

(a) i.

$$\mathrm{Var}(X) \;=\; \mathbb{E}\big(\mathrm{Var}(X \mid C)\big) \;+\; \mathrm{Var}\big(\mathbb{E}(X \mid C)\big)$$

$$= \;\Big[\mathbb{P}(C{=}1)\cdot\mathrm{Var}(X \mid C{=}1) + \mathbb{P}(C{=}2)\cdot\mathrm{Var}(X \mid C{=}2)\Big] \;+\; \mathrm{Var}\big(\mu_C\big)$$

$$= \;\delta_1\sigma_1^2 + \delta_2\sigma_2^2 \;+\; \mathbb{E}\big(\mu_C^2\big) - \big(\mathbb{E}(\mu)\big)^2$$

$$= \;\delta_1\sigma_1^2 + \delta_2\sigma_2^2 \;+\; \delta_1\mu_1^2 + \delta_2\mu_2^2 \;-\; (\delta_1\mu_1 + \delta_2\mu_2)^2$$

$$= \;\delta_1\sigma_1^2 + \delta_2\sigma_2^2 \;+\; \delta_1\mu_1^2 + \delta_2\mu_2^2 \;-\; \big(\delta_1^2\mu_1^2 + 2\delta_1\delta_2\mu_1\mu_2 + \delta_2^2\mu_2^2\big)$$

$$= \;\delta_1\sigma_1^2 + \delta_2\sigma_2^2 \;+\; (\delta_1 - \delta_1^2)\mu_1^2 + (\delta_2 - \delta_2^2)\mu_2^2 \;-\; 2\delta_1\delta_2\mu_1\mu_2$$

$$= \;\delta_1\sigma_1^2 + \delta_2\sigma_2^2 \;+\; \delta_1\delta_2\,\mu_1^2 + \delta_1\delta_2\,\mu_2^2 \;-\; 2\delta_1\delta_2\mu_1\mu_2$$

$$= \;\delta_1\sigma_1^2 + \delta_2\sigma_2^2 \;+\; \delta_1\delta_2\,(\mu_1 - \mu_2)^2.$$

Nota 1:

$$\delta_1+\delta_2=1, \qquad \delta_1-\delta_1^2=\delta_1(1-\delta_1)=\delta_1\delta_2, \qquad \delta_2-\delta_2^2=\delta_2(1-\delta_2)=\delta_1\delta_2.$$

Nota 2: $\mathrm{Var}(\mu_C)=\mathbb{E}(\mu_C^2)-(\mathbb{E}\mu_C)^2$, *con* $\mu_C\in\{\mu_1,\mu_2\}$.

(a) ii. *Usando el resultado anterior y considerando que para Poisson* $\mathbb{E}[N]=\mathrm{Var}(N)=\mu$ *(es decir,* $\sigma_i^2=\mu_i$*):*

$$\mathrm{Var}(X) \;=\; \delta_1\mu_1+\delta_2\mu_2 \;+\; \delta_1\delta_2(\mu_1-\mu_2)^2.$$

Así,

$$\mathrm{Var}(X) \;=\; \mathbb{E}(X) \;+\; \delta_1\delta_2(\mu_1-\mu_2)^2.$$

Para una mezcla no-trivial *(*$\mu_1\neq\mu_2$ *y* $0<\delta_1,\delta_2<1$*) se tiene* $\delta_1\delta_2(\mu_1-\mu_2)^2>0$*, luego*

$$\mathrm{Var}(X) \;>\; \mathbb{E}[X].$$

(b) i.

$$\mathrm{Var}(X) \;=\; \mathbb{E}\big[\mathrm{Var}(X\mid C)\big] \;+\; \mathrm{Var}\big(\mathbb{E}(X\mid C)\big)$$

$$\textit{Eq. 1} \qquad \mathbb{E}\big[\mathrm{Var}(X\mid C)\big] \;=\; \sum_{i=1}^{m}\delta_i\,\sigma_i^2.$$

$$\textit{Eq. 2} \qquad \mathrm{Var}\big(\mathbb{E}(X\mid C)\big) \;=\; \mathrm{Var}(\mu_C) \;=\; \mathbb{E}[M_C^2] \;-\; \big(\mathbb{E}[M_C]\big)^2$$

$$=\; \sum_{i=1}^{m}\delta_i\,\mu_i^2 \;-\; \Big(\sum_{i=1}^{m}\delta_i\,\mu_i\Big)^2$$

$$=\; \sum_{i=1}^{m}(\delta_i-\delta_i^2)\,\mu_i^2 \;-\; 2\sum_{i<j}\delta_i\delta_j\,\mu_i\,\mu_j$$

$$\textit{Nota 3:} \qquad \delta_i\,(1-\delta_i) \;=\; \delta_i\cdot\sum_{\substack{j\\ j\neq i}}\delta_j \;=\; \sum_{\substack{j\\ j\neq i}}\delta_i\,\delta_j$$

Sustituyendo lo obtenido en Nota 3, notemos que la primera parte de Eq. 2 se convierte en,

$$\sum_{i=1}^{m}(\delta_i-\delta_i^2)\,\mu_i^2 = \sum_{i=1}^{m}\Big(\sum_{\substack{j\\ j\neq i}}\delta_i\delta_j\Big)\,\mu_i^2 \;=\; \sum_{i=1}^{m}\sum_{\substack{j\\ j\neq i}}\delta_i\delta_j\,\mu_i^2$$

$$=\; \sum_{i<j}\left(\delta_i\delta_j\,\mu_i^2+\delta_i\delta_j\,\mu_j^2\right) \;=\; \sum_{i<j}\delta_i\delta_j\,(\mu_i^2+\mu_j^2)$$

Sustituyendo en Eq. 2,

$$\sum_{i<j}\delta_i\delta_j\,\mu_i^2 \;-\; 2\sum_{i<j}\delta_i\delta_j\,\mu_i\mu_j \;+\; \sum_{i<j}\delta_i\delta_j\,\mu_j^2$$

Finalmente, al sumar Eq. 1 y Eq. 2,

$$\Rightarrow \quad \mathrm{Var}(X) \;=\; \sum_{i<j} \delta_i \delta_j \,(\mu_i - \mu_j)^2 \;+\; \sum_{i=1}^{m} \delta_i \,\sigma_i^2.$$

Nota 4: De parejas ordenadas a no ordenadas.
Para cualquier pareja no ordenada $\{i,j\}$ existen 2 contribuciones ordenadas a la suma (de (i,j): $\delta_i\delta_j\,\mu_i^2$, y de (j,i): $\delta_j\delta_i\,\mu_j^2$).
Usamos que

$$\sum_i \sum_{\substack{j \\ j\neq i}} (\cdot) \;=\; \sum_{i<j} \big[(\cdot)_{(i,j)} + (\cdot)_{(j,i)}\big].$$

(b) ii. Cuando cada componente tiene la misma media*:*
Si todas las medias coinciden ($\mu_i = \mu_j$ para todo i, j), entonces los términos $\delta_i\delta_j(\mu_i - \mu_j)^2$ se anulan y queda

$$\mathrm{Var}(X) \;=\; \sum_{i=1}^{m} \delta_i \,\sigma_i^2.$$

Problema 3

Brown and Buckley (2015, p. 308) consider a Poisson mixture likelihood of the form

$$L \;=\; \prod_{i=1}^{n} \sum_{j=1}^{k} w_j \, f(x_i \mid \mu_j).$$

(Here $f(\cdot \mid \mu)$ denotes a Poisson probability function with mean μ.) They write that 'Even for moderate values of n and k, this takes a long time to evaluate as there are k^n terms when the inner sums are expanded', and do not pursue maximum likelihood estimation.

Explain why it is in fact possible to evaluate L or its logarithm in computations which are of order kn rather than k^n.

Solución

$$L \;=\; \prod_{i=1}^{n} \sum_{j=1}^{k} \delta_j \,\rho_j(x_i, \theta_j) \;=\; \prod_{i=1}^{n} \sum_{j=1}^{k} \delta_j \,\frac{\lambda_j^{x_i}\, e^{-\lambda_j}}{x_i!}$$

$$\log L \;=\; \sum_{i=1}^{n} \log\left[\sum_{j=1}^{k} \delta_j \,\rho_j(x_i, \theta_j)\right] \;=\; \sum_{i=1}^{n} \log\left(S_i\right)$$

Para cada S_i hay k términos dentro de la suma $\sum_{j=1}^{k}(\cdot) \Rightarrow$ complejidad $O(k)$

Hay n términos i en la suma externa $\sum_{i=1}^{n}(\cdot) \Rightarrow$ complejidad total de la suma doble: $O(k\,n)$.

Capítulo 2

Problema 8

8. Consider an m-state HMM with the basic dependence structure as depicted in Figure 2.2.

(a) Consider the vector $\boldsymbol{\alpha}_t = \big(\alpha_t(1), \ldots, \alpha_t(m)\big)$ defined by

$$\alpha_t(j) \;=\; \Pr\big(\mathbf{X}^{(t)} = \mathbf{x}^{(t)},\, C_t = j\big), \qquad j = 1, 2, \ldots, m.$$

Use conditional probability and the conditional independence assumptions to show that

$$\alpha_t(j) \;=\; \sum_{i=1}^{m} \alpha_{t-1}(i)\, \gamma_{ij}\, p_j(x_t).$$

(b) Verify for yourself that the result from (a), written in matrix notation, yields the forward recursion

$$\boldsymbol{\alpha}_t \;=\; \boldsymbol{\alpha}_{t-1}\, \Gamma\, \mathbf{P}(x_t), \qquad t = 2, \ldots, T.$$

(c) Hence derive the matrix expression for the likelihood.

Solución

(a) *Para* $j \in \{1, \ldots, m\}$,

$$\begin{aligned}
\alpha_t(j) &= \sum_{i=1}^{m} \Pr\big(\mathbf{X}^{(t)} = \mathbf{x}^{(t)},\, C_{t-1} = i,\, C_t = j\big) \\
&= \sum_{i=1}^{m} \Pr\Big(\mathbf{X}^{(t-1)} = \mathbf{x}^{(t-1)},\, X_t = x_t,\, C_{t-1} = i,\, C_t = j\Big) \\
&= \sum_{i=1}^{m} \Pr\big(X_t = x_t \mid C_t = j\big)\; \Pr\big(C_t = j \mid C_{t-1} = i\big)\; \Pr\big(\mathbf{X}^{(t-1)} = \mathbf{x}^{(t-1)},\, C_{t-1} = i\big) \\
&= \sum_{i=1}^{m} \rho_j(x_t)\; \gamma_{ij}\; \alpha_{t-1}(i).
\end{aligned}$$

(b) *Escribiendo el resultado de (a) en notación matricial:*

$$\alpha_t(j) \;=\; \sum_{i=1}^{m} \alpha_{t-1}(i)\, \gamma_{ij}\, \rho_j(x_t).$$

Sea $\boldsymbol{\alpha}_t = (\alpha_t(1), \ldots, \alpha_t(m))$, $\Gamma = [\gamma_{ij}]$ *y*

$$\mathbf{P}(x_t) = \operatorname{diag}(\rho_1(x_t), \ldots, \rho_m(x_t)).$$

Entonces

$$\left(\boldsymbol{\alpha}_{t-1}\,\Gamma\,\mathbf{P}(x_t)\right)_j = \sum_{i=1}^{m} \alpha_{t-1}(i)\,\gamma_{ij}\,\rho_j(x_t) = \alpha_t(j),$$

y por lo tanto

$$\boldsymbol{\alpha}_t \;=\; \boldsymbol{\alpha}_{t-1}\,\Gamma\,\mathbf{P}(x_t), \qquad t \geq 2.$$

Para el inicio,

$$\boldsymbol{\alpha}_1 \;=\; \boldsymbol{\delta}\,\mathbf{P}(x_1), \qquad \delta_j = \Pr(C_1 = j).$$

(c) *La verosimilitud puede escribirse como*

$$L \;=\; \Pr\big(\mathbf{X}^{(t)} = \mathbf{x}^{(t)}\big) \;=\; \sum_{j=1}^{m} \alpha_t(j) \;=\; \boldsymbol{\alpha}_t\,\mathbf{1},$$

donde $\mathbf{1}$ *es el vector columna de unos. Encadenando la recursión,*

$$L \;=\; \boldsymbol{\delta}\,\mathbf{P}(x_1)\,\Gamma\,\mathbf{P}(x_2)\,\Gamma\,\mathbf{P}(x_3)\,\cdots\,\Gamma\,\mathbf{P}(x_t)\,\mathbf{1}.$$

Problema 12

12. (Interval-censored observations.)

(a) Suppose that, in a series of unbounded counts $x_1, \ldots, x_T$, only the observation x_t is interval-censored, and $a \leq x_t \leq b$, where b may be ∞. Prove the statement made in Section 2.3.5 that the likelihood of a PoissonHMM with m states is obtained by replacing $\mathbf{P}(x_t)$ in the expression (2.12) by the $m \times m$ diagonal matrix of which the ith diagonal element is $\Pr(a \leq X_t \leq b \mid C_t = i)$.

(b) Extend part (a) to allow for any number of interval-censored observations.

Solución

Referencia: *Cuando tenemos todas las observaciones (sin censura),*

$$L_T \;=\; \Pr\big(\mathbf{X}^{(T)} = \mathbf{x}^{(T)}\big) \;=\; \sum_{c^{(T)}} \Pr\big(\mathbf{X}^{(T)} = \mathbf{x}^{(T)},\, C^{(T)} = c^{(T)}\big) \;=\; \boldsymbol{\delta}\,\mathbf{P}(x_1)\,\Gamma\,\mathbf{P}(x_2)\cdots\Gamma\,\mathbf{P}(x_T)\,\mathbf{1}.$$

(a) Observación x_t ***censurada por intervalo.*** *Para observaciones con censura por intervalos, el* L_T *ajustado queda:*

$$L_T \;=\; \Pr\left(A,\, B\right),$$

donde

$$A:\; X_1 = x_1,\, X_2 = x_2, \ldots, X_{t-1} = x_{t-1},\, X_{t+1} = x_{t+1}, \ldots, X_T = x_T, \qquad B:\; a \leq X_t \leq b.$$

$$L_T \;=\; \sum_{C_1,\ldots,C_T} \Pr\big(A,\, B,\, C^{(T)} = c^{(T)}\big).$$

Recordando que en cualquier modelo gráfico dirigido la distribución conjunta se factoriza como

$$\Pr(V_1,\ldots,V_n) \;=\; \prod_{i=1}^{n} \Pr\big(V_i \,|\, \mathrm{pa}(V_i)\big),$$

donde $\mathrm{pa}(V_i)$ *es el conjunto de padres de* V_i*.*

De aquí,

$$\Pr\big(A, B, C^{(T)} = c^{(T)}\big) = \Pr(C_1)\ \Pr(X_1 \mid C_1) \prod_{\substack{k=2\\k\neq t}}^{T} \Pr(X_k \mid C_k)\ \Pr(C_k \mid C_{k-1})$$
$$\cdot \Pr(B \mid C_t)\ \Pr(C_t \mid C_{t-1}).$$

Por lo tanto,

$$L_T \;=\; \sum_{C_1,\ldots,C_T} \delta_{c_1}\, \rho_{c_1}(x_1)\, \gamma_{c_1 c_2}\, \rho_{c_2}(x_2)\, \cdots\, \rho_{c_{t-1}}(x_{t-1})\ \Pr(B \mid C_t)\ \gamma_{c_{t-1}c_t}\, \cdots\, \rho_{c_T}(x_T).$$

Se obtiene la misma forma matricial de antes excepto por el reemplazo en el tiempo t:

$$\Pr(B \mid C_t = i) \;=\; \Pr\big(a \le X_t \le b \,\big|\, C_t = i\big) \;=\; \sum_{x=a}^{b} \Pr(X_t = x \mid C_t = i).$$

Denotando

$$\mathbf{P}_{[a,b]}(x_t) \;=\; \mathrm{diag}\Big(\Pr(a \le X_t \le b \mid C_t = 1),\, \ldots,\, \Pr(a \le X_t \le b \mid C_t = m)\Big),$$

tenemos

$$L_T \;=\; \boldsymbol{\delta}\, \mathbf{P}(x_1)\, \Gamma\, \mathbf{P}(x_2)\, \cdots\, \Gamma\, \mathbf{P}_{[a,b]}(x_t)\, \cdots\, \Gamma\, \mathbf{P}(x_T)\, \mathbf{1}.$$

Caso particular: PoissonHMM.

$$\Pr(B \mid C_t = i) = \Pr\big(a \le X_t \le b \,\big|\, C_t = i\big) = \sum_{x=a}^{b} \frac{e^{-\lambda_i}\lambda_i^x}{x!}.$$

Si $b \longrightarrow \infty$ *entonces*

$$\Pr\big(a \le X_t \le \infty \mid C_t = i\big) = 1 - F_{\mathrm{Pois}(\lambda_i)}(a-1).$$

(b) Extensión a cualquier número de observaciones censuradas. *Se hace el mismo reemplazo en cada tiempo censurado* t_k*:*

$$\mathbf{P}(x_{t_k})$$

se convierte en

$$\mathbf{P}_{[a_{t_k},\, b_{t_k}]}(x_{t_k}) = \mathrm{diag}\Big(\Pr(a_{t_k} \leq X_{t_k} \leq b_{t_k} \mid C_{t_k} = 1), \ldots, \Pr(a_{t_k} \leq X_{t_k} \leq b_{t_k} \mid C_{t_k} = m) \Big).$$

Finalmente,

$$L_T \;=\; \boldsymbol{\delta}\, \mathbf{P}(x_1)\, \Gamma\, \mathbf{P}(x_2) \,\cdots\, \Gamma\, \mathbf{P}_{[a_{t_1},\, b_{t_1}]}(x_{t_1}) \,\cdots\, \Gamma\, \mathbf{P}_{[a_{t_K},\, b_{t_K}]}(x_{t_K}) \,\cdots\, \Gamma\, \mathbf{P}(x_T)\, \mathbf{1}.$$

Capítulo 3

Problema 11

11. (Embeddability of discrete-time Markov chain in continuous-time.) It is not always possible to embed a discrete-time Markov chain uniquely in a continuous-time chain. That is, given a t.p.m. Γ, there does not always exist a unique generator matrix $\mathbf{Q}$ such that $\Gamma = \exp(\mathbf{Q})$. The following examples show that there may, even in simple cases, be more than one corresponding generator matrix, or there may be none.

(a) (Example taken from Israel, Rosenthal and Wei (2001, p. 256).) Consider the matrices Γ, $\mathbf{Q}_1$ and $\mathbf{Q}_2$ given by

$$\Gamma = \frac{1}{5}\begin{pmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix} - \frac{e^{-4\pi}}{5}\begin{pmatrix} -3 & 2 & 1 \\ 2 & -3 & 1 \\ 2 & 2 & -4 \end{pmatrix}, \qquad \mathbf{Q}_1 = 2\pi\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & 0 & -2 \end{pmatrix}, \qquad \mathbf{Q}_2 = \frac{4\pi}{5}\begin{pmatrix} -3 & 2 & 1 \\ 2 & -3 & 1 \\ 2 & 2 & -4 \end{pmatrix}.$$

Verify that $\exp(\mathbf{Q}_1) = \exp(\mathbf{Q}_2) = \Gamma$.

(b) Theorem 3.1 of Israel *et al.* (2001, p. 249) states the following.

Let $\mathbf{P}$ *be a transition [probability] matrix, and suppose that*

i. $\det(\mathbf{P}) \leq 0$, *or*

ii. $\det(\mathbf{P}) > \prod_i p_{ii}$, *or*

iii. *there are states* i *and* j *such that* j *is accessible from* i*, but* $p_{ij} = 0$.

Then there does not exist an exact generator for $\mathbf{P}$.

Use this theorem to conclude that there is no corresponding generator matrix for the following t.p.m.s:

$$\Gamma = \begin{pmatrix} 0.4 & 0.6 \\ 0.5 & 0.5 \end{pmatrix}, \qquad \Gamma = \begin{pmatrix} 0.9 & 0.1 & 0 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}.$$

Solución

(a) *Verificar que* $\exp(\mathbf{Q}_1) = \exp(\mathbf{Q}_2) = \Gamma$.

Parte A.1: usando $\mathbf{Q}_2$ y un proyector idempotente.

Sea

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} -3 & 2 & 1 \\ 2 & -3 & 1 \\ 2 & 2 & -4 \end{pmatrix}.$$

Entonces del enunciado

$$\Gamma = \frac{1}{5}A - \frac{e^{-4\pi}}{5}B.$$

Obsérvese que $A = B + 5I$. *Por tanto,*

$$\Gamma = \frac{1}{5}A - \frac{e^{-4\pi}}{5}(A - 5I) = \frac{1 - e^{-4\pi}}{5}A + e^{-4\pi}I.$$

Además,

$$\left(\frac{1}{5}A\right)^2 = \frac{1}{25}\begin{pmatrix} 10 & 10 & 5 \\ 10 & 10 & 5 \\ 10 & 10 & 5 \end{pmatrix} = \frac{5}{25}A = \frac{1}{5}A,$$

de modo que $P := \frac{1}{5}A$ *es idempotente y*

$$\Gamma = e^{-4\pi}I + (1 - e^{-4\pi})\,P.$$

Nuestro objetivo será $\exp(\mathbf{Q}_2) = \Gamma$.

Como $A = 5P$, *se tiene* $B = 5(P - I)$ *y*

$$\mathbf{Q}_2 = \frac{4\pi}{5}B = 4\pi(P - I) = -4\pi(I - P).$$

Notar que

$$(I - P)^2 = I - 2P + P^2 = I - 2P + P = I - P,$$

así que $S := I - P$ *es idempotente.*

> ***Fórmula para exponencial de idempotente.*** *Si* $S^2 = S$, *entonces* $\exp(\alpha S) = (I - S) + e^{\alpha}S = I + (e^{\alpha} - 1)S$.

Aplicando la fórmula con $S = I - P$ *y* $\alpha = -4\pi$*:*

$$\exp(\mathbf{Q}_2) = \exp\left(-4\pi(I-P)\right) = (I-(I-P))+e^{-4\pi}(I-P) = P+e^{-4\pi}(I-P) = e^{-4\pi}I+(1-e^{-4\pi})\,P = \Gamma.$$

$\Rightarrow \boxed{\exp(\mathbf{Q}_2) = \Gamma}$.

Parte A.2: usando $\mathbf{Q}_1$ y proyectores de Lagrange.

Escribamos

$$M := \frac{\mathbf{Q}_1}{2\pi} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & 0 & -2 \end{pmatrix}.$$

El polinomio característico de M *es* $\chi_M(z) = z(z^2 + 4z + 5)$, *por lo que*

$$z_0 = 0, \qquad z_{1,2} = -2 \pm i.$$

Entonces los autovalores de $\mathbf{Q}_1 = 2\pi M$ *son*

$$\lambda_0 = 0, \qquad \lambda_{1,2} = 2\pi(-2 \pm i) = -4\pi \pm 2\pi i.$$

Proyectores de Lagrange. *Para M diagonalizable y raíces z_i, los* polinomios base de Lagrange $\ell_i(z) = \prod_{j\neq i} \frac{z-z_j}{z_i-z_j}$ *inducen proyectores $L_i(M) = \ell_i(M)$ con $L_i^2 = L_i$, $L_i L_j = 0$ si $i \neq j$ y $\sum_i L_i = I$. Además, para funciones polinomiales/analíticas,*

$$f(M) = \sum_i f(z_i)\, L_i(M).$$

En particular, $e^M = \sum_i e^{z_i} L_i(M)$ *y* $e^{\mathbf{Q}_1} = \sum_i e^{\lambda_i} L_i(\mathbf{Q}_1)$.

Construimos $L_0(\mathbf{Q}_1)$ *(el proyector asociado a* $\lambda_0 = 0$*):*

$$L_0(\mathbf{Q}_1) = \prod_{j\neq 0} \frac{\mathbf{Q}_1 - \lambda_j I}{\lambda_0 - \lambda_j} = \frac{(\mathbf{Q}_1 - \lambda_1 I)(\mathbf{Q}_1 - \lambda_2 I)}{(0-\lambda_1)(0-\lambda_2)}.$$

Desarrollando el numerador:

$$(\mathbf{Q}_1 - \lambda_1 I)(\mathbf{Q}_1 - \lambda_2 I) = \mathbf{Q}_1^2 - (\lambda_1 + \lambda_2)\mathbf{Q}_1 + \lambda_1\lambda_2\, I = \mathbf{Q}_1^2 + 8\pi\, \mathbf{Q}_1 + 20\pi^2 I,$$

y el denominador es $(-\lambda_1)(-\lambda_2) = \lambda_1\lambda_2 = 20\pi^2$. *Por tanto,*

$$L_0(\mathbf{Q}_1) = \frac{\mathbf{Q}_1^2 + 8\pi\, \mathbf{Q}_1 + 20\pi^2 I}{20\pi^2}.$$

Usando $\mathbf{Q}_1 = 2\pi M$*, se obtiene*

$$L_0(\mathbf{Q}_1) = \frac{(2\pi)^2 M^2 + 8\pi(2\pi)M + 20\pi^2 I}{20\pi^2} = \frac{M^2 + 4M + 5I}{5}.$$

Pero

$$M^2 + 4M + 5I = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 2 & 1 \\ 2 & 2 & 1 \end{pmatrix} = A, \qquad \Rightarrow \qquad L_0(\mathbf{Q}_1) = \frac{1}{5}A = P.$$

Finalmente, como $\mathbf{Q}_1$ *es diagonalizable,*

$$e^{\mathbf{Q}_1} = e^{\lambda_0} L_0 + e^{\lambda_1} L_1 + e^{\lambda_2} L_2 = L_0 + e^{-4\pi+2\pi i} L_1 + e^{-4\pi-2\pi i} L_2.$$

Usando la periodicidad $e^{\pm 2\pi i} = 1$,

$$e^{\lambda_1} = e^{-4\pi}, \qquad e^{\lambda_2} = e^{-4\pi},$$

y

$$e^{\mathbf{Q}_1} = L_0 + e^{-4\pi}(L_1 + L_2) = L_0 + e^{-4\pi}(I - L_0) = e^{-4\pi} I + (1 - e^{-4\pi})\, L_0 = e^{-4\pi} I + (1 - e^{-4\pi})\, P = \Gamma.$$

$\Rightarrow \boxed{\exp(\mathbf{Q}_1) = \Gamma}$.

(b) *Aplicación del Teorema 3.1 de Israel* et al.

Caso 1: $\Gamma = \begin{pmatrix} 0.4 & 0.6 \\ 0.5 & 0.5 \end{pmatrix}$. *Aquí*

$$\det(\Gamma) = 0.4 \cdot 0.5 - 0.6 \cdot 0.5 = -0.1 \leq 0.$$

Por el punto (i) del teorema, no *existe generador exacto para* Γ.

Caso 2: $\Gamma = \begin{pmatrix} 0.9 & 0.1 & 0 \\ 0.1 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}$. *Se verifica que* $\Gamma^{(2)}_{13} \neq 0$ *(hay accesibilidad de* 1 *a* 3 *en dos pasos), mientras que* $\gamma_{13} = 0$*. Por el punto (iii) del teorema (accesibilidad con probabilidad de un paso nula),* no *existe generador exacto para* Γ.